

Centro Brasileiro de Pesquisas Físicas
Programa de Doutorado

Marco Antonio Guimarães Auad Barroca

Model, Sample, Solve — A Sampling-Based Quantum Computing Framework for Materials Science

Rio de Janeiro/RJ

2025

"MODEL, SAMPLE, SOLVE - A SAMPLING-BASED QUANTUM COMPUTING FRAMEWORK FOR MATERIALS SCIENCE"

MARCO ANTONIO GUIMARÃES AUAD BARROCA

Tese de Doutorado em Física apresentada no Centro Brasileiro de Pesquisas Físicas do Ministério da Ciência Tecnologia e Inovação. Fazendo parte da banca examinadora os seguintes professores:

Documento assinado digitalmente
ALEXANDRE MARTINS DE SOUZA
Data: 05/01/2026 15:59:44-0300
Verifique em https://validar.iti.gov.br

Alexandre Martins de Souza - Orientador/CBPF

Documento assinado digitalmente
IVAN DOS SANTOS OLIVEIRA JUNIOR
Data: 05/01/2026 15:26:18-0300
Verifique em https://validar.iti.gov.br

Ivan dos Santos Oliveira Junior – Coorientador/ CBPF

Documento assinado digitalmente
EDUARDO INACIO DUZZIONI
Data: 21/12/2025 15:51:44-0300
CPF: ***.513.789-**
Verifique as assinaturas em https://v.ufsc.br

Eduardo Inacio Duzzioni – UFSC

Documento assinado digitalmente
MARCELO SILVA SARANDY
Data: 22/12/2025 11:18:57-0300
Verifique em https://validar.iti.gov.br

Marcelo Sarandy - UFF

Documento assinado digitalmente
CELSO JORGE VILLAS BOAS
Data: 22/12/2025 13:16:19-0300
Verifique em https://validar.iti.gov.br

Celso Jorge Villas Boas - UFSCar

Documento assinado digitalmente
DIOGO DE OLIVEIRA SOARES PINTO
Data: 05/01/2026 14:10:11-0300
Verifique em https://validar.iti.gov.br

Diogo de Oliveira Soares Pinto – USP

Rio de Janeiro, 18 de dezembro de 2025.

Model, Sample, Solve — A Sampling-Based Quantum Computing Framework for Materials Science

Marco Antonio Guimarães Auad Barroca

Tese de Doutorado submetida a coordenação de formação científica do Centro Brasileiro de Pesquisas Físicas como parte dos requisitos para a obtenção do título de Doutor em Física.
Orientador: Alexandre Martins de Souza
Coorientador: Ivan dos Santos Oliveira Junior

Rio de Janeiro/RJ

2025

Abstract

This thesis investigates whether sampling-based quantum algorithms can deliver chemically meaningful predictions for realistic systems on pre–fault-tolerant hardware. We develop and evaluate a "Model, Sample, Solve" workflow: construct compact, material-informed Hamiltonians; generate shallow, physics-guided trial states to sample configuration support; and solve the resulting measurement-defined subspace classically. After establishing a variational control baseline (qubit-ADAPT-VQE) to quantify accuracy vs. depth/measurement cost, we pivot to Sample-Based Quantum Diagonalization (SQD) and its extended form (Ext-SQD), which decouple state preparation from fragile variational optimization and instead rely on configuration recovery, post-selecteed excitations, and subspace diagonalization.

We instantiate this program in two settings. (i) *Embedded surface reactions* for Li–O_2 interfaces: periodic Density Functional Theory (DFT) geometries feed density-difference analysis and coupled-cluster natural orbitals to define HONO/LUNO-centered active spaces; shallow Local Unitary Cluster Jastrow (LUCJ) trials enable SQD/Ext-SQD estimates of ground and reaction energies that track selected-Configuration Interaction (CI) references within chemical-accuracy tolerances across feasible sizes. (ii) *Periodic band edges* in compact oxides: KS-DFT is projected to tight-binding matrices and augmented with self-consistent on-site U and inter-site V to form Extended-Hubbard models. Direct gaps at Γ are computed from many-body ground states in metal oxides with 38-46 spin orbitals.

Beyond results, the thesis contributes standardized, data-driven interfaces (tight-binding / Extended-Hubbard instances and descriptors) that tie model building to quantum experiments with transparent resource reporting (qubits, depth, two-qubit counts). The overall conclusion is pragmatic: sampling-plus-diagonalization provides a credible near-term route to CI-grade benchmarks for reactivity and band edges on today's devices. Furthermoe, we outline a staged roadmap positioning these workflows for incremental growth while maintaining verification against classical baselines.

Key-words: Quantum Simulations. Materials Science. Quantum Chemistry.

Resumo

Esta tese investiga se algoritmos quânticos baseados em amostragem podem fornecer previsões quimicamente significativas para sistemas realistas em hardware pré-tolerante a falhas. Desenvolvemos e avaliamos um fluxo de trabalho "Modelar, Amostrar, Resolver": construir Hamiltonianos compactos informados pelo material; gerar estados-teste rasos, guiados pela física, para amostrar o suporte de configurações; e resolver classicamente o subespaço resultante definido por medições. Após estabelecer uma linha de base de controle variacional (qubit-ADAPT-VQE) para quantificar precisão vs. profundidade/custo de medição, migramos para a Diagonalização Quântica por Amostragem (Sample-Based Quantum Diagonalization, SQD) e sua forma estendida (Ext-SQD), que desacoplam a preparação do estado da frágil otimização variacional e, em vez disso, baseiam-se na recuperação de configurações, em excitações pós-selecionadas e na diagonalização em subespaços.

Instanciamos esse programa em dois cenários. (i) *Reações de superfície embutidas* em interfaces Li–O_2: geometrias periódicas de DFT alimentam análises de diferença de densidade e orbitais naturais de coupled-cluster para definir espaços ativos centrados em HONO/LUNO; ensaios rasos de Local Unitary Cluster Jastrow (LUCJ) permitem estimativas via SQD/Ext-SQD de energias do estado fundamental e de reação que acompanham referências de Interação de Configurações (CI) selecionada dentro de tolerâncias de exatidão química em tamanhos viáveis. (ii) *Bordas de banda periódicas* em óxidos compactos: KS-DFT é projetada em matrizes de tight-binding e aumentada com parâmetros U locais e V intersítio auto-consistentes para formar modelos de Hubbard Estendido. Lacunas diretas em Γ são calculadas a partir de estados fundamentais de muitos corpos em óxidos metálicos com 38–46 orbitais de spin.

Além dos resultados, a tese contribui com interfaces padronizadas e orientadas a dados (instâncias e descritores de tight-binding/Hubbard Estendido) que conectam a construção de modelos a experimentos quânticos com relato transparente de recursos (qubits, profundidade, contagem de portas de dois qubits). A conclusão geral é pragmática: amostragem + diagonalização fornece uma rota crível no curto prazo para benchmarks de nível CI de reatividade e bordas de banda nos dispositivos atuais. Além disso, delineamos um roteiro em etapas que posiciona esses fluxos de trabalho para crescimento incremental, mantendo a verificação em relação a linhas de base clássicas.

Palavras-chave: Simulações Quânticas. Ciência de Materiais. Química Quântica.

List of Figures

Figure 1 – **Example SWAP network for $n = 6$ qubits.** In $O(n)$ steps, each of the $O(n^2)$ pair of qubits (colors) performs an interaction, even with just linear connectivity. *Reproduced from Y. Aleexev, et al., [1] (DOI) © 2024 Elsevier. Used with permission.* 19

Figure 2 – **Two circuits implementing the same sequence $R_{\mathrm{YY}}(\theta_3) \cdot R_{\mathrm{XZ}}(\theta_2) \cdot R_{\mathrm{XX}}(\theta_1)$.** *Reproduced from Y. Aleexev, et al., [1] (DOI) © 2024 Elsevier. Used with permission.* 20

Figure 3 – **Schematic overview of representative use cases in materials discovery.** The central challenge in quantum theory is to determine the properties of molecules and solids (blue). This task can be addressed by solving the electronic Schrödinger equation to compute, for example, ground-state, excited-state, and non-equilibrium dynamical properties (vertical arrow). Because exact solutions are computationally prohibitive, approximations are introduced that reduce the quantum problem to fit within available computational budgets (red). Accurate solutions of the electronic Schrödinger equation serve as the foundation (green) for downstream applications in materials science, including vibrational structure calculations, coarse-grained modeling (e.g., corrosion or deformation), and materials discovery (e.g., catalysis, metamaterials). *Reproduced from Y. Aleexev, et al., [1] (DOI) © 2024 Elsevier. Used with permission.* 28

Figure 4 – Fe_2S_2 (a) and Fe_4S_4 (b) clusters in ferredoxins; the nitrogenase P-cluster in the resting state P^N (c); and the FeMo cofactor (d). Orange, yellow, teal, red, blue, gray, and white sticks denote Fe, S, Mo, O, N, C, and H, respectively. *Reproduced from Y. Aleexev, et al., [1] (DOI) © 2024 Elsevier. Used with permission.* 32

Figure 5 – Left: crystal structure of α-$RuCl_3$, showing the van der Waals gap between $RuCl_6$ honeycomb layers (Ru/Cl as gray/green) and a zigzag-ordered configuration (red arrows). Right: the 2D Heisenberg–Kitaev model; spins are white circles and $x/y/z$ bonds are green/blue/red. Current quantum simulations target 1–2 hexagons (6–10 spins) [2]. *Reproduced from Y. Aleexev, et al., [1] (DOI) © 2024 Elsevier. Used with permission.* 35

Figure 6 – (a) Schematic of a $[CuO_2]_4$ cell in a cuprate CuO_2 plane (Cu/O shown as orange/red; orbital phases in blue/green). (b) Parameters of the three-band Hubbard (Emery) model (curved connectors depict kinetic couplings).*Reproduced from Y. Aleexev, et al., [1] (DOI) © 2024 Elsevier. Used with permission.* 36
Figure 7 – **Estimated classically tractable sizes for Frenkel exciton Hamiltonians using hierarchical equations of motion on 64 AMD Opteron cores with 250 GB RAM.** Three photosynthetic complexes are indicated: Fenna–Matthews–Olson (Green Sulfur Bacteria), Light Harvesting I/II (Purple Bacteria), and Photosystem II (higher plants). Adapted from [3].*Reproduced from Y. Aleexev, et al., [1] (DOI) © 2024 Elsevier. Used with permission.* 40
Figure 8 – Schematic of a planar multilayer transparent radiative-cooling metamaterial design that admits a binary encoding of layer choices (adapted from [4]).*Reproduced from Y. Aleexev, et al., [1] (DOI) © 2024 Elsevier. Used with permission.* 42
Figure 9 – **Potential Energy Surface for LiH calculated by VQE UCCSD (blue), ADAPT-VQE UCCSD (green).** The result of the direct diagonalization of the fermionic Hamiltonian using NumPy (red) is shown as a target. The minimum energy point should tell us the equilibrium distance for the Li-H bond. We'll give an explanation of the methods used to calculate later on. 48
Figure 10 – **A 3-qubit circuit with eight θ parameters which parameterized the *Ry* rotations.** The circuit also includes CNOTs as the entangling operations. 73
Figure 11 – **Mg-MOF-74 structure:** (a) conventional unit cell (162 atoms) and (b) primitive unit cell (54 atoms). *Reproduced from* M.A. Barroca, et al., [5] (DOI) CC BY 4.0 (link). 80
Figure 12 – **Molecular geometries and bond axes used in the PES calculations for (a) H_2O, (b) N_2 and (c) CO_2 molecules in the vicinity of a metal ion M.** For H_2O, the ion is moved along the C_2 rotation axis. For N_2, the ion is moved along the N-N axis. For CO_2, the ion is moved along the O-C-O axis. *Reproduced from* M.A. Barroca, et al., [5] (DOI) CC BY 4.0 (link). 82

Figure 13 – **Convergence to the ground-state energy from quantum simulations using UCCSD-VQE (green), $R_Y R_Z$-VQE (red), and qubit-ADAPT-VQE (blue) as a function of the total number of VQE iterations.** The black line shows the target from direct diagonalization of the fermionic Hamiltonian in NumPy (FCI-equivalent). All ansätze converge to within chemical accuracy ($\sim$ 1 mHa) of the target. Notably, qubit-ADAPT-VQE reaches this tolerance in far fewer iterations than UCCSD-VQE and $R_Y R_Z$-VQE. Energies correspond to the $(6, 7)$ active-space calculation. *Reproduced from* M.A. Barroca, et al., [5] (DOI) CC BY 4.0 (link). 84

Figure 14 – **Potential energy surfaces and energy differences with respect to direct diagonalization with NumPy, for (a) $Mg^{+2} + H_2O$, (b) $Mg^{+2} + N_2$ and (c) $Mg^{+2} + CO_2$ systems.** The black dashed line in the ΔE plot represents the chemical accuracy threshold. Noiseless quantum simulations of UCCSD-VQE (green circles) and qubit-ADAPT-VQE (blue circles) match the NumPy (magenta triangles) benchmark in all three cases, however, $R_Y R_Z$-VQE (red circles) only matches the benchmark for the simplest system, namely, $Mg^{+2} + H_2O$. The black dashed line in the ΔE plot represents the chemical accuracy threshold. *Reproduced from* M.A. Barroca, et al., [5] (DOI) CC BY 4.0 (link). . . 86

Figure 15 – **Potential Energy Surfaces and energy differences with respect to direct diagonalization with NumPy (purple triangles) for VQE quantum experiments on hardware (black diamonds) and qubit-ADAPT-VQE simulation (blue circles).** The black dashed line in the ΔE plot represents the chemical accuracy threshold. The adjusted Morse potentials exhibit a considerable overlap, despite the level of noise present in the hardware data points, which leads to a reasonable agreement between the observables shown in Table 4. *Reproduced from* M.A. Barroca, et al., [5] (DOI) CC BY 4.0 (link). . . 87

Figure 16 – **Potential energy surfaces and energy differences with respect to direct diagonalization with NumPy, for (a) $Mg^{+2} + H_2O$, (b) $Mg^{+2} + N_2$ and (c) $Mg^{+2} + CO_2$ systems.** The black dashed line in the ΔE plot represents the chemical accuracy threshold. Results obtained from quantum experiments on hardware using (3,4) active space (black diamonds) and quantum simulations in a (6,7) active space (blue circles) compared to PySCF-based, all-electron classical simulations using the HF (red squares), DFT (green squares) and NumPy direct diagonalization (purple triangles). 88

Figure 17 – **VQE PES for Zn^{2+} adducts: VQE_UCCSD (green), qubit-ADAPT-VQE (blue), and NumPy exact diagonalization (magenta) agree across all ligands**; VQE_RyRz (red) departs from the benchmark for N_2. 91
Figure 18 – **qubit-ADAPT-VQE (blue) compared to PySCF HF (red) and DFT (green) for Zn^{2+} adducts.** HF and qubit-ADAPT-VQE agree closely; DFT predicts slightly shorter r_e, especially for N_2. 91
Figure 19 – **Hybrid quantum–classical framework for studying lithium–oxygen reactions.** Geometry optimization and DFT calculations lead into active-space selection via DDA and natural orbitals, identifying key orbitals near the reactive site. A truncated LUCJ ansatz is sampled on quantum hardware (or compressed simulators), and SQD/Ext-SQD compute reaction energies ($\Delta E = E_{\text{prod}} - E_{\text{reac}}$), providing insights into lithium oxide formation. *Reproduced from* M.A. Barroca, et al., [6] (DOI) CC BY 4.0 (link). 129
Figure 20 – **Active Space selection** We start with geometry optimization (a). Move to the Density difference analysis (b) and evaluate each orbital individual contribution to the density difference. Finally, we calculate the active space using the Natural Orbitals (c), where orbitals for the product are pictured. They were sorted based on their contribution to the Density difference. Natural Orbitals are calculated and the active space grows "inside out" from the HONO-LUNO level. *Reproduced from* M.A. Barroca, et al., [6] (DOI) CC BY 4.0 (link). 131
Figure 21 – **Comparison of ground state and reaction energy results for classical methods.** Classical simulations of the ground state energies as a function of active space size. The left curves are for the reactant and the right ones are for the product. *Reproduced from* M.A. Barroca, et al., [6] (DOI) CC BY 4.0 (link). 132
Figure 22 – **Circuit-to-chip mapping.** Green qubits represent α spin orbitals, red qubits represent β spin orbitals, while blue qubits act as ancillae. Numbers inside the qubits represent their index with respect to the device ordering. *Reproduced from* M.A. Barroca, et al., [6] (DOI) CC BY 4.0 (link). 133

Figure 23 – **Histograms of electron configurations sampled from the quantum hardware for the (14e,20o) reactant active space.** We display only the top-10 configurations. Filled (unfilled) circles represent occupied (unoccupied) orbitals. (a) Histogram for the non-truncated one-layer LUCJ ansatz. (b) Histogram for the truncated two-layer LUCJ ansatz. We observe that the truncated variant leads to a more relevant distribution of electronic configurations, that is not dominated by the reference Hartree–Fock state. *Reproduced from* M.A. Barroca, et al., [6] (DOI) CC BY 4.0 (link). 134

Figure 24 – **Comparison of ground state and reaction energy results for SQD methods.** SQD calculations of the ground state energies as a function of active space size. The left curves are for the reactant and the right ones are for the product. *Reproduced from* M.A. Barroca, et al., [6] (DOI) CC BY 4.0 (link). 135

Figure 25 – **Comparison of ground state and reaction energy results for quantum experiments** (a) Simulated (yellow diamonds) and experimental (green diamonds) Ext-SQD results compared to Ext-HCI, the best classical method considered. Experimental Ext-SQD scales up to 32 orbitals, only. (b) Reaction energies for quantum and classical methods computed at the Γ-point. DFT results are shown as a reference but are known to be overestimated. *Reproduced from* M.A. Barroca, et al., [6] (DOI) CC BY 4.0 (link). 136

Figure 26 – **DFT→lattice workflow.** KS densities are projected to a localized AO basis to yield TB hoppings t_{pq}; intra-site U_p and inter-site V_{pq} are obtained self-consistently within DFT+$U+V$. Each AO becomes a lattice site; orbitals on the same atom map to different sites. The final model is the Extended-Hubbard Hamiltonian in Equation 6.1. *Reproduced from* Duriez et al. [7] (DOI), CC BY 4.0 (link). 144

Figure 27 – Validation of the TB projection and resulting sparsity. **Left:** DFT (dashed) vs. PAOFLOW TB (solid) bandstructures for Si, ZrO_2, and HfO_2. **Right:** False-color $|t_{ij}|^2$ at Γ showing block structure (intra-site vs. inter-site) and dominant near-neighbor couplings. *Reproduced from* Duriez et al. [7] (DOI), CC BY 4.0 (link). 145

Figure 28 – **Hubbard parameters and band–gap shifts in the TBHubbard EH subset. (a–b)** Distributions of intra–site U and inter–site V across metals for d–p and d–s manifolds. **(c–d)** Scatter of E_g^{DFT} vs. $E_g^{\mathrm{DFT+U+V}}$ for d–p and d–s perturbations, respectively. Many Zr/Hf frameworks lie near the diagonal (small corrections), whereas late–row d systems show larger positive shifts, consistent with correlation–driven gap opening. Reproduced from Carvalho et al. [8] (DOI), CC BY 4.0 (link). 147

Figure 29 – **Dataset at a glance. (a)** Coverage of QMOF and its TB (green) and EH (blue) subsets. **(b)** t-SNE of TB matrices indicates distinct clustering among source databases and subsets. **(c)** t-SNE of SOAP–3 Å metal–center descriptors (after PCA) highlights chemically coherent neighborhoods—useful for model transfer and active learning. *Reproduced from* Carvalho et al. [8] (DOI), CC BY 4.0 (link). 148

Figure 30 – **Representative tight–binding matrices at Γ for two MOFs.** False–color plots of $|t_{ij}|^2$ reveal strong intra–site blocks and chemically interpretable inter–site couplings (e.g., M–N, C–N, C–H). This sparsity/locality is exactly what near–term quantum solvers exploit when defining active spaces and excitation pools. *Reproduced from* Carvalho et al. [8] (DOI), CC BY 4.0 (link). 149

Figure 31 – **Direct band gaps at Γ for HfO_2 and ZrO_2 from DFT and from lattice Hamiltonians under different interaction settings (TB, U, V, $U+V$).** Lattice gaps are computed from ground-state energies in $N_e \pm 1$ and N_e sectors using HCI. GW and experimental references from [9, 10]. *Reproduced from* Duriez et al. [7] (DOI), CC BY 4.0 (link). 151

Figure 32 – **SQD/Ext-SQD on superconducting hardware for (a) HfO_2 (38 qubits) and (b) ZrO_2 (46 qubits). Errors vs. HCI (reference) as a function of CI-subspace fraction for N_e-1, N_e, and N_e+1 sectors.** Ext-SQD attains HCI-level accuracy with fewer configurations and surpasses CCSD(T). *Reproduced from* Duriez et al. [7] (DOI), CC BY 4.0 (link). 152

Figure 33 – **HCI convergence of ground-state energies for HfO_2 and ZrO_2 in the N_e (a), N_e-1 (b), and N_e+1 (c) electron-number sectors.** For HfO_2, we plot the deviation from FCI (exact) energies; for ZrO_2, we report absolute energies. The horizontal axis shows the fraction of the total configuration space included in the CI diagonalization (as in Fig. 32). Color bars encode the ground-state variance, $\mathrm{Var}(H) = (\langle H^2\rangle - \langle H\rangle^2)/\langle H\rangle^2$, with smaller values indicating higher-fidelity eigenstates. *Reproduced from* Duriez et al. [7] (DOI), CC BY 4.0 (link). 153

List of Tables

Table 1 – Qualitative comparison of fermion–to–qubit mappings (for K spin–orbitals). "Update" refers to the number of qubits affected when flipping an occupation; "Parity" refers to the number needed to compute fermionic parity. Shorter support generally yields shorter Pauli strings and shallower two-qubit ladders. 72

Table 2 – The qubit-ADAPT-VQE ansatz converges in 3x and 116x fewer iterations, and is 5x and 25x faster than the UCCSD-VQE and $R_Y R_Z$-VQE ansatze, respectively. The number of iterations for qubit-ADAPT-VQE is the sum of all VQE iterations. *Reproduced from* M.A. Barroca, et al., [5] (DOI) CC BY 4.0 (link). 83

Table 3 – **Comparing the circuit sizes generated by different ansätze.** qubit-ADAPT-VQE manages to get satisfactory results while using the shallowest circuits. Circuits are described in terms of the number of gate operations and depth required to run on IBM quantum hardware. *Reproduced from* M.A. Barroca, et al., [5] (DOI) CC BY 4.0 (link). . . . 84

Table 4 – **Comparison of Equilibrium Distance and Binding Energy for $Mg^{+2} + H_2O$ obtained from the fitted Morse potential to results obtained with different methods.** The uncertainties represent the error bars of the curve fit parameters. Taking the fit uncertainties and the inherent noise of the quantum hardware data points, we consider all methods to be in agreement. *Reproduced from* M.A. Barroca, et al., [5] (DOI) CC BY 4.0 (link). 88

Table 5 – **Comparison of Equilibrium Distances and Binding Energies for Mg^{+2}-based systems obtained from the fitted Morse potential across different methods, both quantum (simulations and hardware experiments) and classical.** The Binding Energy values all agree to within the numeric uncertainty. The equilibrium distances agree between qubit-ADAPT-VQE and HF, but are 2-3% higher than what is predicted by DFT. *Reproduced from* M.A. Barroca, et al., [5] (DOI) CC BY 4.0 (link). 89

Table 6 – **Morse-fit equilibrium distance r_e and binding energy D_e for $Zn^{2+} + H_2O$ from the three VQE ansätze versus NumPy.** Uncertainties reflect fit residuals. 91

Table 7 – **Morse-fit equilibrium distances r_e and binding energies D_e for Zn^{2+} adducts from qubit-ADAPT-VQE and PySCF HF/DFT.** . 92

Table 8 – **Sampling–based quantum chemistry and many–body case studies on hardware.** . . . 121
Table 9 – **Quantum metrics for select circuits.** q is the total number of qubits used, which is the sum of s qubits representing spin orbitals and a ancilla qubits. d is the 2-qubit circuit depth, CX is the number of CNOT gates and u the number of 1-qubit gates. Notice the small addition of two electrons in the product's active space makes circuits slightly larger after we include all occupied orbitals. *Reproduced from* M.A. Barroca, et al., [6] (DOI) CC BY 4.0 (link). . . . 133
Table 10 – **Select Ext-SQD and Ext-HCI calculations for product and reactant.** $|\tilde{\chi}|$ is the number of samples taken from **ibm_aachen**, K is the number of configuration recovery iterations used in SQD, D is the dimension of the subspace being diagonalized by SQD, D_E is the dimension of the subspace being diagonalized by Ext-SQD, while D_{EH} is the dimension of the subspace being diagonalized by Ext-HCI. *Reproduced from* M.A. Barroca, et al., [6] (DOI) CC BY 4.0 (link). . . . 138
Table 11 – **Reaction Energies.** ΔE (in eV) compared across methods. *Reproduced from* M.A. Barroca, et al., [6] (DOI) CC BY 4.0 (link). . . . 138
Table 12 – **Direct gap of cubic ZrO_2 at Γ:** literature vs. this work (EH value corresponds to the $U{+}V$ point in Fig. 31). *Reproduced from* Duriez et al. [7] (DOI), CC BY 4.0 (link). . . . 150
Table 13 – **Direct band gaps at Γ (E_g^{Γ}, eV) for ZrO_2, HfO_2, and Si.** TB and Extended-Hubbard (EHM) values are from HCI on the projected lattice Hamiltonians; literature values are shown for context. *Reproduced from* Duriez et al. [7] (DOI), CC BY 4.0 (link). . . . 154

Contents

1 INTRODUCTION . . . 1
1.1 A Pragmatic Paradigm: Quantum-Centric Workflows . . . 3
1.2 Thesis Goals and Research Questions . . . 4
1.3 Summary of Contributions . . . 5
1.4 Thesis Structure . . . 5
1.5 Author Contributions and Publication Disclosure . . . 6

2 QUANTUM COMPUTING IN MATERIALS SCIENCE: CHALLENGES AND OUTLOOK . . . 8
2.1 Introduction . . . 8
2.2 Key Quantum Algorithms for Materials-Science Workflows . . . 9
2.3 Classical Information Processing . . . 11
2.3.1 Mapping to qubits . . . 11
2.3.2 Hamiltonian encodings for dynamics simulation . . . 13
2.3.3 Quantum measurement . . . 15
2.3.4 Problem-level preprocessing and optimization . . . 16
2.3.5 Circuits, transpilation, and architecture . . . 19
2.3.6 Error mitigation . . . 21
2.4 Classical Simulation of Quantum Systems . . . 23
2.4.1 Why use classical simulators . . . 23
2.4.2 Quantum circuit simulators . . . 24
2.4.2.1 Density matrix simulators . . . 25
2.4.2.2 Tensor network simulators . . . 25
2.4.2.3 Open-system Lindblad simulators . . . 26
2.5 Use-Case Identification . . . 26
2.5.1 Criteria for selecting relevant use cases . . . 26
2.5.2 Overview of use cases . . . 28
2.5.3 Relevant use cases . . . 30
2.5.3.1 Ground-state electronic structure . . . 30
2.5.3.2 Embedding . . . 33
2.5.3.3 Model systems . . . 34
2.5.4 Excited states . . . 37
2.5.5 Applications of electronic structure . . . 38
2.6 Outlook . . . 42

3 METHODS FOR QUANTUM SIMULATION OF FERMIONIC SYSTEMS . . . 44
3.1 Classical Computational Chemistry Methods . . . 44
3.1.1 Second-quantized electronic Hamiltonian . . . 44
3.1.2 The Born-Oppenheimer Approximation . . . 46
3.1.3 The variational principle . . . 47
3.1.4 Slater Determinants and the Fock operator . . . 49
3.1.5 The Hartree-Fock Method and the Self-Consistent Field Algorithm . . . 51
3.1.6 Density Functional Theory . . . 55
3.1.7 Basis sets, frozen cores, and effective core potentials . . . 56
3.1.8 Full Configuration Interaction and Truncated CI Methods . . . 59
3.1.9 Coupled Cluster Theory: CCSD and CCSD(T) . . . 64
3.2 Quantum Computational Chemistry Methods . . . 69
3.2.1 The VQE algorithm . . . 69
3.2.1.1 Problem mapping . . . 70
3.2.1.2 The ansatz . . . 73
3.2.1.3 Measurement cost and optimization pathologies . . . 74
3.2.2 Reducing the problem . . . 76
3.2.3 VQE variations . . . 77
3.2.3.1 ADAPT-VQE . . . 77
3.2.3.2 qubit-ADAPT-VQE . . . 78
3.3 Adapted Article: Exploring qubit-ADAPT-VQE for materials discovery 79
3.3.1 Problem Statement . . . 79
3.3.2 Methods . . . 81
3.3.3 Results and discussion . . . 83
3.3.4 Conclusion . . . 89
3.4 Post-publishing extension: Zn^{2+} adsorption complexes . . . 90
3.5 Summary and motivation for sampling-based methods . . . 92

4 SAMPLE-BASED DIAGONALIZATION: CONCEPTS, METHODS, AND APPLICATIONS . . . 94
4.1 Motivation and problem setting . . . 94
4.2 Core algorithms . . . 97
4.2.1 Quantum Selected Configuration Interaction (QSCI) . . . 97
4.2.2 Sample-based Quantum Diagonalization (SQD) . . . 101
4.2.3 Extended SQD (Ext-SQD) for Ground and Excited States . . . 104
4.2.4 Sample-Based Krylov Quantum Diagonalization (SKQD) . . . 107
4.3 Reducing required qubit count and scaling-out . . . 113
4.4 Hardware and chemistry case studies . . . 117
4.5 Practical guidance and pitfalls . . . 121

4.6 Summary . . . 124

5 QUANTUM SIMULATION OF BATTERY ELECTRODE REACTIONS 126

5.1 Adapted Article: Scaling active spaces in simulations of surface reactions through sample-based quantum diagonalization . . . 126

5.1.1 Systems and reaction workflow . . . 127

5.1.2 Methods and Results . . . 129

5.1.3 Discussion and Summary . . . 137

6 BAND GAP ESTIMATION IN PERIODIC MATERIALS VIA SAMPLE-BASED METHODS . . . 141

6.1 Adapted Articles: Band–Gap Prediction via Sample-Based Quantum Diagonalization and lattice modeling of metal-organic frameworks for quantum computers . . . 141

6.1.1 Motivation . . . 142

6.1.2 From DFT to a Material-Informed Lattice Hamiltonian for Crystals . . . 143

6.1.3 The TBHubbard database . . . 144

6.1.4 From Descriptors to Hamiltonians: How TBHubbard Feeds Band–Gap Prediction and Quantum Workflows . . . 146

6.1.5 Results . . . 150

6.1.6 Discussion . . . 153

7 CONCLUSIONS AND FUTURE DIRECTIONS . . . 155

7.1 Consolidated Answers to Goals & Questions . . . 155

7.2 Consolidated Contributions . . . 156

7.3 Review . . . 157

7.4 Implications and future directions . . . 160

BIBLIOGRAPHY . . . 162

1 Introduction

The prediction and design of novel catalysts, next-generation battery technologies, and quantum materials represent grand challenges for the physical sciences. At their core, these pursuits hinge on a deceptively simple question: how do electrons behave in realistic systems? Answering this question is a highly difficult quantum mechanical problem, and our ability to solve it is a direct gatekeeper to transformative technological progress. In energy storage, for example, the performance of lithium-ion or next-generation lithium-air batteries is dictated by complex electrochemical reactions at the solid-electrolyte interphase (SEI) [11–14]. Understanding and controlling the formation of this layer, or preventing the growth of lithium dendrites that lead to battery failure, requires a predictive model of electron transfer and chemical bonding at a level of quantum detail that is currently inaccessible [14, 15]. Similarly, in catalysis, the efficiency of processes like nitrogen fixation or CO_2 reduction depends on the subtle electronic structure of transition-metal active sites within complex molecular or crystalline environments. The d- and f-block elements at the heart of these catalysts exhibit strong electron correlation, a phenomenon where the motion of each electron is intricately linked to many others, which defies simple descriptions [16].

This fundamental challenge is so central to modern science and engineering that a substantial fraction of supercomputing resources is already dedicated to materials simulations [1, 17]. Even with this immense classical power, a scaling wall persists. The "gold standard" of quantum chemistry, the full configuration interaction (FCI) method, provides a formally exact solution to the electronic Schrödinger equation within a given basis set. However, its computational cost grows factorially with the number of electrons and orbitals, rendering it intractable for all but the smallest benchmark systems [18, 19]. Its existence serves as a crucial theoretical reference point, but offers no practical path forward for the complex materials that drive modern technology.

To circumvent this dire scaling, a hierarchy of approximation methods has been developed. Density Functional Theory (DFT) has become the workhorse of the field, offering a remarkable balance of cost and accuracy for a wide range of problems, particularly in ground-state predictions of weakly correlated systems [20, 21]. Its success is predicated on mapping the interacting many-electron problem onto a simpler system of non-interacting electrons moving in an effective potential. However, the exact form of this potential (the exchange-correlation functional) is unknown and must be approximated [22, 23]. This leads to well-known failures in systems where electron correlation is strong, such as in the aforementioned transition-metal catalysts, or where long-range interactions like van der Waals forces are dominant [24, 25]. Furthermore, standard DFT approximations suffer from self-interaction error, leading to incorrect predictions for charge localization and band gaps in many materials [26, 27]. On the other hand, wave-function-based methods from chemistry, such as coupled-cluster theory with single, double, and perturbative triple excitations (CCSD(T)), can achieve "chemical accuracy" but at a steep polynomial cost (scaling as $\mathcal{O}(N^7)$ for CCSD(T)). As the method is not strictly variational, its high accuracy comes without a guaranteed lower bound, meaning the calculated energy can fall below the true value, further complicating its application profile, which is already limited to relatively small systems [28, 29]. The result is a vast and critical space of scientific

problems, complex, strongly correlated, and technologically vital, that remain beyond the reach of reliable classical simulation.

Quantum computing has emerged as a promising paradigm to dismantle this wall. By encoding quantum information in qubits and performing operations using the principles of quantum mechanics, these devices promise to simulate quantum systems with an efficiency that classical computers never could. The initial excitement in chemistry and materials science was fueled by the Variational Quantum Eigensolver (VQE), a hybrid quantum-classical algorithm that seemed perfectly suited for noisy intermediate-scale quantum (NISQ) hardware [30–33]. VQE translates the problem of finding a system's ground-state energy into an optimization task. A parameterized quantum circuit, or "ansatz," is used to prepare a trial wave function on the quantum computer. Its energy is measured, and a classical optimizer updates the circuit parameters to minimize this energy, guided by the variational principle [34].

However, as the field has pushed VQE towards systems of genuine scientific interest, its profound limitations have become apparent. The optimization process itself is fraught with challenges. For many problems of interest, the optimization landscape is riddled with local minima, and the classical optimizers struggle to find the true ground state [35]. More fundamentally, the field has had to confront the "barren plateau" phenomenon, a condition where the gradient of the cost function vanishes exponentially with increasing system size for many common ansatz choices [36, 37]. This makes training the quantum circuit impossible for all but the smallest systems. Compounding these issues are the immense measurement overheads required to obtain a statistically significant energy estimate and the fragility of the algorithm in the presence of hardware noise [38–41].

The limitations of VQE do not signal a failure of the quantum paradigm, but rather the need for a more robust class of algorithms. In response, the field is witnessing a critical shift toward new techniques that are less reliant on fragile variational optimization loops. This thesis focuses on one such family: **sample-based quantum algorithms** [6, 7, 42–47]. These methods, including Sample-Based Quantum Diagonalization (SQD) and its variants, pivot the computational paradigm. Instead of optimizing a complex parameterized circuit, they use the quantum computer to perform statistical sampling of the many-body configuration space. From these samples, a compact, effective Hamiltonian can be constructed in a smaller, more relevant subspace and then diagonalized classically. This approach is inherently more robust to certain types of noise and, by avoiding the optimization landscape entirely, sidesteps the barren plateau problem.

These methods are part of a broader evolution in quantum algorithm design. Another promising family of near-term techniques is quantum Krylov methods. These algorithms use short-time quantum evolution to construct a basis for a Krylov subspace, which provides a compact representation of the Hamiltonian's action on an initial state. Like sample-based methods, they then solve a small eigenvalue problem classically, offering provable convergence guarantees under certain conditions [48]. Both sample-based and Krylov methods represent a move toward leveraging the unique strengths of quantum computers (sampling, time evolution) while offloading the difficult optimization task to classical machines. Looking further ahead to the fault-tolerant era, algorithms like Quantum Phase Estimation (QPE) promise an exponential speedup in finding eigenvalues. However, QPE requires long, coherent quantum circuits and access to error correction, placing it beyond the reach of current and near-term hardware [49–51]. The work in

this thesis is therefore positioned at the practical forefront of this algorithmic evolution, validating the power of sample-based methods as a viable and scalable approach for today's scientific challenges.

1.1 A Pragmatic Paradigm: Quantum-Centric Workflows

The practical integration of these advanced quantum algorithms into scientific discovery workflows requires a pragmatic vision. The prevailing model is not of quantum computers as standalone replacements, but as specialized quantum co-processors embedded within a larger, classical high-performance computing (HPC) framework. This **quantum-centric supercomputing (QCSC)** ecosystem acknowledges that quantum computers will, for the foreseeable future, be a scarce and specialized resource best suited for specific tasks [1]. This hybrid approach is driven by necessity. A typical quantum simulation workflow involves a tight interplay between classical and quantum resources. The classical computer handles a host of critical tasks: pre-processing (e.g., performing initial DFT calculations to define an active space, optimizing molecular geometries, generating the Hamiltonian), concurrent processing (e.g., managing the feedback loop in adaptive algorithms, processing measurement outcomes for error mitigation schemes), and post-processing (e.g., analyzing results, calculating derivative properties).

This hybrid perspective naturally motivates the use of fragment-based and embedding methods. These powerful techniques provide a formal bridge between quantum and classical. The core idea is to partition a large, intractable problem into a smaller, computationally difficult part and a larger, more manageable environment. For instance, in simulating a reaction on a catalyst's surface, the active site and the reacting molecule can be defined as the "active space" or "fragment" to be treated with a high-level quantum method. The surrounding catalyst support and solvent, which interact with the active site primarily through electrostatic and steric effects, can be treated with a more efficient classical method like DFT or even a molecular mechanics force field. This allows the precious resources of the quantum computer to be focused exclusively on the region where strong electron correlation and quantum effects are most critical. The success of the entire simulation then hinges on the fidelity of the "quantum-classical" interface and the methods used to couple the two parts [52, 53].

At the technical level, quantum applications within this paradigm are organized around two primary pursuits:

Quantum simulation aims to solve the Schrödinger equation for these carefully chosen quantum fragments. As discussed, the central task is often finding the ground-state energy, a problem known to be computationally hard in the worst case (belonging to the Quantum Merlin-Arthur-hard (QMA-hard) complexity class [54–56]). As the field moves beyond the initial explorations with VQE, the new benchmark for progress is the ability of methods like sample-based diagonalization and quantum Krylov techniques to deliver scientifically relevant results such as reaction energy barriers and electronic band gaps for realistic, embedded systems. The goal is to achieve "chemical accuracy" (errors below 1.6 mHa or 1 kcal/mol), the threshold at which computational predictions become truly useful for guiding experimental efforts.

Quantum optimization, in contrast, addresses problems that can be framed as a discrete search over binary variables. This is crucial for accelerating other parts of the scientific workflow that lie beyond direct simulation. For example, discovering new materials involves searching vast chemical spaces, a combinatorial problem that can be mapped to a Quadratic Unconstrained Binary Optimization (QUBO) model [57–59]. Similarly, optimizing an experimental synthesis protocol or finding the ideal placement of sensors can be framed as QUBOs. Here, algorithms like the Quantum Approximate Optimization Algorithm (QAOA) or quantum annealing are employed [60–62]. While algorithmically distinct from quantum simulation, quantum optimization represents a complementary avenue for achieving quantum advantage in science, tackling the higher-level logic of discovery and design.

This thesis is situated squarely within this evolving, pragmatic paradigm. Our focus is the deployment and validation of sample-based quantum simulation methods for challenging materials problems that are intractable for classical computers. The following sections will detail the specific research questions that guide this work, summarize the contributions made, and outline the structure of the thesis.

1.2 Thesis Goals and Research Questions

Our stated goal is to investigate how quantum computing methods can be applied to realistic scientific problems, with a particular focus on the transition from variational to sample-based techniques as system complexity increases in realistic scenarios. While most of the thesis focuses on quantum simulation for chemically relevant systems.

We are guided by the following research questions:

- **RQ1** - What are the limitations of variational quantum algorithms for modeling correlated systems, and how do these limitations motivate the transition to sample-based methods for more realistic simulations?
- **RQ2** - How can sample-based quantum diagonalization methods be adapted to study chemically embedded fragments and periodic Hamiltonians with realistic system sizes and parameter regimes?
- **RQ3** - Can these approaches achieve chemically meaningful accuracy when modeling surface reactions (e.g., in lithium–oxygen batteries) or estimating electronic band gaps (e.g., in metal–organic frameworks)?
- **RQ4** - What challenges arise when integrating classical embedding techniques with quantum sampling methods, and how do design choices in problem representation affect performance?
- **RQ5** - To what extent can variational and sampling-based algorithms be applied to simplified systems to benchmark resource requirements and algorithmic stability?

These can be consolidated into one central question: ***Can shallow, sampling-driven quantum workflows deliver chemically meaningful predictions?***

1.3 Summary of Contributions

Our focus is the application of existing quantum computing methods, particularly sample-based diagonalization and variational algorithms, to chemically and physically realistic systems. While the methods themselves are established in the literature, their integration into quantum-classical workflows for materials simulation represents a meaningful step toward practical applications. The primary contributions are:

- **Early-Stage Variational Benchmarks:** Evaluated variational quantum algorithms on simplified test systems to identify resource bottlenecks and convergence issues. These findings motivated the shift toward sample-based methods for realistic simulations.
- **Method Integration and Deployment:** Applied and extended SQD methods to realistic fermionic systems by combining them with classical embedding techniques. Demonstrated how these methods scale to complex systems of practical relevance.
- **Embedded Li–O_2 Surface Reactions.** Engineered a quantum–classical workflow for oxygen-reduction steps at Li metal surfaces: periodic DFT geometries feed density-difference analysis (DDA) and CC natural-orbital refinement to define HONO–LUNO–centered active spaces. SQD/Ext-SQD yielded ground-state and reaction energies, up to 32 orbitals, that track classical benchmarks within chemical accuracy and provide variationally consistent improvements. Tensor-network sampling extended the same circuits to 40 orbitals, quantifying accuracy–cost trade-offs and hardware limits.
- **SQD for Periodic Hamiltonians (Oxide Band Gaps).** Adapted SQD/Ext-SQD to periodic, material-informed Extended-Hubbard Hamiltonians derived from DFT→lattice projections, and computed direct band gaps at Γ for compact oxides (ZrO_2, HfO_2) on pre–fault-tolerant hardware (38–46 qubits), with results consistent with higher-level classical references.
- **TBHubbard Dataset for Transferable Lattice Models.** Helped curate and analyze a CC-licensed resource of tight-binding (TB) models and self-consistent U, V parameters for MOFs, along with aligned descriptors. TBHubbard provides standardized EH instances and descriptors that could inform active-space selection, parameter trends, and benchmarking for periodic SQD workflows today and pave a data-ready path to larger framework targets as hardware scales.

1.4 Thesis Structure

This document is organized into seven chapters, progressing from high-level perspectives and theoretical foundations to methodological application and optimization strategies. The structure follows a natural trajectory: beginning with a strategic outlook on the role of quantum computing in materials science, then exploring variational algorithms and their limitations, and culminating in the deployment of sample-based quantum

simulation methods as a more scalable alternative. A final chapter presents a parallel investigation into quantum optimization for structured problems. Each chapter is briefly described below:

- **Chapter 2 – Quantum Computing in Materials Science: Challenges and Outlook** Provides a strategic overview of how quantum computing fits into materials discovery workflows. This chapter, based on a co-authored perspective article, outlines the scientific bottlenecks and emerging opportunities for quantum-enhanced computation in real-world materials problems.
- **Chapter 3 – Methods for Quantum Simulation of Fermionic Systems** Reviews essential background in second quantization, qubit mappings, and variational quantum algorithms. Classical techniques and the motivation for sample-based methods such as Sample-Based Quantum Diagonalization, are introduced as well as published results for early variational method testing on small molecules based on original work authored by the candidate.
- **Chapter 4 – Sample-based diagonalization: concepts, methods, and applications** Presents the theoretical formulation and implementation of Sample Based Quantum Diagonalization and its variants. The chapter highlights recently published work, discusses their results, and algorithmic performance of various methods.
- **Chapter 5 – Quantum Simulation of Battery Electrode Reactions** Applies the Extended Sample Based Quantum Diagonalization framework to simulate surface chemical reactions on lithium–oxygen battery interfaces using embedded Hamiltonians. Results demonstrate the practical applicability of SQD to real materials. This chapter is based on work where the candidate is first author.
- **Chapter 6 – Band Gap Estimation in Periodic Materials via Sample-Based Methods** Extends the sampling paradigm to crystalline solids, focusing on cubic ZrO_2 and HfO_2. The chapter reports resource usage for quantum experiments, clarifies how TBHubbard serves as a standardized source of TB/EH models and descriptors for future scaling, and is based on collaborative work.
- **Chapter 7 – Conclusions and Future Directions** Summarizes the thesis findings and reflects on the broader impact of the work. Key directions for future research in quantum chemistry and materials simulation.

1.5 Author Contributions and Publication Disclosure

This thesis contains material adapted from peer-reviewed publications and preprints authored or co-authored by the candidate. Where applicable, these materials have been reformatted to fit the thesis structure, with appropriate attribution. Below is an itemized summary of the candidate's contributions to each publication.

- **Future Generation Computer Systems, Volume 160, 2024, Pages 666-710 – "Quantum-centric Supercomputing for Materials Science: A Perspective on Challenges and Future Directions"** (Co-author)

  - Contributed to conceptual framing and drafting of the manuscript.
  - Authored sections related to quantum simulation workflows and hybrid strategies.
  - Conducted literature review on scalable quantum methods for materials science.

- **APL Quantum 1 December 2024; 1 (4) – "Exploring qubit-ADAPT-VQE for materials discovery in direct air capture"**
  - Applied qubit-ADAPT-VQE to compute potential energy surfaces for simplified representations of Metal Organic Frameworks.
  - Designed and executed all numerical experiments in classical and quantum hardware.
  - Wrote core sections of the manuscript.

- **arXiv:2503.10923 – "Scaling active spaces in simulations of surface reactions through sample-based quantum diagonalization"** (Submitted to peer-review)
  - Constructed embedded Hamiltonians for Li–O_2 battery interfaces.
  - Applied SQD and Extended SQD to compute reaction energies and validate against classical methods.
  - Designed and executed all numerical experiments.
  - Wrote core sections of the manuscript.

- **arXiv:2503.10901 – "Computing band gaps of periodic materials via sample-based quantum diagonalization"** (Co-author; Submitted to peer-review)
  - Contributed code that was integrated into the (Ext-)SQD simulation workflow.
  - Provided technical assistance for adapting SQD to periodic systems used in band gap estimation.
  - Helped guide and deploy the computational setup and interpretation of numerical results.

- **Scientific Data 12 November 2025; 12 (1776) – "TBHubbard: tight-binding and extended Hubbard model dataset for metal-organic frameworks"** (Co-author)
  - Provided technical feedback on the parameterization of extended Hubbard models.
  - Participated in discussions guiding the structure and intended scope of the database.
  - Advised on potential quantum simulation use cases and integration opportunities.

2 Quantum Computing in Materials Science: Challenges and Outlook

This chapter is adapted-condensed, and lightly updated-from the perspective article [1].[1] *The candidate contributed to Sections III.A, III.C, III.F, and VII of the article, covering qubit mapping, measurement, circuit transpilation, and use-case identification. The condensed version presented here surveys algorithmic and infrastructural bottlenecks that arise when integrating quantum computing into materials science workflows and purposefully omits sections IV and V of the original work. It thereby sets the stage for the variational algorithms discussed in Chapter 3, the sample-based simulation methods developed in Chapters 4–6.*

2.1 Introduction

Materials-science use cases are being investigated as some of the first applications of quantum computing to show practical quantum advantages. The strongest motivation for this belief is that a great deal of materials-science problems are quantum in nature [63]. Consequently, any computational advance in solving many-body quantum problems will reverberate from basic discovery to industrial deployment and sustainability goals.

Today, high-performance supercomputing centres devote a substantial share of their resources to electronic-structure and atomistic simulations [64–67]. Yet *ab-initio* calculations carried out in the full configuration interaction (CI) limit suffer from exponential or factorial growth in memory and runtime requirements with system size [68–73]. Approximation hierarchies can delay but not remove this wall: they either fail for strongly correlated materials or become intractable as the system size and complexity increase [74–85].

Quantum computers offer a route around this classical scaling crisis, encoding the wave function directly in entangled qubits and thus avoiding the exponential memory blow-up [63, 86]. In realistic deployments, these devices will serve as accelerators inside a classical high-performance environment—a model known as *quantum-centric supercomputing* (QCSC). Heavy classical processing remains indispensable: to pre-screen chemical spaces, steer adaptive algorithms, mitigate hardware noise, and decode fault-tolerant syndromes [50].

In this chapter, we motivate the need for QCSC in materials research, identify the principal algorithmic and infrastructural challenges that stand between today's hardware and a genuine quantum advantage, and propose concrete directions for addressing them. Specifically, we examine key algorithms, such as ground-state, excited-state, and discrete-optimisation approaches with potential for quantum advantage in materials settings. As

[1] Published version: Y. Alexeev *et al.*, "*Quantum-centric supercomputing for materials science: A perspective on challenges and future directions,*" *Future Generation Computer Systems* **160**, 666–710 (2024). © Elsevier BV, reproduced here in accordance with the publisher's policy permitting inclusion in a non-commercial doctoral thesis.

well as, materials-science use cases, applications that satisfy three criteria: (i) classical hardness; (ii) compatibility with NISQ or fault-tolerant devices, depending on the scenario; and (iii) scientific relevance.

The remainder of the chapter is organised as follows. Section 2.2 surveys core quantum algorithms; Section 2.3 analyses the classical information-processing steps in hybrid workflows; Section 2.4 summarises classical quantum-circuit simulators; and Section 2.5 brings these strands together in a set of representative materials science use cases.

2.2 Key Quantum Algorithms for Materials-Science Workflows

This section surveys the quantum algorithms that are most relevant to materials-science applications. Although simulating quantum many-body systems is the most natural target for a quantum computer, certain partial differential equations (PDEs) such as Maxwell's equations or the heat equation, also arise in multiscale materials workflows. We therefore treat both categories, beginning with real-time dynamics and progressing to ground-state, open-system, finite-temperature, and PDE methods.

Simulating real-time evolution under a Hamiltonian is arguably the most direct use case for quantum computers, and three algorithmic families span the range from near-term practicality to asymptotic optimality. At the high-end is *qubitization*, the culmination of quantum signal-processing ideas, which implements the propagator e^{-iHt} with only $O\big(t + \log(1/\epsilon)\big)$ oracle queries for evolution time t and target error ϵ [87]; the price is an involved Hamiltonian oracle and ancillary qubits, making qubitization a natural fit for early fault-tolerant, rather than NISQ-era, devices. On the opposite extreme lie the Trotter–Suzuki product-formula methods that inaugurated quantum simulation [86, 88]: a p-th-order formula requires $O\big(t^{1+1/p}/\epsilon^{1/p}\big)$ gate operations, permits a fine trade-off between circuit depth and accuracy, and exploits any commutativity present in the Hamiltonian, keeping resource demands low enough for present-day hardware. Bridging these regimes is the *HHKL* algorithm [89], which patches together local Trotter steps using Lieb–Robinson bounds to deliver nearly optimal scaling for geometrically local or even power-law decaying interactions on a lattice [90], effectively combining the structural economy of Trotterization with the precision guarantees of signal-processing approaches.

Predicting ground- and other low-lying energy states is the quintessential materials-science workload for quantum simulation, yet even deciding whether a local Hamiltonian is gapped is QMA-hard in the worst case [54, 55]. In practice, three algorithmic tiers have emerged. At the shallow-circuit end, VQE dominates current prototypes: by preparing a low-depth, parameterised ansatz and variationally minimising the measured energy, it has been demonstrated on platforms ranging from NMR and photonics to superconducting and trapped-ion qubits [30–32, 91–110]. Related finite-depth QAOA styles have been proposed for frustrated lattices [111, 112]. VQE's flexibility, however, comes with steep measurement overheads and optimisation landscapes pitted with barren plateaus, leaving no guarantee of convergence [36, 37]. At the opposite extreme, a fault-tolerant machine can invoke QPE [113] and its modern refinements [114–116], which project onto an eigenstate with success probability proportional to the squared overlap between the actual ground state and an initial guess; assumptions on that overlap can be relaxed by appealing to

dissipative mixing times [117].

Bridging these two ends is a fast-growing family of *quantum Krylov* or subspace-diagonalization methods [33, 118–137]: by repeatedly applying short-time evolutions and building a small Krylov basis, these algorithms confine the eigenproblem to a classically diagonalized subspace. Many variants [119, 120, 123–125, 127–129, 131–135, 138, 139] enjoy provable convergence bounds comparable to QPE while requiring circuits only modestly deeper than VQE.

Furthermore, other sample-based approaches provide an alternative middle path: methods such as SQD build an effective subspace from snapshots of a sampled circuit and solve the eigenproblem in that sampled subspace without variational optimization [6, 7, 42–47]. Because both Krylov and SQD techniques exploit the locality of spin or fermionic lattice models, they are promising candidates for chemically accurate ground-state energies on noisy hardware.

Many phenomena in chemistry, condensed matter, and nanoscale devices cannot be captured by purely unitary dynamics; coupling to phonons, photons, or other bath modes must be included explicitly. Quantum computers can attack this *open-system* regime by simulating the Lindblad master equation, which augments coherent Hamiltonian evolution with dissipators that model environmental interactions [140]. Early proposals embed the bath in auxiliary qubits and evolve the joint system directly [141–143], but this becomes impractical when the environment is large. Digital Trotterization offers a scalable alternative: it approximates the evolution of k-local Lindbladians [142] and can even handle non-Markovian kernels by interleaving coherent and jump operators [144]. For sparse or long-range generators the cost can be reduced further by implementing sparse Stinespring isometries [145] or by decomposing the Lindbladian as a linear combination of Pauli channels and applying a variant of the Linear Combination of Unitaries (LCU) framework [146]. Imaginary-time techniques provide yet another route, replacing dissipative evolution by a sequence of non-unitary projections that can be digitised on gate hardware [147]. More recently, qubitization has been generalised to Liouvillian dynamics [148], wave-matrix "Lindbladisation" schemes have been proposed to map open evolution into enlarged Hamiltonian problems [149, 150], and partial probabilistic error-cancellation protocols have been suggested to embed dissipation within noise-mitigated circuitry [151]. Together, these approaches lay the groundwork for first-principles modelling of charge transport, exciton diffusion, and light-harvesting processes [117, 152–163].

Accurate finite-temperature expectation values are indispensable for analysing phase transitions, transport, and battery operating windows. Unlike ground-state preparation which inherits QMA-hardness in the worst case [54], preparing a Gibbs state at inverse temperature β is typically easier once β is not too large. Two algorithmic lines have emerged. The first adapts projection ideas from phase estimation: Grover-assisted rejection sampling cools an infinite-temperature ensemble to a target β [164, 165], and a quantum heat-bath trick combines two Gibbs states at β to obtain one at 2β [166]. The second line seeks to mimic physical thermalisation. Discrete-time quantum Metropolis algorithms over energy eigenstates converge provably to the Gibbs distribution [167], and quantum walks provide quadratic speed-ups over classical Markov-chain Monte Carlo [168, 169]. Continuous-time approaches simulate Davies generators—Liouvillians believed to model natural thermalisation [170]—and recent work shows how to implement them efficiently even with long-range interactions [155, 171, 172]. Some proposals reduce quantum re-

sources to a single ancilla qubit plus classical post-processing [117], while others probe the fundamental link between thermalization and the computational power of quantum dynamics [173].

Beyond the Schrödinger equation, materials modelling frequently involves partial differential equations, such as Maxwell's equations for electromagnetic response, the heat equation for thermal transport, and elasticity or eigenmode equations for structural analysis, problems that are usually discretized with the finite-element method (FEM). When an FEM discretization turns a boundary-value problem into a large, sparse linear system, quantum linear-systems algorithms (QLSAs) provide a natural accelerator [174]. Because FEM matrix elements are often known analytically, one can load them without a QRAM, preserving the polynomial speedups demonstrated for generic finite-element workflows [175] and for specific applications such as heat diffusion [176] and electromagnetic cavity design [177]. A complementary literature treats initial-value problems: linear cases have been tackled with Hamiltonian embeddings and sparse-grid techniques [178, 179], while nonlinear PDEs can sometimes be recast into a higher-dimensional linear Schrödinger form [180]. Such embeddings work only when the resulting wavefunction remains normalisable [181] and when the target dynamics do not amplify chaos so strongly that Trotter errors explode [182]; nevertheless, they illustrate how a surprisingly broad class of classical dynamics can be folded into a quantum-simulation framework [183].

Taken together, the dynamical, ground-state, open-system, finite-temperature, and PDE algorithms surveyed above constitute a menu from which complete, quantum-centric materials workflows can be assembled. *This thesis, however, narrows its attention to a specific subset of that menu.* In the chapters that follow, we will focus on sample-based techniques for electronic-structure problems, and their integration with embedding strategies on classical high-performance computing (HPC) resources. Other methods such as qubitized real-time dynamics or quantum finite-element solvers, are discussed here for context but remain outside the scope of the original research presented in this thesis. The next sections will focus on how we process and encode classical information for quantum computers.

2.3 Classical Information Processing

This section sets the technical levers (encodings, simulation primitives, measurement, compilation, mitigation) that determine qubits, depth, and shots for the applications studied later.

2.3.1 Mapping to qubits

Quantum computers inherently operate on two-level systems, yet most physical models of interest in materials science are expressed in terms of fermionic or bosonic modes. Thus, a preliminary but crucial step in quantum simulations involves translating the original fermionic or bosonic Hamiltonian into an equivalent representation using qubits. A suitable mapping must preserve the underlying operator algebra: canonical anticommutation relations (CAR) for fermions, and canonical commutation relations (CCR)

for bosons, while efficiently utilizing quantum resources. Although qubits themselves can be considered as hard-core bosons [184], directly encoding generic bosonic or fermionic modes into qubits typically requires careful consideration.

Bosonic modes, having infinitely many accessible states, first require a truncation to a finite occupancy cutoff to be represented digitally. Once this cutoff is chosen, encoding the occupation numbers into qubits involves balancing qubit count against the locality of operators. Unary encodings yield simpler, more localized operators at a linear qubit cost, while binary encodings drastically reduce qubit count at the price of increased operator non-locality. This trade-off is well documented and explored in detail in the literature [185–191].

For fermionic modes, the antisymmetric nature of fermionic wavefunctions introduces additional complexity. Fermionic operators anticommute, while spatially separated qubit operators naturally commute. The earliest known transformation addressing this issue is the Jordan–Wigner (JW) mapping [192], which represents fermionic operators as strings of Pauli matrices, incorporating non-local "string" terms to preserve anticommutation. The field has subsequently developed numerous alternative mappings, including parity, Bravyi–Kitaev (BK), and Bravyi–Kitaev superfast (BKSF), each designed to optimize different resources such as operator locality, qubit count, or opportunities for symmetry-based reductions [107,110,193–203]. Some mappings differ from JW by a simple unitary transformation, such as a Clifford conjugation, while others target specific lattice structures, preserving anticommutation relations only among a carefully selected subset of fermionic operators.

The choice of fermion-to-qubit mapping profoundly impacts the computational cost and practicality of simulations. In the near-term quantum era, hardware constraints such as limited qubit count and shallow circuit depths strongly favor mappings optimized for compactness and operator locality. In this context, the parity transform, BK transform, and their variants (particularly when coupled with automated circuit optimization and compiler techniques [105, 204]) can deliver substantially more efficient quantum circuits than the JW mapping [107, 110, 196]. Additionally, lattice-aware mappings that exploit specific geometric structures and limited qubit connectivity in quantum hardware have also shown promise in reducing resource demands [107, 200, 201, 203, 205, 206].

Although most current fermion-to-qubit mappings fall into the general framework of mapping products of Majorana operators to Pauli strings [203], further exploration and innovation within this framework remain valuable. Future research directions involve systematically exploring the combinatorial space of mappings to identify even more resource-efficient strategies or moving beyond the Majorana-to-Pauli paradigm altogether to develop entirely new mapping concepts and thus significantly expand our simulation capabilities.

Regarding the work in this thesis the fermion–to–qubit mapping (and symmetry tapering) sets qubit counts and operator locality for all embedded Hamiltonians used in our studies on Chapters 3,5, and 6

2.3.2 Hamiltonian encodings for dynamics simulation

The intractability of classically simulating generic quantum dynamics has long motivated quantum computational research [63], representing perhaps the clearest path towards achieving practical quantum advantage [86]. At the core of such dynamics simulations is the ability to implement the time-evolution operator e^{-iHt} for a given Hamiltonian H. Typically, the Hamiltonian is decomposed into a sum of simpler operators, $H = \sum_i H_i$, each of which can be diagonalized using straightforward unitary transformations as $H_i = U_i D_i U_i^\dagger$. Examples include exact squares of one-electron fermionic operators in quantum chemistry, weighted sums of commuting Pauli operators, or one-sparse operators. The classical pre-processing choice of decomposition, analysis of term sparsity, and commutation structure significantly impact the efficiency of subsequent quantum algorithms [207]. Depending on these choices, well-established methods such as Trotterization and linear combination of unitaries (LCU) can be employed for compiling the time-evolution operator into basic quantum gates.

Trotterization

The fundamental principle of Trotterization is the approximation of a complex quantum evolution through the sequential application of simpler evolutions. For a Hamiltonian decomposed as $H = \sum_i H_i$, the simplest first-order Trotter approximation is given by:

$$\left(\prod_{k=1}^{m} e^{-iH_k \frac{t}{n}}\right)^n = e^{-iHt} + O\left(\frac{m^2 t^2}{n}\right). \tag{2.1}$$

The approximation error decreases with increasing Trotter steps n. Higher-order formulas, such as the second-order Trotter-Suzuki approximation:

$$\left(\prod_{k=1}^{m} e^{-iH_k \frac{t}{2n}} \prod_{k=m}^{1} e^{-iH_k \frac{t}{2n}}\right)^n = e^{-iHt} + O\left(\frac{m^3 t^3}{n^2}\right), \tag{2.2}$$

offer improved error scaling at the expense of implementation complexity. Specific applications may favor particular Trotter orders; for instance, the second-order formula is preferred in certain lattice gauge theory simulations due to its manageable complexity and favorable scaling properties [191, 208], while the fourth-order formula has demonstrated superior accuracy in medium-sized Heisenberg Hamiltonian simulations [209, 210].

The accuracy of Trotterization scales as $O(t^{p+1}/n^p)$ for a product formula of order p [86, 211]. However, since this error arises from non-commuting Hamiltonian terms, tighter bounds can be obtained by careful term ordering and grouping, exploiting the Hamiltonian's commutation structure. Recent works have made progress towards such refined error analyses [88, 207, 212]. In practice, the accuracy that matters most is the fidelity of the state resulting from applying the Trotterized evolution to a particular initial state, highlighting the importance of further state-dependent optimization.

While Pauli operator decompositions are standard, Trotterization applies generally to any decomposition into easily simulated operators. For example, one-sparse Hamiltonians have been simulated via Trotterization [211, 213–216], though detailed error analysis for these remains limited.

For Hamiltonians expressed in terms of Pauli operators $H = \sum_i c_i P_i$, traditional upper bounds on Trotter steps use the $l1$-norm $|||H||| = \sum_i |c_i|$. However, this approach neglects commutation relations, often leading to significant overestimation. More refined analyses involving nested commutators offer tighter step count bounds, greatly reducing required quantum resources [88]. Nevertheless, these methods incur high classical computational costs ($O(N^{p+1})$) for a Hamiltonian with N terms. Efficient classical strategies, leveraging high-performance computing to optimize term ordering and grouping, are thus crucial to harnessing quantum hardware effectively and achieving practical quantum advantage.

Ultimately, the optimal Trotterization approach depends heavily on the specific simulation application, initial state, and desired accuracy. Analyzing these factors carefully will invariably yield further reductions in quantum resources compared to generic Hamiltonian-based estimates.

Linear Combination of Unitaries

An alternative to Trotterization involves representing the Hamiltonian as a linear combination of unitary operators (LCU). In this framework, one constructs a block-encoding of H into a larger unitary operator $U(H)$, typically expressed as:

$$H = \sum_n c_n U_n, \quad U(H) = \sum_n \mathcal{P}n \otimes U_n, \tag{2.3}$$

where the operators $\mathcal{P}n$ form a complete, orthogonal set of projectors acting on ancilla qubits. The performance of LCU-based simulation algorithms strongly depends on the $l1$-norm $\lambda = \sum_n |c_n|$ [217]. Consequently, significant efforts have aimed at reducing λ through optimal choices of U_n operators [218–220]. A fundamental lower bound on λ is set by half the Hamiltonian's spectral range, $(E\text{max} - E\text{min})/2$, thus motivating modifications to the Hamiltonian itself to minimize this quantity [219].

Specifically, one can create an effective Hamiltonian that faithfully reproduces dynamics within a targeted symmetry subspace, reducing the spectral range and hence the $l1$-norm. Such effective Hamiltonians were explored in Ref. [220] and demonstrated to significantly improve LCU efficiency. Further reductions could be realized by explicitly modifying or excluding high-energy states irrelevant to the targeted dynamics.

Moreover, beyond standard Pauli products, alternative unitary decompositions (e.g., grouping anti-commuting Pauli operators or using tensor hypercontraction) have been shown to further lower the $l1$-norm [218, 219]. Evaluating these decompositions involves balancing classical computational costs, achieved $l1$-norm reductions, and quantum implementation complexity, including ancillary qubit requirements and gate counts. Such comprehensive analyses are critical for fully leveraging LCU approaches in realistic quantum simulations.

Despite not realizing any work related to dynamics simulations in this thesis this section is necessary to underscore a very relevant application of Quantum Computers in chemistry and Materials Science. It would be negligent to not include any mentions.

2.3.3 Quantum measurement

Many quantum algorithms involve obtaining numerical values of observables, such as energies and dipole moments, as expectation values of associated operators. Realistic use cases may involve observables on systems comprising 50 to 100 qubits, leading to Pauli operator decompositions with potentially millions of terms. This is particularly notable in electronic structure Hamiltonians, which scale approximately as N^4 Pauli products. Individually measuring these terms is impractically time-consuming, even on superconducting architectures with relatively rapid measurements. Thus, grouping Pauli products to reduce the number of required measurements becomes crucial.

This subsection focuses on methods for estimating the expectation value of a single operator, such as a Hamiltonian, for a single wavefunction. While the general problem involves estimating expectation values for multiple operators and wavefunctions, a scenario that poses additional optimization challenges important for certain applications [221], we limit our discussion here for clarity and brevity.

All measurement schemes described here yield estimators for the Hamiltonian expectation value by combining results of projective measurements of Pauli operators. These measurements can be conceptualized either directly as measurements of generic Pauli operators or as measurements of diagonal (Ising) Hamiltonian terms in suitably rotated quantum states. Physical Hamiltonians, such as electronic structure Hamiltonians, generally cannot be transformed directly into a single diagonal form through local unitary transformations. Hence, the Hamiltonian $\hat{H}$ is usually decomposed into simpler fragments $\hat{H}_n$ that can each be rotated into diagonal (Ising) form via unitary transformations $\hat{U}_n$:

$$\hat{H} = \sum_n \hat{H}_n, \quad \hat{H}_n = \hat{U}_n^\dagger \hat{Z}_n \hat{U}_n, \tag{2.4}$$

and the expectation value is estimated as

$$\mathbb{E}(\hat{H}) = \sum_{n,k} \frac{Z_n^{(k)}}{M_n}, \tag{2.5}$$

where $Z_n^{(k)}$ are the measurement outcomes of $\hat{Z}_n$ in rotated states $\hat{U}_n|\psi\rangle$, with M_n representing the number of measurements per fragment.

Measurement methods can be categorized based on several criteria:

1. **Measurement approach**: Direct measurements within the Hilbert space versus measurements using ancilla qubits for positive operator-valued measures (POVMs).
2. **Functional relations**: How estimators for measurable operators combine to yield the estimator for the full Hamiltonian expectation value, either as simple linear combinations or more complex functional forms [222].
3. **Unitary transformations**: The complexity of the unitary rotations ($\hat{U}_n$) required to diagonalize the measurable operators.
4. **Estimator error scaling**: Error scaling with measurement count; either shot-noise limited ($1/\sqrt{M}$) or quantum-limited Heisenberg scaling ($1/M$). Near-term methods typically achieve the $1/\sqrt{M}$ scaling due to practical limitations on circuit depth.

Classical shadow tomography [223] is a recent and popular measurement approach. In this method, random unitary transformations ($\hat{U}_k$) typically drawn from the Clifford group (with subsequent extensions to other unitary groups [224, 225]) are applied to states before measurement. Each random transformation yields a Hamiltonian decomposition into diagonal terms $\hat{Z}_k$ and non-diagonal remainder terms $\hat{R}_k$, the latter contributing no net expectation value upon measurement. This motivates optimizing the selection of random unitary transformations to maximize the diagonal components, thereby improving sampling convergence [223, 226–233]. Recent benchmarks for these randomized measurement approaches have been performed on molecular electronic structure Hamiltonians [234].

Adaptive informationally complete (IC) POVMs offer a more general framework for measurement optimization [235,236]. These methods iteratively tailor the measurement basis to specific observables and states, and naturally incorporate noise models [237]. Additionally, IC-POVMs support estimating multiple observables simultaneously, with subsequent post-processing to mitigate measurement noise [238].

The practical efficiency of measurement schemes hinges on several factors: (1) the classical optimization cost involved in identifying optimal measurable fragments; (2) quantum overhead due to implementing additional unitary transformations; and (3) measurement overhead necessary to achieve a desired accuracy. Near-term implementations place particular emphasis on minimizing quantum overhead. For instance, methods using fermionic-based rotations [39, 239–241] may require fewer total measurements but typically demand greater quantum resources compared to simpler Clifford-based approaches [242,243]. Further measurement efficiency improvements are possible by selecting fragments that preserve symmetries, enabling symmetry-based error mitigation by discarding symmetry-violating outcomes [125, 244, 245].

A crucial ongoing research direction is identifying physical Hamiltonians amenable to measurement schemes for near-term hardware. Spin models often decompose naturally into just a few measurement bases, whereas molecular electronic structure Hamiltonians, with their N^4 scaling, present considerable measurement complexity. Intermediate cases, such as the Fermi-Hubbard Hamiltonian, offer potential tractability by permitting partitioning into relatively simple commuting sets that can be diagonalized using shallow Bell measurement circuits, simultaneously enabling error mitigation through symmetry checks [246, 247]. A detailed exploration of measurement strategies across diverse materials science simulations can be found in Ref. [248].

The measurement strategy (Pauli grouping, randomized bases, symmetry filters) determines the shot budgets and error bars for energies and band gaps reported later. In practice, we will not be concerned too much with this as we'll see that other aspects are more relevant to the quality of results at the scale we are studying with sampling based algorithms. The whole idea behind sampling based algorithms is to circumvent measuremnet bottlenecks after all.

2.3.4 Problem-level preprocessing and optimization

To achieve quantum advantage with near-term devices, optimizations must be applied at every level of the computational stack, including the problem definition itself. Since it's impossible to survey all possible input problems for quantum algorithms in

materials science, we present two illustrative examples of such problem-level preprocessing and optimization: *active space selection* and *VQE ansatz design*. This highlights a critical step that will likely be necessary for any successful application on noisy quantum computers.

Active Space Selection and Embedding

A central task in quantum chemistry is solving the eigenvalue problem for a system's Hamiltonian. However, diagonalizing the full Hamiltonian is computationally prohibitive and often unnecessary. For many systems of interest, such as point defects in semiconductors [249], catalysts [250], extended molecules [251], spin defects in solids [252–256], and molecular qubits [257, 258] the chemically relevant physics is localized to a small subset of orbitals and electrons. This allows for the definition of an *active space*, which yields a smaller, effective Hamiltonian that can be solved on a quantum computer while retaining the system's essential properties [259, 260]. The selection of this space often relies on chemical intuition, as a universal, automated approach is still lacking. This is particularly crucial for systems with strong electronic correlation (multiconfigurational systems), whose accurate description requires computationally expensive multireference methods like *Complete Active Space Self Consistent Field* (CASSCF) [261–263] or *Density matrix Renormalization Group* (DMRG) [264] that scale poorly.

An *ab-initio* method for systematically defining a tractable quantum problem is *Density Matrix Embedding Theory (DMET)* [53, 265–275]. In DMET, a user selects a "fragment" of interest (e.g., orbitals around a defect). The environment's orbitals that are entangled with this fragment are then identified via a Schmidt decomposition of the one-body reduced density matrix (1-RDM) of an approximate wavefunction for the whole system. This is done by diagonalizing the environmental block (D_{env}) of the 1-RDM:

$$D_{\text{env}} = \mathbf{U}\lambda\mathbf{U}^{\dagger} \tag{2.6}$$

The fragment and its corresponding entangled "bath" orbitals form an embedding subspace, creating a self-contained quantum problem that's much smaller than the original. The reader can look at chapter 4 for a more in-depth explanation.

This embedding framework serves as a foundation for powerful hybrid methods. For instance, in *Complete Active Space Density Matrix Embedding Theory* (CAS-DMET), a CASSCF calculation is performed on the embedding subspace to capture static correlation. Dynamic correlation can then be added using multireference perturbation theories, as in *2nd order n-electron valence state perturbation theory* (NEVPT2-DMET) [253–255, 276–278]. A more cost-effective approach is *Density Matrix Embedded Pair-Density Functional Theory* (DME-PDFT) [274], which leverages multiconfiguration pair-density functional theory (MC-PDFT) [279–281]. MC-PDFT computes the total energy using the 1- and 2-RDMs from the embedded calculation, offering an efficient way to capture both static and dynamic correlation. This approach is particularly well-suited for embedding because it uses semi-local functionals, making it less sensitive to the size of the embedded fragment than methods like NEVPT2-DMET [274].

Alternative partitioning schemes also exist, such as the *localized active space self-consistent field (LASSCF)* method [268, 282]. LASSCF divides the active space into smaller, localized subspaces that are treated with a high level of theory, connected only

through a mean field. While LASSCF captures strong intra-fragment correlation, the weaker inter-fragment correlations must be added separately. This can be done classically with methods like State Interaction (LASSI) [283] or on a quantum computer via *LAS-UCC* [284], where VQE is used to variationally determine the correlation between the fragments.

Ultimately, accurately modeling the "large" active spaces found in challenging chemical systems remains a significant hurdle. Classical methods like DMRG and selected CI [69, 71, 285–287] must rely on approximations to remain tractable. Quantum computing promises to solve these problems with higher accuracy as hardware improves. Hybrid quantum-classical approaches like LAS-UCC and quantum implementations of perturbation theory [288] are promising avenues for bridging this gap.

VQE Ansatz Design and Initialization

The VQE is a leading algorithm for solving eigenvalue problems on near-term quantum hardware [30], with numerous demonstrations on small systems [30–32, 91–104, 106, 108, 109]. The success of VQE depends critically on the choice of the parameterized quantum circuit, or *ansatz*, $U(\vec{\theta})$. An effective ansatz must strike a balance between expressivity (the ability to accurately represent the true ground state) and practicality (modest circuit depth and trainable parameters). This creates a fundamental trade-off. On one end, chemically-inspired ansatzes like unitary coupled-cluster (UCC) [30, 94, 105, 289] are highly expressive but often result in circuits that are too deep for current hardware. On the other end, hardware-efficient ansatzes [32] are designed for specific devices but may be difficult to train due to barren plateaus [36, 37] and can fail to capture the necessary physics [290].

Significant research aims to find a middle ground. Adaptive methods like *ADAPT-VQE* [291] and qubit coupled-cluster [292–294] build the ansatz iteratively, using a pre-screening step to add only the most significant operators. This strategy reduces both circuit depth and the number of parameters. While the scalability of training these adaptive ansätze remains an open question, some techniques like HMP2 [107] show promise. Furthermore, the increased measurement overhead associated with the operator selection process can be mitigated with advanced measurement schemes like IC-POVMs [295].

Beyond the ansatz structure, the initialization of the variational parameters is equally critical for successful optimization. Random initialization can lead to poor convergence, especially on noisy hardware. A powerful strategy is to use classical pre-computation to find a high-quality starting point. A prime example is the *Clifford Ansatz For Quantum Accuracy (CAFQA)* [296]. This method first constructs an ansatz using only Clifford gates, which can be efficiently simulated and optimized on a classical computer. The resulting parameters from this classical search are then used as the initial guess for a standard VQE optimization with a more expressive ansatz. This idea of bootstrapping VQE can be extended with other classically tractable methods, such as near-Clifford circuits or tensor network approaches [297, 298]. These initialization techniques can be made even more powerful by incorporating problem-specific knowledge to guide the classical search.

Active-space/embedding defines the reduced Hamiltonians we utilize on Chapters 3, and 5, while ansatz design and initialization provide the variational baselines and

comparisons in Chapter 3. These are pivotal to execute the experiments later described in those chapters.

2.3.5 Circuits, transpilation, and architecture

Just like in classical computing, quantum computations must be broken down into a sequence of hardware-compatible operations. These operations are typically expressed in the form of a quantum circuit, which must be transpiled into native gates that respect the target device's connectivity and architecture. The efficiency and feasibility of executing a quantum algorithm in the context of scientific applications like materials simulation critically depend on circuit design, compilation strategies, and hardware capabilities. This section surveys key developments and considerations at the circuit level, with a focus on connectivity and circuit optimization.

Connectivity

Many quantum algorithms require pairwise interactions across qubits that may not be directly connected in hardware. On devices with restricted connectivity, such as nearest-neighbor architectures, this constraint is typically overcome through the use of SWAP networks. A well-known example is the linear-depth SWAP network [299–301], which enables all $\binom{n}{2}$ qubit pairs to interact using only $O(n)$ circuit depth. This technique allows efficient Hamiltonian simulation on linearly connected devices and serves as a compelling alternative to more complex all-to-all architectures [302, 303].

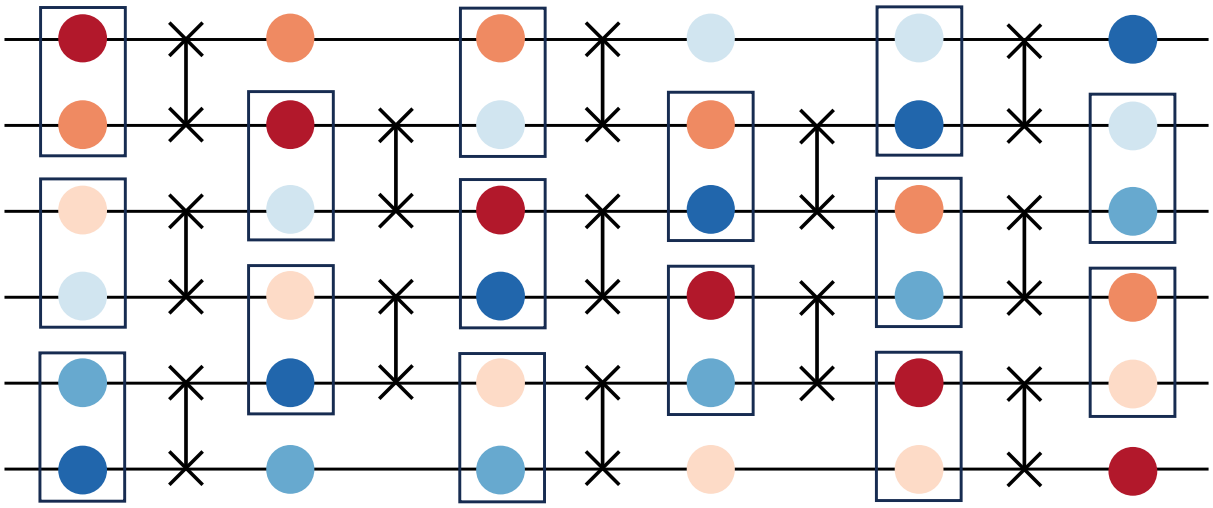


Figure 1 – **Example SWAP network for $n = 6$ qubits.** In $O(n)$ steps, each of the $O(n^2)$ pair of qubits (colors) performs an interaction, even with just linear connectivity. *Reproduced from Y. Aleexev, et al., [1] (DOI) © 2024 Elsevier. Used with permission.*

As quantum processors scale, hardware limitations make it increasingly difficult to fit large circuits onto a single monolithic QPU. This has led to growing interest in distributed quantum computing, where multiple QPUs are interconnected through quantum links to act as a larger composite system. Modular architectures as those of companies

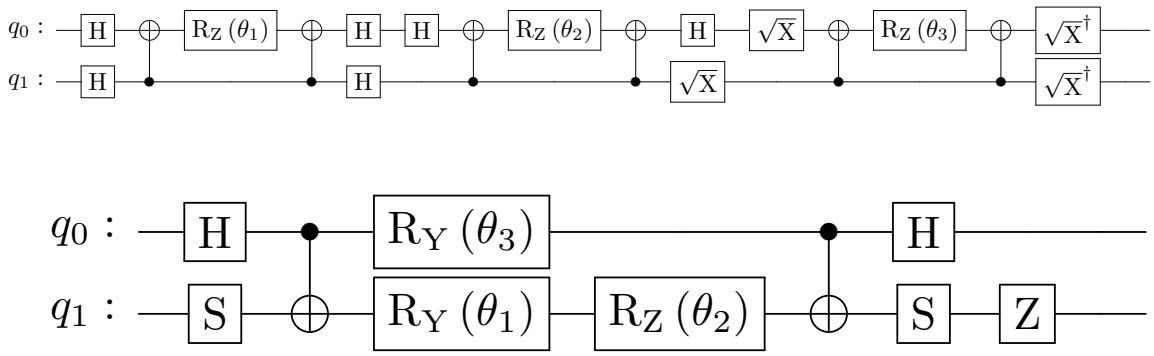


Figure 2 – **Two circuits implementing the same sequence** $R_{\mathrm{YY}}(\theta_3)\cdot R_{\mathrm{XZ}}(\theta_2)\cdot R_{\mathrm{XX}}(\theta_1)$**.** *Reproduced from Y. Aleexev, et al., [1] (DOI) © 2024 Elsevier. Used with permission.*

like IBM, Quantinuum, and IonQ promise this kind of extensibility. Supporting work has already explored hardware modeling [304], interconnect-aware compilation [305], and domain-specific mappings for algorithms such as variational quantum eigensolvers [306]. While promising, these approaches must contend with real-world challenges like quantum communication latency and limited entanglement fidelity between modules.

Complementing these hardware efforts, circuit cutting has emerged as a software-driven strategy to circumvent current device limitations. In this approach, large quantum circuits are divided into smaller, independently executable fragments [307, 308]. Each fragment runs on a smaller QPU, and classical post-processing reconstructs the global output. While early implementations incurred exponential costs, more recent techniques such as ScaleQC [309] have reduced overheads to sub-exponential in favorable scenarios. Circuit cutting has thus expanded the practical scope of quantum workloads, though its scalability depends on improvements in classical reconstruction and tighter integration with backend infrastructure [310].

Circuit optimization

Once connectivity constraints are addressed, optimizing the circuit structure itself becomes critical. In Hamiltonian simulation, this often involves synthesizing sequences of Pauli rotations $R_{P_i}(\theta_i)$. Finding compact decompositions while respecting hardware topology can significantly reduce circuit depth and gate count. Prior work [107,209,311,312] has drawn connections between this problem and parity network synthesis [313,314], offering heuristics that exploit commutation and structure. However, extending these methods to include ancilla qubits or mid-circuit measurement remains an open and active research direction.

A particularly powerful optimization tool is the use of dynamic circuits that incorporate mid-circuit measurements and feedforward. Though the principle of deferred measurement guarantees logical equivalence, mid-circuit measurements can yield significant practical advantages. For instance, feedforward allows for more efficient implementations of controlled rotations [315] and constant-depth constructions of GHZ states, even on

devices with limited connectivity [316]. These methods have proven especially useful in optimizing simulation subroutines such as CNOT ladders and merit further compiler-level support.

Circuit optimizations must also be tailored to the cost model of the target hardware. On pre-fault-tolerant devices, single-qubit gates are generally cheaper than entangling gates, whereas in fault-tolerant settings, Clifford gates are much less costly than T gates. Effective compilers can exploit these asymmetries, producing backend-specific code without requiring the user to manage these details manually [107, 204]. Such abstractions are crucial for bridging usability and performance in future quantum software stacks.

Beyond exact implementations, allowing approximate unitary synthesis can further reduce overhead. Widely used techniques include approximate quantum Fourier transforms [317], approximate SWAPs [318], and approximate arithmetic in Shor's algorithm [319]. In many cases, approximations tailored to a known input state can drastically reduce resource usage while preserving accuracy [320, 321]. These methods mirror classical approximation techniques and suggest fruitful directions for quantum compiler research.

Taken together, these developments in connectivity, circuit decomposition, and approximation form the foundation of quantum circuit design. As hardware matures, compiler technologies that account for architectural constraints, resource tradeoffs, and distributed execution will be key to unlocking scalable quantum applications in materials science and beyond.

Connectivity-aware synthesis (SWAP networks), transpilation, and optional circuit cutting translate operator counts into backend depth/width in our resource estimates available later in chapters 3,5 and 6. one needs to be aware of these mechanisms to minimize depth and improve the quality of results from quantum calculations.

2.3.6 Error mitigation

In the absence of large-scale fault-tolerant quantum computers, noise remains a defining challenge for near-term quantum computation. Although full quantum error correction (QEC) promises scalable protection against decoherence and gate errors, its overhead is still out of reach for most contemporary devices. In the interim, a class of techniques known as quantum error mitigation (QEM) has emerged as a practical alternative. Rather than correcting errors during computation, QEM methods seek to reduce the impact of noise on observables through repeated measurements and classical post-processing [100, 245, 322, 323]. These approaches are especially critical in tasks such as estimating expectation values of quantum operators, where noise introduces bias and variance into the output statistics.

One such method, probabilistic error cancellation (PEC), provides a theoretically rigorous route to recovering ideal expectation values from noisy circuits [324]. This technique requires learning a detailed noise model for the hardware, then constructing a weighted ensemble of noisy operations that, on average, cancel the effects of the noise. The accuracy of the results improves with the quality of the noise model, and the estimator remains unbiased under the assumption that the noise remains stable during execution. However, this comes at a steep price: the number of samples required grows exponentially

with the noise rate and circuit size. Recent proposals have explored mitigating this overhead by replacing sampling with tensor-network simulations [238], which can reduce the cost quadratically at the expense of additional classical computation. Other efforts focus on improving noise parameter estimation to more tightly bound the sampling burden.

Zero-noise extrapolation (ZNE) represents another prominent approach within QEM, one that avoids explicit noise inversion. Instead, it infers the ideal observable by evaluating the quantum circuit at artificially amplified noise levels and fitting the resulting data to extrapolate toward the zero-noise limit. Noise amplification can be performed via analog methods such as pulse stretching [325], digital repetition of subcircuits [326–328], or probabilistic circuit resampling based on a learned noise model [329]. While ZNE does not offer unbiased results in the strict sense, it is often more practical than PEC due to significantly lower sampling costs. Nevertheless, its accuracy depends heavily on the quality of the extrapolation, the noise scaling technique, and the choice of fitting model, factors that continue to be the subject of ongoing optimization [328]. Extensions of ZNE to error-corrected circuits have also recently been proposed [330].

A complementary, in-circuit suppression tactic is dynamical decoupling: sequences of π pulses (e.g., CPMG, XY4/XY8, UDD) inserted during idle windows to average out slow dephasing and other quasi-static errors by open-loop control [331, 332]. DD's effect is well captured by the filter-function formalism, which connects a chosen sequence to the device's noise power spectrum and predicts which frequencies are attenuated [333]. Experiments across platforms (NV centers and superconducting qubits) show sizable coherence extensions and fidelity gains under appropriate sequences [334–336]. In near-term workflows, DD is best viewed as error suppression that reduces the effective noise scale in QEM: it can flatten ZNE extrapolation curves and lower the sampling overhead for PEC, but it is not unbiased correction and may offer limited benefit for T_1-dominated noise or when pulse imperfections/crosstalk accumulate. Practical use therefore depends on scheduler integration (apply on idle wires only, avoid collisions with two-qubit gates) and on empirical calibration against the device's spectral noise features.

An alternative strategy for improving quantum output fidelity relies on leveraging ensemble correlations across different circuit configurations. The Quancorde protocol [337], for instance, uses a suite of classically simulable "canary circuits" designed to resemble the target circuit structure to estimate how different mappings or devices respond to noise. These structural proxies allow for ranking circuit variants by expected fidelity. Measurement outcomes from the target computation are then reweighted according to this ranking, effectively enhancing the probability of correct outcomes while suppressing spurious contributions. This method is particularly useful in cloud-based or multi-device contexts, where multiple candidate configurations are available.

Taken together, QEM techniques have played a central role in pushing the capabilities of current quantum hardware, enabling demonstrations of quantum simulation and other tasks at the edge of classical tractability [329]. As quantum processors improve, QEM is expected to remain a key component of hybrid quantum-classical workflows, both as a stopgap before widespread QEC and as a potential complement to it. One exciting direction is the co-design of quantum algorithms and mitigation protocols: tailoring circuits to specific noise profiles, or designing mitigation strategies that exploit known structure in the application. Such synergies will likely be essential for practical near-term quantum advantage.

Lightweight error-mitigation techniques are used and detailed in all hardware experiments in Chapter 3, Chapters 5 and 6.

2.4 Classical Simulation of Quantum Systems

2.4.1 Why use classical simulators

Although quantum computers are envisioned to eventually surpass classical machines in certain computational tasks, classical simulation of quantum circuits continues to play a central role in the current, pre–fault-tolerant era [130, 297, 298, 338–342]. For researchers developing algorithms, workflows, and applications in quantum computing classical simulators offer capabilities that complement and, in many cases, enable the effective use of quantum hardware.

Several practical motivations underpin their use:

1. **Efficient development and debugging:** Quantum hardware access is limited and often incurs queue times or usage costs. Iterative development of quantum circuits, including debugging and design-space exploration, benefits greatly from rapid turnaround on classical simulators before deployment to real devices.

2. **Noise and error modeling:** Real quantum processors are susceptible to decoherence, gate errors, and readout imperfections. Classical simulators allow noise models to be selectively applied and tuned, enabling controlled studies of algorithmic robustness and the evaluation of mitigation strategies such as error suppression or correction.

3. **Scaling studies:** While hardware qubit counts remain modest, simulators can be configured to explore larger problem sizes, thereby informing resource scaling estimates and guiding algorithm design for future architectures.

4. **Flexible inspection of quantum states:** Simulators can expose the full quantum state vector or density matrix, allowing the researcher to probe intermediate results, entanglement measures, or subsystem properties that would not be directly accessible on hardware.

5. **Algorithm–hardware co-design:** Simulators can incorporate hardware-specific constraints without incurring physical-device limitations, allowing developers to test mapping and transpilation strategies in a controlled environment.

6. **Variational pre-optimization:** Many variational algorithms require repeated quantum–classical iterations. Running early-stage parameter sweeps on simulators can reduce hardware execution time by starting from parameters that are already near optimal [130, 297, 298]. CAFQA is an example we mentioned earlier.

7. **Hybrid HPC integration:** Classical HPC resources can be coupled with quantum hardware, for example, in schemes such as "Operator Backpropagation" [342], where part of a circuit is executed on hardware and the remainder is simulated classically

before combining results. This can effectively extend achievable circuit depth for certain workflows.

In short, classical simulation environments remain a critical tool not only for algorithm prototyping and benchmarking but also for hybrid execution workflows in HPC contexts. As quantum computing matures, such simulators are expected to remain embedded within multi-node computing environments, serving as an additional execution backend for subcircuits that fall within their feasible scope.

2.4.2 Quantum circuit simulators

Classical simulation of quantum circuits can be carried out using a variety of numerical techniques, each optimized for particular state representations, noise models, or problem structures. In this section, we outline the most widely used simulation paradigms, discuss their computational characteristics, and highlight their applicability to quantum computing research, including materials-science–oriented workloads.

It is worth noting that not every quantum circuit presents a challenge for classical hardware. For example, Clifford circuits [343–345] and certain structured unitaries such as the quantum Fourier transform [346, 347] can be simulated efficiently for large qubit counts. Circuits with sparse entanglement or restricted input states may also admit efficient simulation strategies [345, 348, 349]. Likewise, circuits with a modular structure, such as sparsely coupled clusters of highly connected qubits, can sometimes be simulated via decomposition techniques similar to circuit cutting (Sec. 2.3.5) [350].

State-vector simulators

The state-vector formalism represents a pure n-qubit state as

$$|\psi\rangle = \sum_{i=0}^{2^n-1} \alpha_i |i\rangle$$

where the amplitudes α_i are complex-valued and stored explicitly in memory. Applying a sequence of gates $U_{m-1} \cdots U_1 U_0$ corresponds to updating these amplitudes via small dense matrix multiplications. In practice, 1-qubit gates are 2×2 matrices and 2-qubit gates are 4×4 matrices, consistent with standard gate decompositions [351].

For example, applying a 1-qubit gate $U_{2\times2}$ to qubit q updates the relevant amplitude pairs:

$$\begin{bmatrix} \alpha_{s_i} \\ \alpha_{s_i+2^q} \end{bmatrix} \rightarrow U_{2\times2} \cdot \begin{bmatrix} \alpha_{s_i} \\ \alpha_{s_i+2^q} \end{bmatrix} \tag{2.7}$$

where $s_i = \lfloor i/2^q \rfloor 2^{q+1} + (i \bmod 2^q)$, for $i \in [0, 2^{n-1} - 1]$. Two-qubit gates $U_{4\times4}$ similarly act on quadruples of amplitudes according to

$$\begin{bmatrix} \alpha_{s_i} \\ \alpha_{s_i+2^p} \\ \alpha_{s_i+2^q} \\ \alpha_{s_i+2^p+2^q} \end{bmatrix} \rightarrow U_{4\times4} \cdot \begin{bmatrix} \alpha_{s_i} \\ \alpha_{s_i+2^p} \\ \alpha_{s_i+2^q} \\ \alpha_{s_i+2^p+2^q} \end{bmatrix} \tag{2.8}$$

where $s_i = \lfloor \lfloor i/2^p \rfloor / 2^{q-p-1} \rfloor 2^{q+1} + (\lfloor i/2^p \rfloor \% 2^{q-p-1}) 2^{p+1} + (i \% 2^p)$ for every integer $i \in [0, 2^{n-2} - 1]$.

The primary limitation of this method is its exponential memory footprint: storing all 2^n amplitudes in double precision requires 16×2^n bytes. To scale beyond the memory of a single node, HPC implementations distribute the state vector across many nodes [352–355]. Depth scaling is linear for this approach [350], so deep circuits on modest qubit counts can still be handled efficiently; for instance, low-energy nuclear physics simulations with $\mathcal{O}(10^8)$ gates on 21 qubits have been completed in under an hour on GPU accelerators [356].

Approximate state-vector methods relax accuracy requirements to achieve larger simulated qubit counts. Truncated amplitude representations have been used to simulate chemistry-oriented VQE instances on ∼64 qubits with negligible loss in accuracy [298, 357]. Hybrid strategies such as Feynman path summation can also trade depth for qubit count [350, 358].

2.4.2.1 Density matrix simulators

For open systems or noise modeling, the pure-state description is replaced by the density operator

$$\rho = \sum_s p_s |\psi_s\rangle\langle\psi_s|$$

which encodes a statistical mixture of pure states. This formalism can represent both classical and quantum uncertainty and is well suited for simulating decoherence and dissipation.

The computational overhead is significant: an n-qubit density matrix has 4^n complex entries, leading to 2^n times the memory footprint of the equivalent state-vector simulation. Under a gate G, the evolution is

$$\rho(t) = G(t)\rho(0)G(t)^\dagger \tag{2.9}$$

with noise channels represented via Kraus decompositions. The general n-qubit G is constructed from tensor products of the active gate with identities on unaffected qubits.

This approach is indispensable for high-fidelity noise studies and is often combined with stochastic unraveling methods to reduce computational demands in large-n simulations [359].

2.4.2.2 Tensor network simulators

Tensor network (TN) methods [360–366] represent quantum states as networks of tensors connected along shared indices, exploiting entanglement structure to reduce storage requirements.

Different architectures target different problem geometries:

- **MPS** — Efficient for 1D systems or low-entanglement circuits, capturing short-range correlations in a scalable manner.

- **PEPS** — Extends MPS concepts to higher dimensions; suited for 2D materials and lattice models, though computationally heavier.
- **TTN** — Hierarchical structures useful for circuits with layered entanglement or natural tree-like decompositions.
- **MERA** — Designed for critical systems with scale-invariant entanglement, common in phase transition studies.
- **CVTN** — Adapted for continuous-variable systems, important in photonics and field theory.

With truncation of bond dimensions, TN simulators have reached qubit counts in the 200 range for certain circuit classes [362]. Approximation quality depends on entanglement growth: shallow or low-entanglement circuits are tractable, while generic deep circuits remain challenging. These methods will be particularly relevant in Chapter 5 where we utilize a truncated MPS simulation to achieve results for higher orbital counts.

2.4.2.3 Open-system Lindblad simulators

When the interaction between a quantum system and its environment cannot be neglected, Lindblad master equations offer a compact formalism for non-unitary dynamics. The equation combines Hamiltonian evolution with dissipative terms modeling decoherence and noise processes.

Open-system simulators based on Lindblad dynamics are valuable for investigating relaxation, dephasing, and transport in condensed matter systems, as well as for validating error mitigation in realistic device settings. Recent work incorporates noise directly into the simulated gates [367], an approach extensible to non-Markovian noise models via colored-noise processes.

By enabling direct simulation of decoherence pathways and environment-induced dynamics, these tools bridge the gap between idealized quantum circuits and laboratory-scale quantum devices.

In summary, classical quantum-circuit simulation spans a spectrum from exact state-vector and density-matrix approaches to approximate tensor-network and Lindbladian methods for open systems. Each offers different trade-offs in scalability, accuracy, and ability to capture noise effects. For this thesis, we primarily employ state-vector, density-matrix, and tensor-network (specifically MPS) methods, as they best capture the dynamics and system sizes relevant to our materials-science workloads.

2.5 Use-Case Identification

2.5.1 Criteria for selecting relevant use cases

Identifying where quantum computing could provide meaningful advantages in real-world materials discovery requires a clear, structured set of selection criteria [368]. The

goal is not to compile an exhaustive catalogue, but rather to pinpoint use cases that both address pressing scientific or industrial needs and exhibit computational characteristics that make them plausible candidates for quantum acceleration.

Three high-level attributes shape our selection process:

1. **Potential impact.** We prioritize application domains where improvements in accuracy, throughput, or insight could significantly influence technology development. Representative targets include semiconductor integration and fabrication, advanced electronics and optoelectronics, carbon capture and storage, catalysis, energy storage, materials recycling, and emerging biopharmaceutical applications. The nature of "impact" varies: in some cases it stems from achieving unprecedented resolution in the description of underlying mechanisms; in others, from enabling high-throughput screening to rapidly isolate the most promising material candidates.

2. **Relevant physical scales.** Practical materials design problems rarely reside at a single scale. Instead, they demand simulations spanning length and time scales over more than ten orders of magnitude. A comprehensive workflow may involve quantum-mechanical electronic structure calculations, molecular dynamics at the atomistic or coarse-grained level, and macroscopic thermodynamic or continuum models. The increasing integration of machine learning and AI-driven surrogates further expands this toolkit, accelerating generative materials design, large-scale screening, and multi-fidelity simulation.

3. **System complexity.** Many frontier applications involve structurally and chemically complex materials, where critical physical or chemical events occur within regions containing thousands of molecular building blocks and millions of atoms. To make simulation tractable, one must identify the smallest representative domain capable of capturing the essential phenomena. This could range from the active space of a catalyst, to a reactive surface patch, or even to macroscopic engineered structures such as an aircraft composite component.

These attributes help narrow the scope to problems that are both important and challenging, but they do not in themselves determine whether quantum computing is the right tool. Once a candidate domain is identified, a second layer of analysis is needed, one that compares the computational challenge from both the classical and quantum perspectives, and evaluates the feasibility of the required workflows.

For each potential use case, we therefore assess:

1. the significance of the problem in a real-world context,

2. the computational difficulty of the best-known classical simulation methods,

3. the expected resource demands of the quantum formulation, both on near-term devices and in the fault-tolerant regime,

4. the practical challenges of executing the resulting quantum circuits on hardware and in classical simulators, and

5. the potential for hybrid execution strategies that combine HPC resources with quantum processors in a tightly integrated workflow.

Framing the criteria in this two-stage manner, first by identifying impactful, physically relevant, and complex problems, and then by evaluating their computational feasibility from both classical and quantum standpoints, ensures that selected use cases represent not only areas of high scientific and technological value but also realistic opportunities for early quantum advantage.

These criteria guided how we picked embedded fragments and periodic models and shaped the benchmarking/verification baselines introduced in Chapters 5 and 6

2.5.2 Overview of use cases

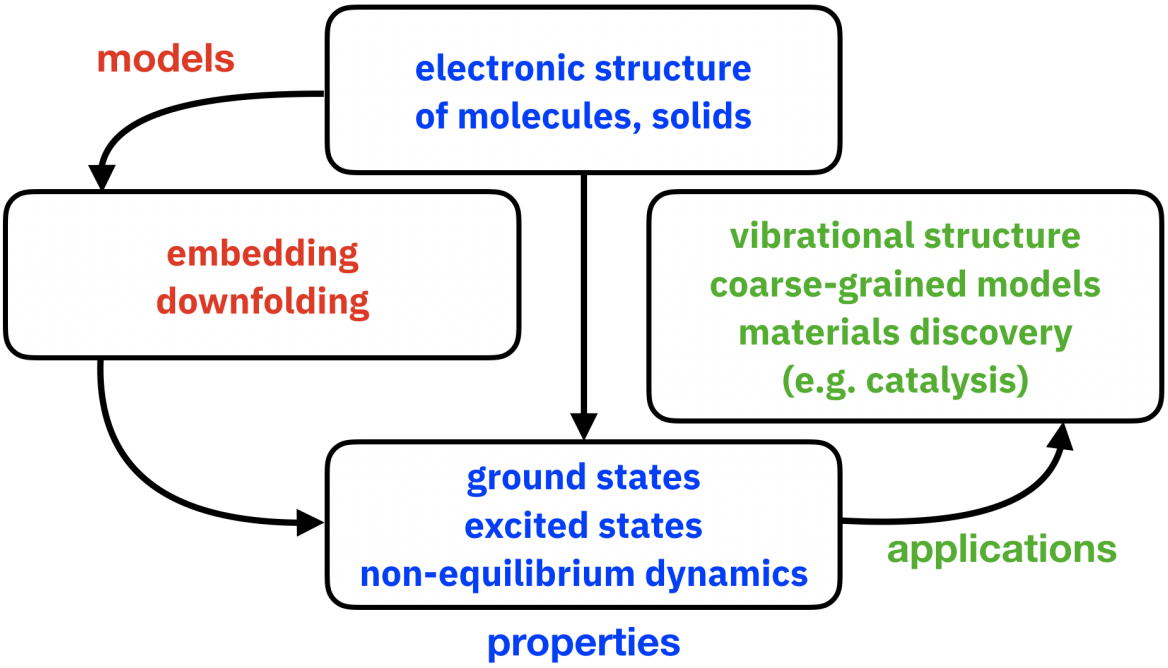


Figure 3 – **Schematic overview of representative use cases in materials discovery.** The central challenge in quantum theory is to determine the properties of molecules and solids (blue). This task can be addressed by solving the electronic Schrödinger equation to compute, for example, ground-state, excited-state, and non-equilibrium dynamical properties (vertical arrow). Because exact solutions are computationally prohibitive, approximations are introduced that reduce the quantum problem to fit within available computational budgets (red). Accurate solutions of the electronic Schrödinger equation serve as the foundation (green) for downstream applications in materials science, including vibrational structure calculations, coarse-grained modeling (e.g., corrosion or deformation), and materials discovery (e.g., catalysis, metamaterials). *Reproduced from Y. Aleexev, et al., [1] (DOI) © 2024 Elsevier. Used with permission.*

In quantum condensed matter physics and quantum chemistry, one of the fundamental objectives is to predict the properties of interacting atomic, molecular, amorphous,

and crystalline systems. Beyond deepening our understanding of quantum mechanics itself, the ability to solve such problems with high fidelity would have direct implications for technological domains as varied as the design of new magnetic materials, high-efficiency photovoltaics, heterogeneous catalysts, and quantum information hardware, as well as pharmaceutical drug discovery.

The computational demands for an accurate *ab initio* treatment of realistic systems, both in terms of qubit count and gate depth, are well beyond the reach of current and near-term quantum hardware, except for very small cases. This has been made explicit in resource estimates by several authors, including Reiher *et al.* [369], Tubman *et al.* [370], Elfving *et al.* [371], and Goings *et al.* [372], who analyzed the use of quantum phase estimation for challenging electronic structure targets such as the nitrogenase cofactor, the chromium dimer, and cytochrome P450 on fault-tolerant devices.

Given these limitations, it is common to introduce approximations that reduce the problem size to a form that can be addressed within available quantum or classical resources (Figure 3, left). Two broad strategies are prevalent:

- **Embedding approaches**, as we saw earlier, isolate a subset of strongly correlated "active" degrees of freedom (often specific orbitals or lattice sites) and formulate an auxiliary quantum problem for these while treating the remainder of the system with a classical approximation. Systematic improvement is possible by enlarging the embedded subspace or by using higher-accuracy classical solvers.
- **Model system approaches** replace the full physical system with a minimal effective low-energy model, then incrementally enrich the model until it reproduces the target physics. For example, the Heisenberg spin model captures the low-energy behavior of the t–J model, which in turn approximates certain regimes of the Hubbard model.

Both methods can yield problems that are much more tractable for quantum hardware than the full electronic structure problem. Typically, embedding techniques produce compact but dense Hamiltonians (few qubits, many operations), while model systems produce larger but sparse Hamiltonians (many qubits, fewer operations). Sparsity can be exploited on devices with limited connectivity to keep gate counts manageable.

The specific physical observables of interest also play a decisive role in defining the computational approach. Standard electronic structure workflows focus on ground-state properties, such as total energies and band structures, from which structural, vibrational, and mechanical characteristics can be inferred. However, real materials are probed at finite temperature and under conditions that reveal excited-state and dynamical properties.

Experimental techniques such as photoemission spectroscopy (single-particle excitation spectra and band gaps) [373], nuclear magnetic resonance and neutron scattering (nuclear spin excitations), and optical conductivity measurements (two-particle current–current correlations) require access to excited states. Traditional density functional theory (DFT) often needs additional approximations to provide these quantities, and DFT-derived band gaps are notoriously unreliable for semiconductors [374–376]. Quantum methods that directly encode excitation spectra could therefore complement and extend existing computational tools.

In many experimental scenarios, an external perturbation probes the material. If the perturbation is weak, linear response theory provides an accurate description. Strong perturbations, however, drive the system far from equilibrium and require time-dependent modeling. Examples include quenches (instantaneous changes to the Hamiltonian), steady states after bias application, and periodically driven systems such as those illuminated by continuous-wave lasers.

While classical simulation of ground-state physics is well established, excited-state and real-time dynamics remain computationally intensive, often pushing the limits of current HPC resources. Accurately modeling such nonequilibrium phenomena is still an emerging capability and thus presents a promising application space for quantum computing.

Finally, reliable solutions to the electronic structure problem also enable downstream calculations: solving the nuclear Schrödinger equation for vibrational modes, developing coarse-grained models to describe mesoscale phenomena, and predicting macroscopic material behavior such as corrosion, mechanical stress, and fracture propagation (Figure 3, right). These multiscale connections reinforce the importance of improving electronic structure capabilities whether through classical acceleration, quantum algorithms, or hybrid approaches.

This motivates our focus on (i) ground-state energetics for embedded fragments (Chapter 5) and (ii) spectral/band-gap proxies in periodic models (Chapter 6), executed with pre-fault-tolerant sampling methods from Chapter 4.

2.5.3 Relevant use cases

Some of the use cases are not relevant to our work, but still, they are noted here because of their importance and potential in near-term quantum computing for materials science.

2.5.3.1 Ground-state electronic structure

A central objective of electronic structure theory is to determine the ground and low-lying eigenstates of a molecular Hamiltonian, typically within the Born–Oppenheimer approximation at fixed nuclear coordinates $\mathbf{R}$. Mapping out the potential energy surface $E(\mathbf{R})$ is essential for understanding reactivity, reaction rates, and product distributions. In particular, predicting reaction barrier heights with ~ 1 kcal/mol accuracy relative to experiment is critical for applications such as catalyst discovery, battery electrode modeling, and process design, yet it still remains challenging for standard approaches, especially DFT [376].

As an illustration, process design often seeks reactant sets for which desired reactions have lower barriers than competing pathways. Experimental screening across many candidate reactants can be slow, costly, and hazardous. Computation promises orders-of-magnitude gains in speed, safety, and cost for barrier screening, but state-of-the-art classical methods are frequently not accurate enough to resolve small barrier differences.

The accuracy target for reaction barriers (∼1 kcal/mol) and multireference challenges justifies our embedded-fragment studies on Chapter 5 and the shift from pure VQE baselines (Chapter 3) to sample-based diagonalization (Chapter 4).

Static and dynamic correlation

Electronic wavefunctions are commonly categorized as single-reference or multireference. Single-reference cases are dominated by one electronic configuration; most molecular ground states fall in this category. Multireference character becomes prominent for excited states, stretched bonds, transition-metal chemistry, or small HOMO–LUMO gaps (e.g., extended π systems).

Single-reference problems are well served by classical quantum chemistry, from DFT (routinely $O(10^3)$ atoms [20]) to coupled-cluster methods (kcal/mol accuracy at $O(10)$ atoms [377]). By contrast, extending high accuracy to genuinely multireference systems remains difficult; despite progress, accuracy for systems larger than $O(1)$ atoms is typically far lower than in the single-reference regime.

Transition-metal chemistry provides many multireference exemplars [69, 378]. Transition-metal oxides underpin catalysis, semiconductor technology, pigments, and beyond, yet DFT benchmarks for their bond dissociation energetics show mean absolute errors of $\gtrsim$30 kJ/mol (7 kcal/mol) [379], underscoring the challenge for DFT in transition-metal energetics. Representative hard instances include:

1. *Enzyme active sites with transition metals*, often with multiple coupled centers (Fig. 4):
   - Fe_2S_2 and Fe_4S_4 clusters in ferredoxins [380], with active spaces of (30e,20o) and (36e,54o), respectively [381, 382];
   - the FeMo cofactor of nitrogenase (biological N_2 reduction) [369, 383], modeled with a (113e,76o) active space [384, 385];
   - the nitrogenase P-cluster [386] and the O_2-evolving complex of photosystem II [387, 388]. For the P-cluster, relevant charge states (P^N, P^{1+}, P^{OX}) correspond to (114e,73o), (117e,75o), and (120e,77o) active spaces, respectively.

   These active sites motivate highly multireference treatments to disentangle spin coupling and metal–metal delocalization essential for interpreting spectroscopy [380, 386, 388–390], beyond what DFT-based studies alone capture [391–393]. The iron clusters example is particularly relevant as it was recently studied using sample-based quantum algorithms [42]

2. *Nanocatalysts and surfaces with transition metals.* Mechanistic modeling for heterogeneous catalysts is comparably demanding. Even single-reference treatments at accuracy well beyond DFT are difficult, while DFT adsorption energies can be unreliable [394, 395]. Multireference effects are expected in various oxide catalysts and along reaction paths [394, 396–398].

Across these systems, the core computational task is obtaining the ground (or low-lying) eigenstates of the electronic Hamiltonian. A primary benchmark is whether

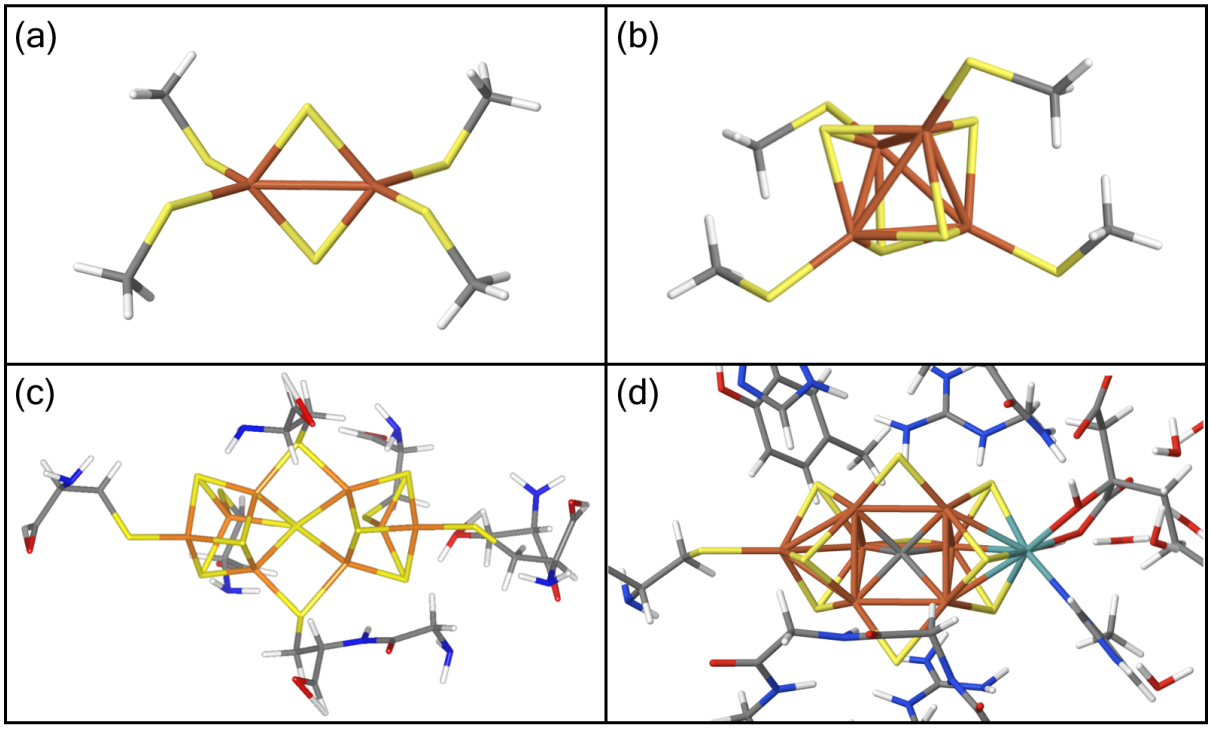


Figure 4 – Fe_2S_2 (a) and Fe_4S_4 (b) clusters in ferredoxins; the nitrogenase P-cluster in the resting state P^N (c); and the FeMo cofactor (d). Orange, yellow, teal, red, blue, gray, and white sticks denote Fe, S, Mo, O, N, C, and H, respectively. *Reproduced from Y. Aleexev, et al., [1] (DOI) © 2024 Elsevier. Used with permission.*

quantum algorithms can deliver more accurate energies (for comparable resources) than the best classical methods. Proof-of-principle studies can target the active spaces listed above, but realistic chemistry ultimately requires (i) larger electron–orbital spaces, (ii) environment, dynamics, and solvation effects, and (iii) dynamical correlation. Notably, even nominally single-reference systems become classically challenging at large scale [399].

Relativistic effects

Many systems (e.g., Au atoms, U dimers) demand explicit relativistic corrections to Hamiltonians to achieve accurate ground and excited states. Omitting relativistic effects imposes a fundamental precision limit on quantum (and classical) predictions [400].

Correlated electronic structure in solids

In crystalline materials, weakly interacting electrons are well described by band theory and DFT, but strong correlations can invalidate the non-interacting picture. Mott insulators, for example, are insulating due to interactions despite band-theory metallicity. While Mott physics is well established, other strongly correlated phenomena still lack predictive microscopic theories. Two prominent examples are:

- *High-T_c superconductivity.* The mechanisms—especially the roles of the pseudogap and strange-metal regimes—remain unsettled [401–403].
- *Frustrated spin systems.* These host topological and gapless spin liquids [404] and have long served as benchmarks for advanced numerical methods, making them attractive for quantum simulation.

2.5.3.2 Embedding

Highly relevant real-world instances of the electronic structure problem (e.g., catalysis) require substantial quantum resources, creating tension between realism and near-term feasibility. Quantum embedding and effective low-energy modeling offer practical paths forward that are also compelling near-term quantum applications.

In embedding, the system is partitioned into an active space plus an environment. The environment is treated at a lower level of theory, while the active space receives a higher-accuracy (and more expensive) treatment. Because the active space spans a much smaller Hilbert space, quantum simulation becomes feasible on quantum hardware or high-end emulators. Typical active spaces include point defects in semiconductors/insulators and catalytic active sites at surfaces or interfaces.

Representative embedding strategies include:

- **Density-based embedding**: define a spatial active region; treat the environment with DFT and the active region with higher-level DFT (DFT-in-DFT) or wavefunction methods (WF-in-DFT) [405–408].
- **Density matrix embedding theory (DMET)** [53, 265–275]: define a spatial fragment; couple to bath orbitals obtained from a lower-level calculation and enforce consistency via a one-body potential.
- **Green's function–based embedding**: define an active subspace in terms of single-particle states and partition the nonlocal, frequency-dependent self-energy into high- and low-level components with a double-counting correction (e.g., DMFT+DFT [409, 410], DMFT+GW [411–416], SEET [417–419], and QDET [259, 260, 420, 421]).

Practically, embedding workflows begin with a low-level calculation of the full system to parameterize the many-body effective Hamiltonian for the active space. In QDET, this Hamiltonian includes (i) a one-body term with an exact double-counting correction for G_0W_0 as the low-level theory, and (ii) a two-body term incorporating environmental screening. QDET avoids explicit sums over empty states, enabling large-scale applications. With GPU-accelerated classical HPC, systems of $> 10^3$ atoms and active spaces of $> 10^2$ orbitals are feasible. The resulting second-quantized Hamiltonian can be solved with classical FCI or mapped to qubits for quantum eigensolvers [259, 294, 422–428]. Remaining challenges include high-throughput data movement between HPC and quantum devices, fitting gate counts and depths to hardware constraints, and noise mitigation to prevent unphysical outcomes. Notably, quantum computers themselves are used to probe spin defects in solids (promising building blocks for improved quantum hardware) creating a productive feedback loop between materials and quantum computing.

Recent applications include the previoulsy mentioned methods CAS-DMET and NEVPT2-DMET applied to color-center and oxide defects: NV^{-1} and SiV^0 in diamond, and OV^0 in MgO bulk and surface [253–255]. For NV^{-1}, NEVPT2-DMET with a reduced embedding and CAS(10e,9o) yields a triplet–triplet excitation of 2.31 eV (vs. 2.18 eV experiment) and a dark singlet–singlet at 1.02 eV (vs. 1.26 eV). For bulk OV^0 in MgO, NEVPT2-DMET with CAS(2e,5o) predicts 5.24 eV, close to the 5.03 eV absorption maximum. Surface OV^0 excitations have also been treated with CAS-DMET, NEVPT2-DMET, and DME-PDFT, though reported experimental values vary widely (1–5 eV). Using MP2 and CCSD solvers within DMET, CO adsorption energies on MgO computed with the largest embeddings agree with non-embedded references within 1.2 kcal/mol [429]. The LASUCC and LASSI methods have been benchmarked on strongly correlated molecules, including the dissociation of $(H_2)_2$, bond breaking in trans-butadiene, and J-coupling in a tris-(μ-hydroxo)-bridged Cr complex, showing excellent agreement with corresponding CASCI using LAS orbitals [283, 284].

As noted before, embedding is the workhorse interface that yields the reduced Hamiltonians we solve in Chapters 5, setting sizes and couplings for our quantum runs, which feed directly into the resource analyses.

2.5.3.3 Model systems

Low-energy models replace the full electronic problem with effective Hamiltonians for relevant low-energy degrees of freedom, which are then solved for ground and excited states. The real-world utility of such models depends on (i) proposing the right effective form with reliable parameterization and (ii) solving it accurately enough that qualitative and quantitative conclusions are robust to small parameter perturbations and solution errors.

Band theory (and Fermi-liquid phenomenology for metals) is the standard effective description. When it fails, as in high-T_c cuprates or spin-liquid candidates, richer models (e.g., Emery or Heisenberg-type) are invoked. These can illuminate physics but also introduce free parameters whose values are uncertain (often inferred from experiment or *ab initio* downfolding), potentially biasing conclusions. Accurate yet tractable models that admit large-scale simulation are therefore highly desirable.

Downfolding most commonly proceeds from DFT: hopping integrals via Wannier projections [430], and screened interactions via constrained DFT or RPA, with double-counting corrections to remove interactions already included in DFT [431–433, 433]. Alternatives include Löwdin partitioning (potentially with stochastic elements) [434, 435] and canonical transformations [436, 437]. Despite their advantages, it is often hard to quantify whether a chosen model ansatz and its parameters are *a priori* adequate.

In what follows we highlight two widely studied classes: spin Hamiltonians and fermionic lattice models.

Spin Hamiltonians

Spin models capture magnetic properties and broader correlated/ frustrated phenomena. A lattice with n sites and spin-s degrees of freedom has $(2s+1)^n$ configurations. Thermal agitation yields disordered moments at high T, whereas as $T \to 0$ many systems develop long-range order (ferromagnetic or antiferromagnetic). By analogy with gases (paramagnets) and solids (ordered magnets), one may ask whether a "spin liquid" exists.

Quantum spin liquids [438–443] lack magnetic order at any finite T yet feature strong entanglement and quantum fluctuations. Frustration (geometric and/or interaction-driven)generates a large manifold of near-degenerate configurations, suppressing order. Low-spin systems (e.g., $s = \frac{1}{2}$) favor quantum behavior since energy barriers between microstates scale as s^2.

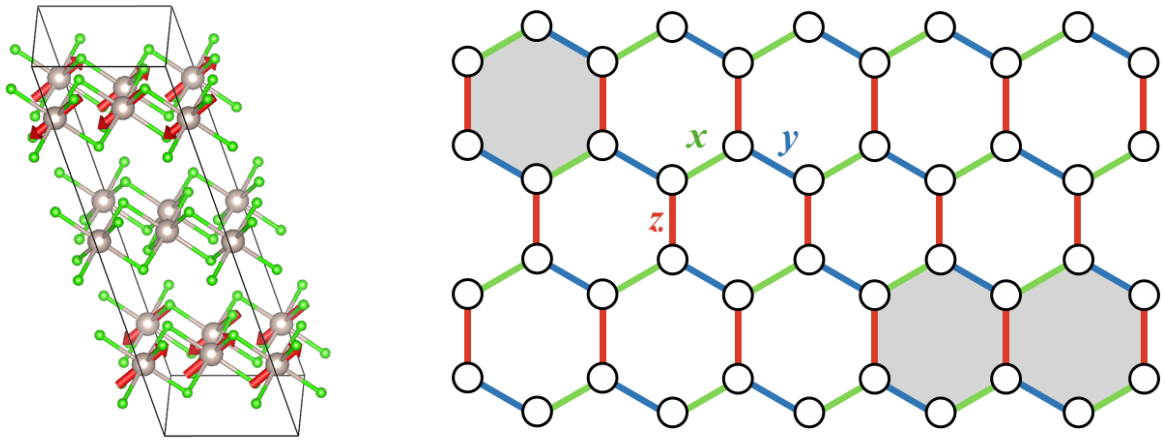


Figure 5 – Left: crystal structure of α-$RuCl_3$, showing the van der Waals gap between $RuCl_6$ honeycomb layers (Ru/Cl as gray/green) and a zigzag-ordered configuration (red arrows). Right: the 2D Heisenberg–Kitaev model; spins are white circles and $x/y/z$ bonds are green/blue/red. Current quantum simulations target 1–2 hexagons (6–10 spins) [2]. *Reproduced from Y. Aleexev, et al., [1] (DOI) © 2024 Elsevier. Used with permission.*

A prominent proximate spin-liquid material is α-$RuCl_3$ (Fig. 5). Stacked honeycomb layers of edge-sharing $RuCl_6$ octahedra exhibit strong spin–orbit coupling, enabling a $j = \frac{1}{2}$ description of Ru^{3+} [444]. The low-energy physics is commonly modeled by the Kitaev–Heisenberg Hamiltonian

$$H_{kh} = \sum_{\langle ij \rangle_\gamma} J\, \vec{S}_i \cdot \vec{S}_j + K\, S_{i\gamma} S_{j\gamma}, \tag{2.10}$$

where $\langle ij \rangle_\gamma$ runs over bonds $\gamma \in \{x, y, z\}$ [445, 446]. Although α-$RuCl_3$ orders in a zigzag pattern at low T [447, 448], Raman and inelastic neutron scattering reveal continua suggestive of fractionalized excitations [447, 449]. Realistic descriptions augment H_{kh} with off-diagonal couplings,

$$\begin{aligned} H = H_{kh} &+ \sum_{\langle ij \rangle_\gamma} \Gamma (S_{i\alpha} S_{j\beta} + S_{i\beta} S_{j\alpha}) \\ &+ \sum_{\langle ij \rangle_\gamma} \Gamma' (S_{i\gamma} S_{j\alpha} + S_{i\gamma} S_{j\beta} + S_{i\alpha} S_{j\gamma} + S_{i\beta} S_{j\gamma}), \end{aligned} \tag{2.11}$$

with $\{\alpha, \beta, \gamma\} = \{x, y, z\}$, $\{z, x, y\}$, $\{y, z, x\}$ for $\gamma = z, y, x$ [450]. Parameter sets inferred from *ab initio* and phenomenology vary appreciably [447, 450–452], though a common scenario takes $K < 0$ (dominant), $\Gamma > 0$ (significant), and a subleading ferromagnetic J [2, 453]. The exactly solvable Kitaev point ($J = \Gamma = \Gamma' = 0$) provides a useful reference; echoes of its two Majorana fermion species may persist in α-$RuCl_3$ observables [454, 455]. These 2-local spin-$\frac{1}{2}$ Hamiltonians are natural for quantum hardware and motivate studies of specific heat and dynamical structure factors.

Lattice models

Competing orders, e.g., metal–insulator [456], para- to ferromagnet [457], and normal- to superconducting [458, 459] are hallmarks of correlated materials. Canonical microscopic descriptions are provided by fermionic lattice models such as the one-band Hubbard model [460–463]:

$$\hat{H} = -t \sum_{\langle ij \rangle, \sigma} \left(\hat{a}^\dagger_{i\sigma} \hat{a}_{j\sigma} + \text{h.c.} \right) + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}. \tag{2.12}$$

Despite its simplicity, the model hosts interaction-driven insulating phases, magnetism, and superconductivity, with details highly sensitive to parameters. Exact solutions exist in 1D [464] and $d \to \infty$ [465]; series expansions are exact only at elevated temperatures [466]. At relevant regimes in 2D/3D, controlled approximations (variational, diagrammatic, and embedding methods) are used [467], with ongoing community benchmarks [467, 468].

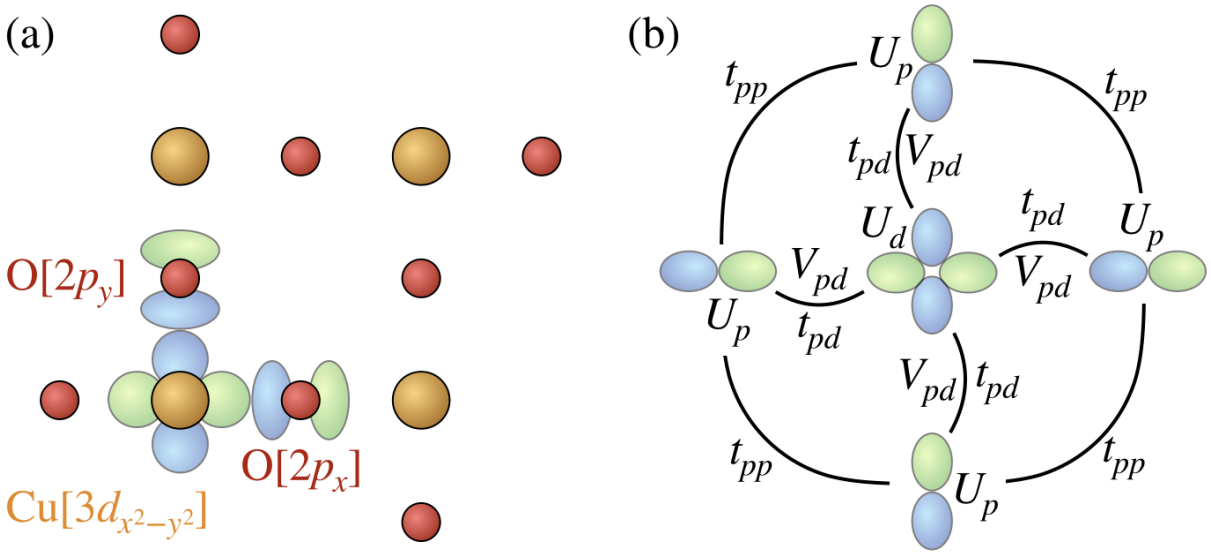


Figure 6 – (a) Schematic of a $[CuO_2]_4$ cell in a cuprate CuO_2 plane (Cu/O shown as orange/red; orbital phases in blue/green). (b) Parameters of the three-band Hubbard (Emery) model (curved connectors depict kinetic couplings). *Reproduced from Y. Aleexev, et al., [1] (DOI) © 2024 Elsevier. Used with permission.*

A technologically relevant extension is the two-dimensional three-band (Emery) model [469], believed to capture essential cuprate physics arising from Cu $d_{x^2-y^2}$ and O

$p_{x,y}$ orbitals (Fig. 6). In the "phase representation,"

$$\begin{aligned}\hat{H} = t_{pd} \sum_{\langle ij\rangle,\sigma} \left(\hat{d}^{\dagger}_{i\sigma}\hat{p}_{j\sigma} + \text{h.c.}\right) + t_{pp} \sum_{\langle ij\rangle,\sigma} \left(\hat{p}^{\dagger}_{i\sigma}\hat{p}_{j\sigma} + \text{h.c.}\right) \\ - \Delta_{pd} \sum_{i,\sigma} \hat{n}^{d}_{i\sigma} + U_d \sum_{i} \hat{n}^{d}_{i\uparrow}\hat{n}^{d}_{i\downarrow} + U_p \sum_{i} \hat{n}^{p}_{i\uparrow}\hat{n}^{p}_{i\downarrow} \\ + V_{pd} \sum_{\langle ij\rangle,\sigma,\tau} \hat{n}^{p}_{i\sigma}\hat{n}^{p}_{j\tau},\end{aligned} \tag{2.13}$$

where $d^{(\dagger)}_{i\sigma}/p^{(\dagger)}_{i\sigma}$ create/annihilate a hole on Cu(3d)/O(2p) orbitals, $\hat{n}^{d/p}_{i\sigma}$ are number operators, and Δ_{pd} is the O–Cu charge transfer gap.

Classically, the Emery model has been tackled by exact diagonalization (up to 24 sites) [470], QMC and DMRG (80–144 sites) [471, 472], and DMET embeddings (12 sites) [473], targeting spin–spin and d-wave correlations to assess stripes and superconductivity [470–473]. Unlike the one-band case, consensus on key aspects is still lacking, opening the door for high-accuracy quantum algorithms to study $O(10^2)$-site lattices.

Although the one-band Hubbard model captures much cuprate phenomenology (e.g., d-wave pairing, charge/spin orders, pseudogap, stripes), there are strong reasons to work directly with the three-band model: (a) possible loss of oxygen-related physics in one-band reductions (e.g., pseudogap signatures) [474]; (b) parameter degeneracies in the one-band model [468] that may be lifted by the additional degrees of freedom; and (c) explicit atomic structure and closer ties to nanoscale experiments [475–477].

Lattice models provide (i) structured benchmarks for our sampling pipelines (Chapter 4) and (ii) the tight-binding/extended Hubbard backdrop used for band-gap estimation in periodic materials (Chapter 6).

2.5.4 Excited states

In many condensed-matter experiments, the focus is on probing dynamical rather than purely equilibrium properties. Techniques such as neutron or photon scattering, and angle-resolved photoemission spectroscopy (ARPES) [478], are commonly employed to investigate quantities like the structure factor or the spectral function [479, 480]. A typical setup involves preparing the system in thermal equilibrium at an inverse temperature β under a Hamiltonian $\hat{H}$, then applying a time-dependent perturbation $\hat{V}(t)$. Within the framework of linear response theory, the measured signal is often expressed as a frequency-dependent correlation function,

$$C_{AB}(\omega) = \int_{-\infty}^{\infty} \frac{dt}{2\pi}\, e^{i\omega t}\, \mathrm{Tr}\left[\hat{\rho}(\beta) e^{it\hat{H}} \hat{A} e^{-it\hat{H}} \hat{B}\right], \tag{2.14}$$

where $\hat{A}$ and $\hat{B}$ are appropriate operators, such as dipole moment components in optical experiments or density operators at different sites for neutron scattering. Notably, in the linear response regime, the perturbation $\hat{V}(t)$ does not explicitly enter $C_{AB}(\omega)$; it is determined entirely from the eigenpairs of $\hat{H}$ or its time-evolution operator $e^{-it\hat{H}}$.

Beyond spectral characterizations, non-equilibrium real-time dynamics has emerged as an area of active investigation, both for its experimental accessibility and its fundamental connection between quantum mechanics and statistical physics. Advances in

ultrafast probes, including free-electron lasers [481, 482] and pulsed laser systems, have enabled pump–probe spectroscopy [483, 484] and similar techniques to resolve quantum processes on femtosecond timescales. These experiments motivate theoretical calculations of time-dependent expectation values of the form

$$B(t) = \mathrm{Tr}\Big[\hat{\rho}(\beta)\hat{U}(t,0)^\dagger \hat{B}\hat{U}(t,0)\Big],$$

where $\hat{U}(t,0)$ evolves the state under $\hat{H} + \hat{V}(t)$. Representing $\hat{U}(t,0)$ as a time-ordered exponential along the Keldysh contour leads to

$$\begin{aligned} B(t) = \sum_{n,m=0}^{\infty} i^{n-m} \int_0^t dr_n \dots dr_1 \int_0^t ds_m \dots ds_1 \\ \times \mathrm{Tr}\Big[\hat{X}(r_n \dots r_1)^\dagger \, \hat{B} \, \hat{X}(s_m \dots s_1)\Big], \end{aligned} \tag{2.15}$$

with $r_1 < \dots < r_n$ and $s_1 < \dots < s_m$, and

$$\begin{aligned} \hat{X}(s_m \dots s_1) = \hat{U}_0(t, s_m)\, \hat{V}(s_m)\, \hat{U}_0(s_m, s_{m-1}) \dots \\ \dots \hat{U}_0(s_2, s_1)\, \hat{V}(s_1)\, \hat{U}_0(s_1, 0), \end{aligned} \tag{2.16}$$

where $\hat{U}_0(t,t') = e^{-i(t-t')\hat{H}}$.

Both frequency-domain correlation functions and real-time non-equilibrium quantities are of high physical relevance, but they are computationally demanding in the classical setting due to the need for explicit time evolution (or, equivalently, access to excited states). These challenges make them strong candidates for fault-tolerant quantum computation, which can implement time evolution operators with polynomial overhead and systematically controllable approximations.

While excited-state/real-time response is not a central deliverable here, the correlation-function and short-time evolution formalism informs our Krylov/sampling discussion (Chapter 4).

2.5.5 Applications of electronic structure

Vibrational structure calculations

Accurately predicting molecular vibrational spectra is computationally demanding on classical hardware. While obtaining reliable electronic structure information is foundational in quantum chemistry and materials design, many practical and scientific objectives require going further to build kinetic models grounded in a detailed vibrational description. From such descriptions, one can extract thermodynamic quantities relevant to, e.g., atmospheric chemistry, catalysis, and combustion modeling. Although small-molecule electronic structures are often tractable with classical methods, high-fidelity vibrational calculations that go beyond the harmonic approximation become challenging even for modest systems. The difficulty arises from anharmonicity and mode–mode coupling among bosonic degrees of freedom, which can be captured by a quartic force–field Hamiltonian [485]:

$$\begin{aligned} H_{\mathrm{anharm}} = \frac{1}{2}\sum_{i=1}^{M} \omega_i \left(q_i^2 + p_i^2\right) + \sum_{\{ijk\}} h_{ijk}\, q_i q_j q_k \\ + \sum_{\{ijkl\}} h_{ijkl}\, q_i q_j q_k q_l + \cdots, \end{aligned} \tag{2.17}$$

where the first term is the harmonic approximation and the remaining terms encode anharmonic corrections via force constants h.

Quantum algorithms for vibrational structure remain comparatively underexplored [321, 368, 486]. Both near-term and fault-tolerant approaches have been proposed, but making them practical requires reductions in qubit counts and circuit depths, along with problem-specific optimizations. Resource studies suggest that achieving quantum advantage for vibrational problems may require fewer total resources than for electronic structure [486, 487]. One reason is that the vibrational qubit Hamiltonian tends to produce more localized Pauli operators, which can lower simulation cost; moreover, more Pauli strings can sometimes be executed in parallel within a single Trotter step than in analogous electronic mappings.

Two complementary strategies can enhance vibrational predictions. The first uses quantum computers to improve the accuracy of the electronic potential energy surface (PES), which is especially effective for relatively rigid molecules whose vibrational spectra are otherwise well-captured by classical methods. The second targets the vibrational problem itself on a quantum device, assuming an already accurate PES (this is particularly important for floppy molecules and for strong anharmonic effects). A pertinent industrially motivated example is polyyne molecules, which feature alternating C–C single and triple bonds, high structural symmetry, and numerous IR/Raman-silent modes. Reliable characterization of their low-frequency vibrations is key to understanding reactivity in processes such as floating catalytic CVD for carbon nanotube growth; recent work provides concrete quantum resource estimates for polyynes [486].

A recent demonstration of the first strategy—hybrid quantum–classical enhancement of vibrational simulations—computes a LiH PES on quantum hardware, then fits analytic forms and classically computes vibrational levels to obtain thermodynamic properties [488]. While successful for the small test system, scaling to more vibrational degrees of freedom will require innovations to avoid combinatorial sampling of molecular distortions.

A classic multiscale setting that couples electronic, vibrational, and transport phenomena is photosynthesis. Light harvesting converts photons into stored chemical energy via absorption and subsequent exciton generation, transport, transduction, and dissipation, with photochemical yields reported up to ∼85% in some subsystems [489]. Recent studies suggest alternate dissipative pathways that mitigate photodamage [490]. Multiscale modeling integrates simplified vibrational descriptions with exciton-transfer dynamics and reactive quantum chemistry [491]. In photosystem II, for instance, the key chemistry involves using an excited electron to drive water oxidation in a compact Mn/Ca/O active-site cluster; despite decades of study, essential mechanistic details remain unsettled [492].

Photosynthetic complexes are well-modeled as open quantum systems interacting with structured vibrational environments. Classical simulations of such dynamics often scale exponentially with system size and environmental complexity, restricting numerically exact treatments to small systems (fewer than ∼20 sites) and simplified baths. Figure 7 summarizes the practical limits reached with hierarchical equations of motion on contemporary supercomputing resources [3]. Proposals to overcome these barriers on quantum hardware range from gate-based and hybrid schemes to analog simulators (see, e.g., [493]). Given current device limitations (e.g., circuit depth), near-term progress will likely hinge

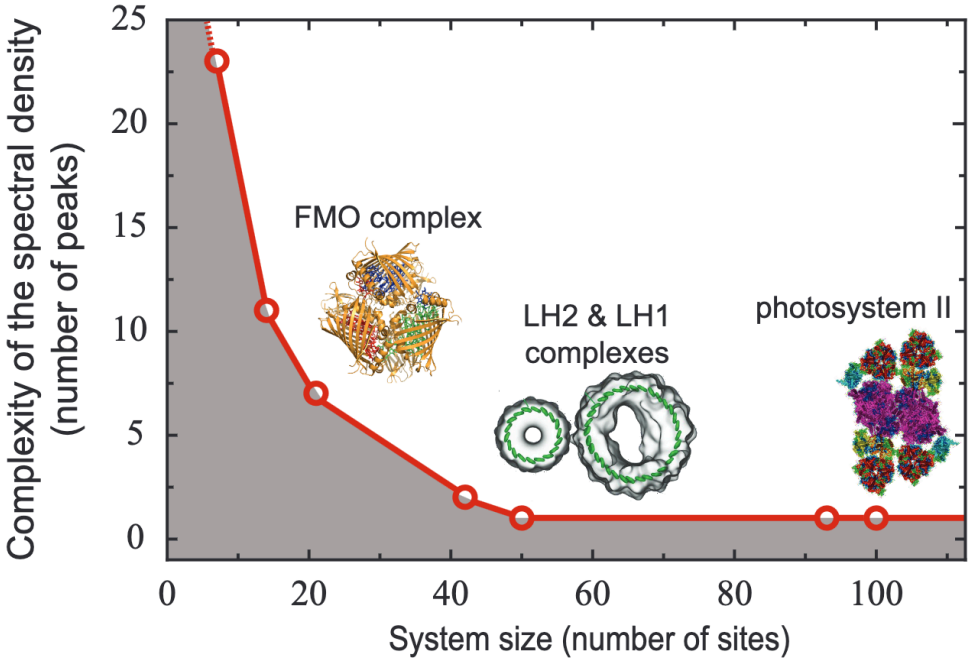


Figure 7 – **Estimated classically tractable sizes for Frenkel exciton Hamiltonians using hierarchical equations of motion on 64 AMD Opteron cores with 250 GB RAM.** Three photosynthetic complexes are indicated: Fenna–Matthews–Olson (Green Sulfur Bacteria), Light Harvesting I/II (Purple Bacteria), and Photosystem II (higher plants). Adapted from [3].*Reproduced from Y. Aleexev, et al., [1] (DOI) © 2024 Elsevier. Used with permission.*

on hybrid algorithms and aggressive resource-optimization techniques.

Vibrational algorithms are not executed in this work; we include them to scope potential near-term advantage and to mirror the resource-comparison logic used for electronic structure in Chapters 5 and 6. Sampling based algorithms for vibrational structure problems is an active area of research with near term impact potential.

Coarse-grained simulations

Translating advanced materials from concept to market typically spans 10–20 years [494]. Much of this cost and delay stems from iterative experimental workflows: preparing high-quality samples, handling hazardous or expensive precursors, and performing extensive characterization only to find many candidates fall short of specifications. Virtual exploration can compress this cycle by shifting experiments from broad exploration to targeted confirmation.

A common computational discovery workflow proceeds in two stages: (i) generate

or identify candidates (often via machine learning), and (ii) filter them by predicting target properties. Two demand signals recur: (a) discovery of materials that are simultaneously light and strong; and (b) durability in the presence of defects (e.g., cracks, corrosion).

For (a), stability screening tests whether a predicted phase is robust to small pressure/temperature changes or phase separation. Standard tools include DFT-based nudged elastic band calculations to estimate reaction barriers and phonon-spectrum analyses to diagnose dynamical stability; see, e.g., the state-of-the-art battery of tests in [495].

For (b), defect modeling spans length scales from atomistic crack tips and corrosion pits to macroscopic failure. Embedding methods are central: a high-accuracy treatment is applied in a strategically chosen active region and consistently matched to a coarser description outside. The fidelity and cost depend on how the active space and its coupling to the environment are defined. If the active space is treated with many-body wavefunction methods, consistently coupling to finite-temperature, coarse-grained phonon descriptions can be nontrivial. Alternatively, if the active space captures precise T=0 atomic configurations, one can "ramp up" thermal vibrations to interface with continuum or mesoscale models. When defects span multiple scales, a recursive, multiscale embedding hierarchy can be employed, matching solutions at successive scales.

First-principles molecular dynamics (with DFT forces) is used both to validate multiscale matching and to study mechanisms directly, e.g., rapid crack propagation across scales, or slower wear/corrosion processes where deeper atomistic modeling is feasible (often again using quantum embedding and, where needed, high-level correlated solvers for the active regions). Other slow, macroscopic phenomena such as creep (time-dependent deformation under load) are typically modeled with differential equations; quantum algorithms for linear systems offer potential acceleration when linearized forms arise. Bulk elastic properties (Young's modulus, Poisson's ratio) can also be computed via continuum models governed by differential/integral equations, with lattice models providing microscopic inputs in driven/dissipative settings.

An example use case: quantum metamaterials

The preceding use cases highlight how quantum and classical computation interleave across the materials pipeline. In line with our criteria, successful targets (i) address real applications, (ii) challenge classical methods, (iii) contain subproblems amenable to quantum speedups, and (iv) naturally integrate HPC with quantum resources. Practically, discovery programs evolve from (1) phenomenological modeling to build intuition, to (2) predictive screening, and finally to (3) direct optimization against target properties. Metamaterials provide a timely case study.

Metamaterials are engineered composites with subwavelength structure that enable phenomena such as negative refractive index and spectrally selective reflection [496–499]. These capabilities can enhance thermophotovoltaics and HVAC efficiency, among other applications. Performance depends sensitively not only on constituent materials but also on geometry, yielding high-dimensional, nonconvex design spaces that defy brute-force or trial-and-error search.

Two core questions arise: (i) how to systematically discover high-performing de-

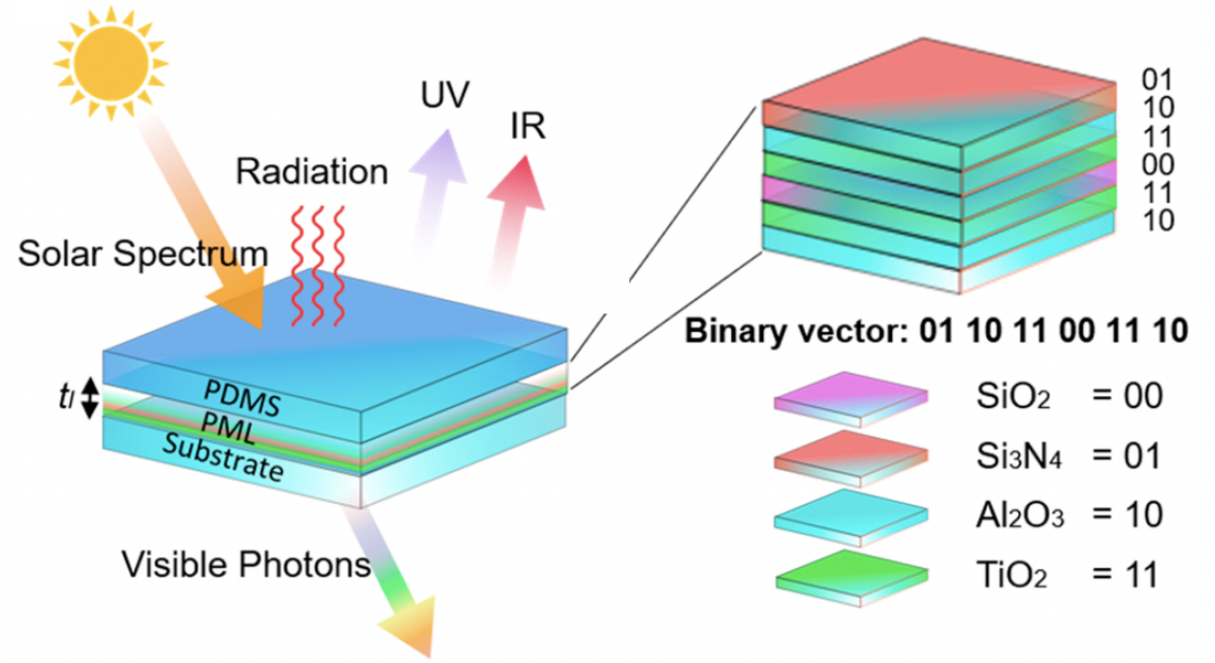


Figure 8 – Schematic of a planar multilayer transparent radiative-cooling metamaterial design that admits a binary encoding of layer choices (adapted from [4]).*Reproduced from Y. Aleexev, et al., [1] (DOI) © 2024 Elsevier. Used with permission.*

signs near the global optimum, and (ii) how to handle complex, mixed discrete–continuous parameterizations. An effective strategy couples active learning with machine learning, quantum optimization, and HPC. Because many metamaterial layouts admit discrete encodings (Fig. 8), the search maps naturally to combinatorial optimization. Quantum algorithms can target such objectives by formulating a Hamiltonian whose ground state encodes the optimal design [500]. Factorization machines, a supervised ML model, compactly capture interactions among design variables and yield bias, linear, and quadratic terms that directly parameterize the Hamiltonian [501, 502]. Iterating between model training and quantum optimization guides sampling toward promising regions of the space.

Current quantum devices are qubit-limited, so large-scale numerical evaluations still rely on parallel classical simulation; integrating HPC for forward solves with quantum subroutines for discrete optimization is therefore natural. To address continuous parameters, recent quantum–classical schemes optimize hybrid discrete–continuous spaces directly [503], broadening applicability to realistic metamaterial design tasks.

2.6 Outlook

Looking ahead, the near– and mid-term path to practical quantum impact in materials science is not a single breakthrough algorithm but the maturation of quantum-centric supercomputing: tightly orchestrated quantum/classical workflows in which HPC performs the heavy lifting before, during, and after focused quantum subroutines. Concretely, this means (i) selecting problems that are classically hard in an appropriate limit, (ii) mapping them to quantum algorithms that are executable on the relevant era of hardware (noisy

or fault-tolerant), and (iii) ensuring the problems matter scientifically or industrially [1]. These principles guide the use cases throughout this chapter and anchor the strategy of the thesis.

On the algorithmic side, the field will likely run a two-track playbook. For fault-tolerant systems, optimal block-encoding and signal-processing techniques (e.g., qubitization) set the asymptotic standard for dynamics and eigenproblems; for pre-FT hardware, structure-exploiting Trotterization and short-time evolution methods (e.g., Krylov builds) remain compelling because they trade depth for accuracy in controllable ways. This thesis aligns with the pre-FT track by advancing sample-based approaches that replace fragile variational training with statistically robust subspace construction, and by pairing these with classical diagonalization and embedding.

Another pillar is classical simulation of quantum circuits, which serves dual roles: approximate verification of quantum results and, crucially, mapping the frontier of classical hardness so that we target genuine quantum advantage regions rather than easy benchmarks. The thesis leverages this by benchmarking sample-based pipelines on classically simulable systems first, then escalating to embedded and periodic models that match or exceed reliable classical baselines.

Putting these strands together, the near-term opportunity space clusters around (a) embedded electronic-structure fragments where quantum resources focus on strongly correlated subspaces; (b) excited-state and finite-temperature response that benefits from short-time evolution and sampling; and (c) workflow acceleration via quantum optimization for discrete design steps. Longer-term, as error correction becomes practical, the same problem classes migrate to deeper time-evolution and phase-estimation primitives, tightening accuracy guarantees without abandoning the quantum-centric workflow stance.

We operationalize the above outlook by (1) replacing fragile variational loops with sample-based diagonalization; (2) integrating these with classical embedding so that quantum effort is reserved for chemically active, correlated subspaces; and (3) extending the methodology to periodic Hamiltonians for band-gap estimation. In short, it instantiates the quantum-centric paradigm advocated by the perspective: algorithmic choices that are executable now, embedded problem formulations that matter scientifically, and HPC-first workflow design to make the whole pipeline runnable at scale.

Building on this, the next chapter turns from strategy to tools. We begin by revisiting the classical electronic-structure toolbox—Hartree–Fock and Density Functional Theory as baselines; post-Hartree–Fock methods such as CCSD(T) for single-reference systems; and multireference and embedding techniques for strongly correlated fragments—emphasizing where cost, accuracy, and systematic improvability collide in realistic materials settings. We then introduce VQE methods and some variants (e.g., ADAPT-VQE and qubit-ADAPT-VQE), outlining measurement overheads, optimizer pathologies, and noise sensitivity that shape practical performance on today's devices. The chapter culminates with a focused case study from our recent work on direct air capture materials, where qubit-ADAPT-VQE is applied to simplified MOF-inspired models to generate potential energy profiles and compare against classical references. The lessons from these experiments, both successes and limitations, motivate the pivot to sampling-based approaches that we discuss in subsequent chapters.

3 Methods for Quantum Simulation of Fermionic Systems

This chapter focuses on providing the theoretical background necessary to understand the context of the project. It goes over Classical methods for computational chemistry, Quantum computational chemistry methods, and a published application of common variational methods in a simplified materials science problem for carbon capture authored by the candidate.

3.1 Classical Computational Chemistry Methods

Classical computational chemistry began with the self-consistent field method introduced by Douglas R. Hartree in the late 1920s to approximate the wave functions and energies of many-body quantum systems, addressing the inability to analytically solve the Schrödinger equation for multi-electron systems [504].

This method evolved into the Hartree-Fock method after corrections by John Clarke Slater and Vladimir Aleksandrovich Fock, incorporating the Pauli Exclusion Principle with the introduction of the Slater determinant [505] [506] [507]. These methods operate within the Born-Oppenheimer approximation, established by Max Born and Julius Robert Oppenheimer, which simplifies quantum chemistry calculations by separating nuclear and electronic wave functions [508].

The computational process relies on the variational principle to minimize energy through iterative calculations [509]. It uses Mean Field Theory, which simplifies many-body problems into a one-body problem by averaging external interactions into a uniform field [510] [511], ignoring electron correlations for tractability.

Increasing problem complexity will lead to the development of post-Hartree-Fock methods like Configuration Interaction, Møller–Plesset perturbation theory, and Coupled Cluster, which consider electron correlations to enhance accuracy [512] [513] [514].

Unlike wave-function methods that vary over many-electron states, alternative approaches such as DFT are also variational but over the electron density via the Hohenberg–Kohn theorems. The unknown is moved into the exchange–correlation functional. This will be discussed further in another section in this chapter.

3.1.1 Second-quantized electronic Hamiltonian

With fixed nuclear coordinates $\mathbf{R}$, the nonrelativistic electronic Hamiltonian in an orthonormal spin–orbital basis $\{\chi_p(x)\}_{p=1}^{K}$ is

$$\hat{H} \;=\; \sum_{pq} h_{pq}\, a_p^\dagger a_q \;+\; \frac{1}{4}\sum_{pqrs} \langle pq\|rs\rangle\, a_p^\dagger a_q^\dagger a_s a_r \;+\; E_{\mathrm{nuc}}, \tag{3.1}$$

where $a_p^\dagger$ and a_p create/annihilate an electron in spin–orbital p, and E_{nuc} is the classical nucleus–nucleus repulsion. Here $x \equiv (\mathbf{r}, \omega)$ denotes the combined spatial and spin coordinate (with $\omega \in \{\alpha, \beta\}$), and the integration measure is

$$\mathrm{d}x \;\equiv\; \mathrm{d}\mathbf{r} \sum_{\omega}.$$

Using $x_1 = (\mathbf{r}_1, \omega_1)$ and $x_2 = (\mathbf{r}_2, \omega_2)$, we define

$$h_{pq} = \int \mathrm{d}x\; \chi_p^*(x)\, h^{(1)}(x)\, \chi_q(x), \qquad h^{(1)}(x) \equiv -\tfrac{1}{2}\nabla^2 - \sum_A \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|}, \tag{3.2}$$

$$(pq|rs) = \iint \mathrm{d}x_1\, \mathrm{d}x_2\; \chi_p^*(x_1)\chi_q^*(x_2)\, \frac{1}{r_{12}}\, \chi_r(x_1)\chi_s(x_2), \tag{3.3}$$

$$\langle pq \| rs \rangle \equiv (pq|rs) - (pq|sr), \tag{3.4}$$

with $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$.

We reserve Roman indices p, q, r, s for spin–orbitals and Greek indices $\mu, \nu, \lambda, \sigma$ for spatial atomic orbitals (AOs). Here K denotes the total number of spin–orbitals in the one-electron basis; e.g., for an RHF reference built from K_{AO} spatial functions, $K = 2K_{\mathrm{AO}}$.

In this notation, particle number $\hat{N} = \sum_p a_p^\dagger a_p$ and spin projection $\hat{S}_z = \frac{1}{2}\sum_p \zeta_p a_p^\dagger a_p$ (with $\zeta_p = \pm 1$ for α/β) commute with $\hat{H}$, so the Hilbert space decomposes into (N, S_z) symmetry sectors. We exploit these symmetries classically (to restrict determinant spaces) and on qubits (e.g., via tapering).

In practice, one builds one- and two-electron integrals in an AO basis $\{\phi_\mu\}$ and transforms to a molecular orbital (MO) basis via C:

$$h_{pq}^{\mathrm{MO}} = \sum_{\mu\nu} C_{\mu p}^*\, h_{\mu\nu}^{\mathrm{AO}}\, C_{\nu q}, \tag{3.5}$$

$$\langle pq \| rs \rangle^{\mathrm{MO}} = \sum_{\mu\nu\lambda\sigma} C_{\mu p}^* C_{\nu q}^*\, \langle \mu\nu \| \lambda\sigma \rangle^{\mathrm{AO}}\, C_{\lambda r} C_{\sigma s}. \tag{3.6}$$

A naive four-index transformation is $\mathcal{O}(K^5)$ and storage of AO electron-repulsion integrals (ERIs) scales as $\mathcal{O}(K^4)$ which makes it impractical to store on current computers' memory. To alleviate this, we employ *density fitting / resolution-of-the-identity* (DF/RI),

$$(\mu\nu|\lambda\sigma) \;\approx\; \sum_{P=1}^{M} B_{\mu\nu}^P B_{\lambda\sigma}^P, \tag{3.7}$$

or *Cholesky decomposition* of ERIs, which reduce storage to $\mathcal{O}(K^2 M)$ and MO transformation to $\mathcal{O}(K^3 M)$ with $M \sim \mathcal{O}(K)$ auxiliary functions. For larger fragments, integral factorization (e.g., tensor hypercontraction) further reduces costs and often lowers the l_1 norm relevant for LCU and improves sparsity for Trotterization.

We distinguish three common references: RHF/RKS (paired spins, spatial orbitals doubly occupied), UHF/UKS (spin-polarized, separate α/β orbitals), and ROHF/ROKS (open-shell with paired core). For RHF, it is convenient to separate spatial-orbital blocks and build spin-adapted excitations; in UHF, the spin labels remain explicit. These choices affect the block structure of h_{pq}, the presence of near-degeneracies, and the active-space definition downstream.

Orbital localization (Boys, Pipek–Mezey, Edmiston–Ruedenberg) can make two-electron tensors and the induced qubit Hamiltonians more sparse and local, which decreases the number and weight of long Pauli strings after fermion–to–qubit mapping. Locality is particularly beneficial for measurement grouping, Trotter error bounds, and circuit layout on devices with limited connectivity.

The fermionic Hamiltonian in Eq. (3.1) has $\mathcal{O}(K^2)$ one-body and $\mathcal{O}(K^4)$ two-body terms, but only matrix elements consistent with (N, S_z) (and point-group irrep) contribute in a given sector. In correlated fragments we also *freeze core* orbitals to reduce K without sacrificing chemically relevant correlation. All such reductions also translate directly into fewer qubits and fewer Pauli terms post-mapping.

Relative to a mean-field reference $|\Phi_0\rangle$, it is often convenient to normal-order $\hat{H}$:

$$\hat{H} \;=\; E_{\mathrm{ref}} \;+\; \sum_{pq} f_{pq} \;:a_p^\dagger a_q:\; + \frac{1}{2}\sum_{pqrs}\langle pq\|rs\rangle \;:a_p^\dagger a_q^\dagger a_s a_r:, \tag{3.8}$$

where f_{pq} is the Fock matrix and $:\cdot:$ denotes normal ordering with respect to $|\Phi_0\rangle$. Normal ordering is standard in coupled-cluster theory, helps expose sparsity, and can slightly reduce the measured operator norm for near-term simulation.

Equation (3.1) is understood at fixed nuclear geometry $\mathbf{R}$; the one- and two-electron integrals inherit this $\mathbf{R}$–dependence and $E_{\mathrm{nuc}} = E_{\mathrm{nuc}}(\mathbf{R})$. In the next subsection we briefly justify this separation by recalling the Born–Oppenheimer approximation and its connection to potential energy surfaces (PES), then return to solving the electronic problem at fixed $\mathbf{R}$.

3.1.2 The Born-Oppenheimer Approximation

The basic idea of the approximation is to separate the motion of the nuclei and the electrons in an atom. We do so by considering that the atomic nucleus has a mass that is orders of magnitude higher than the mass of an electron. This is what we did in the previous section by fixing nuclear coordinates $\mathbf{R}$. On top of that, the nucleus and the electron share an attractive force due to having opposite charges, which in turn causes the electrons to be accelerated with much more intensity than the nucleus.

Let's start by looking at the Hamiltonian for a molecular system. Electrons will be numbered with the index i, and the nuclei will be numbered with the index j. The number of protons in a nucleus will be given by Z_j and its mass by m_j. The mass of all the electrons will be given by m_e. We will also be working in atomic units so that we can omit the constants.

The complete molecular Hamiltonian H will have the format

$$H = T^{\mathrm{N}} + T^{\mathrm{E}} + V^{\mathrm{NE}} + V^{\mathrm{EE}} + V^{\mathrm{NN}} \tag{3.9}$$

where T represents the kinetic energy and $V$the potential energy and the N and E indexes indicate whether the term is related to the nuclei, the electron, or both. Let's expand on each term.

The nuclei's kinetic energy is given by

$$T^{\mathrm{N}} = -\sum_j \frac{1}{2m_j}\nabla_j^2 = -\sum_j \frac{1}{2m_{\mathrm{j}}}\left(\frac{\partial^2}{\partial r_{xj}^2} + \frac{\partial^2}{\partial r_{yj}^2} + \frac{\partial^2}{\partial r_{zj}^2}\right) \tag{3.10}$$

The electrons' kinetic energy is given by

$$T^{\mathrm{E}} = -\sum_i \frac{1}{2m_{\mathrm{e}}}\nabla_i^2 = -\sum_i \frac{1}{2m_{\mathrm{e}}}\left(\frac{\partial^2}{\partial {r_{xi}}^2} + \frac{\partial^2}{\partial {r_{yi}}^2} + \frac{\partial^2}{\partial {r_{zi}}^2}\right) \tag{3.11}$$

The potential energy due to Coulomb attraction between the nuclei and the electrons, electron-electron repulsion, and nucleus-nucleus repulsion is

$$V^{\mathrm{NE}} + V^{\mathrm{EE}} + V^{\mathrm{NN}} = -\sum_i \sum_j \frac{Z_j}{r_{ij}} + \frac{1}{2}\sum_i \sum_{\underline{i}\neq i} \frac{1}{r_{i\underline{i}}} + \frac{1}{2}\sum_j \sum_{\underline{j}\neq j} \frac{Z_j Z_{\underline{j}}}{r_{j\underline{j}}} \tag{3.12}$$

The Born-Oppenheimer approximation then proceeds by first neglecting T^N for the reason we mentioned before, the mass of the nucleus is orders of magnitude higher than that of the electrons. We can also think of V^{NN} as a constant term because the nuclei do not change positions. This means that the Schrödinger equation will be reduced to its time-independent electronic form

$$H_e(r_i, r_j)\varphi_e(r_i, r_j) = E_e(r_j)\varphi_e(r_i, r_j) \tag{3.13}$$

for $H_e(r_i, r_j) = T^E(r_i) + V^{NE}(r_i, r_j) + V^{EE}(r_i)$. We can see that this equation will give the energy ground state E_e as a function of the nuclei positions r_j. This allows us to vary the nuclei's positions and solve the equation repeatedly which in turn will give us a curve known as the Potential Energy Surface(PES). From that curve, we can calculate the minimum which in turn will be the actual ground state energy of our molecular system. A PES for LiH can be seen on figure 9

Although not the focus of this document, it is relevant to mention one can take these calculations further to determine the motion of the nuclei. After obtaining the electronic structure one can solve the equation

$$\left[T^{\mathrm{N}} + E_{\mathrm{e}}(r_j)\right]\varphi(r_j) = E\varphi(r_j), \tag{3.14}$$

where E will now be the total energy of the molecule including electronic energy as well as contributions from nuclear vibration, rotation, and translation.

Solving these equations analytically is only possible for the simplest system of the Hydrogen atom, that's why we need to rely on numerical strategies such as the Hartree-Fock method to arrive at solutions for more complex systems. But before expanding on it let's look at some other relevant results.

3.1.3 The variational principle

To solve the Schrödinger equation the Hartree-Fock method relies on the Variational Principle. It states that given a Hamiltonian H with a discrete eigenvalue spectrum

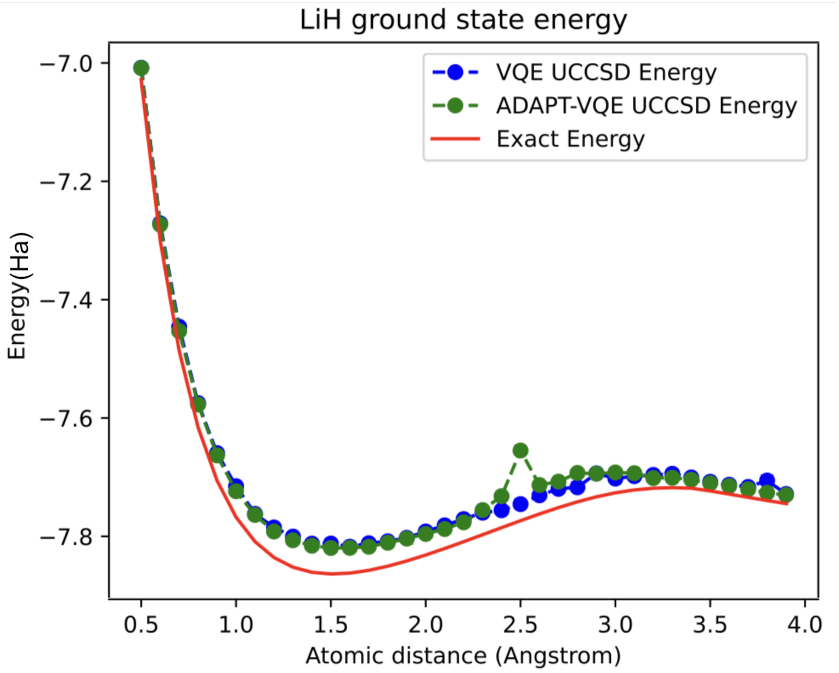


Figure 9 – **Potential Energy Surface for LiH calculated by VQE UCCSD (blue), ADAPT-VQE UCCSD (green).** The result of the direct diagonalization of the fermionic Hamiltonian using NumPy (red) is shown as a target. The minimum energy point should tell us the equilibrium distance for the Li-H bond. We'll give an explanation of the methods used to calculate later on.

and lowest eigenvalue e_0 we have that

$$\langle\phi|H|\phi\rangle \geq e_0 \tag{3.15}$$

for any normalized state $|\phi\rangle$. This is easily proven, assume ψ_i are eigenvectors of H and $|\phi\rangle = \sum_i c_i\psi_i$, then

$$\begin{aligned}\langle\phi|H|\phi\rangle &= \sum_{i,j} c_i c_j^* \langle\psi_j|H|\psi_i\rangle \\ &= \sum_i e_i c_i c_i^* \geq e_0.\end{aligned} \tag{3.16}$$

This means that in order to use the variational method to calculate the ground state we would need to find the state ϕ that minimizes $\langle\phi|H|\phi\rangle$. To do this we parametrize $|\phi(\overrightarrow{\alpha})\rangle$ according to a parameter vector $\overrightarrow{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_n)$ and minimize the function $\epsilon(\overrightarrow{\alpha}) = \langle\phi(\overrightarrow{\alpha})|H|\phi(\overrightarrow{\alpha})\rangle$.

$|\phi(\overrightarrow{\alpha})\rangle$ is called a trial wave function or ansatz. Usually, we want an ansatz that only covers a small subset of the Hilbert space that includes the solution for our minimization problem. Having an ansatz that is too general would increase the search space unnecessarily and choosing a subspace that doesn't cover the solution wouldn't allow us to find it. Choosing a proper ansatz is one of the most crucial elements.

The Hartree-Fock method will provide us with a way to minimize the function $\epsilon(\overrightarrow{\alpha})$, but before doing so in the next section we will understand how to represent the quantum state of our molecular system and build the function that we want to minimize.

3.1.4 Slater Determinants and the Fock operator

In order to approximate the total wavefunction Φ of a multi-electron atom or molecular system Douglas Hartree would first calculate the solution of Schrödinger's equation for each electron separately and assume that the product of these solutions would at least approximate the total solution for the complete system. Each of these one-electron wavefunctions is known as a spin-orbital, which we have been labeling with Roman indices p, q, r, s and counting with K.

This translates to solving the equation for Hamiltonians of the type

$$\hat{h}(i) = -\frac{1}{2}\nabla_i^2 - \sum_j \frac{Z_j}{r_{ij}}, \tag{3.17}$$

which would yield i solutions of the type $\phi_i(r_i)$ that would then form the complete solution $\Phi(\overrightarrow{r}) = \prod_i \phi_i(r_i)$ in what is known as the Hartree product.

As was mentioned before, Hartree failed to realize that the wave-function would have to be anti-symmetric as it's being used to describe fermions [509]. James Slater would then provide a correction in which he would redefine the wave function to adhere to the Pauli exclusion principle.

The new wave-function would take the form of the determinant of a matrix in which every element represents the state ϕ_k of the electron in position r_i. This can be written as

$$\begin{aligned}\Phi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) &= \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(\mathbf{r}_1) & \phi_2(\mathbf{r}_1) & \cdots & \phi_N(\mathbf{r}_1) \\ \phi_1(\mathbf{r}_2) & \phi_2(\mathbf{r}_2) & \cdots & \phi_N(\mathbf{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\mathbf{r}_N) & \phi_2(\mathbf{r}_N) & \cdots & \phi_N(\mathbf{r}_N) \end{vmatrix} \\ &\equiv |\phi_1, \phi_2, \cdots, \phi_N\rangle \\ &\equiv |1, 2, \ldots, N\rangle,\end{aligned} \tag{3.18}$$

where N is the total number of electrons and $\frac{1}{\sqrt{N!}}$ is a normalizing factor and the last two lines are shorthands that we are going to use to refer to Slater determinants.

The calculations that follow are taken from reference [68] and any omissions can be filled in from there. With that in mind, we can write the energy expectation value using equations 3.11 and 3.12.

$$\begin{aligned}E_E &= \langle \phi_1, \phi_2, \cdots, \phi_N | \, H_E \, | \phi_1, \phi_2, \cdots, \phi_N \rangle = \\ E_E &= \left\langle \phi_1, \phi_2, \cdots, \phi_N \left| -\sum_i \frac{1}{2}\nabla_i^2 - \sum_i \sum_j \frac{Z_j}{r_{ij}} + \sum_i \sum_{\underline{i} \neq i} \frac{1}{r_{i\underline{i}}} \right| \phi_1, \phi_2, \cdots, \phi_N \right\rangle,\end{aligned} \tag{3.19}$$

we then group the one-electron terms into $h(i)$

$$E_E = \left\langle \phi_1, \phi_2, \cdots, \phi_N \left| \sum_{i=1}^{N} h(i) + \sum_i \sum_{\underline{i} \neq i} \frac{1}{r_{i\underline{i}}} \right| \phi_1, \phi_2, \cdots, \phi_N \right\rangle, \tag{3.20}$$

and this equation can also be written as

$$E_E = \sum_k \int dr_1 \phi_k^* (r_1)\, h(1)\, (r_1)\, \phi_k (r_1) + \frac{1}{2} \sum_k \sum_{b=1}^{N} \left[\int dr_1 dr_2 \phi_k^* (r_1)\, \phi_k (r_1)\, r_{12}^{-1} \phi_b^* (r_2)\, \phi_b (r_2) - \int dr_1 dr_2 \phi_k^* (r_1)\, \phi_b (r_1)\, r_{12}^{-1} \phi_b^* (r_2)\, \phi_k (r_2) \right]. \tag{3.21}$$

Now we should apply the variational method to minimize equation 3.21. Unfortunately, this isn't straightforward as we are dealing with a constrained function. In fact, this function is constrained by the Slater determinants which have to be orthonormal. This means that we have to rely on Lagrange's method of undetermined multipliers to minimize it [68].

For brevity, we will omit the complete calculation here, but after taking into consideration the constraints the function we need to minimize is

$$\delta E \left[\phi_k^* (r_i) \right] = \delta \left\langle \phi_1, \phi_2, \cdots, \phi_N \left| H_E \right| \phi_1, \phi_2, \cdots, \phi_N \right\rangle - \delta \left[\sum_k \sum_{\underline{k}} \lambda_{k\underline{k}} \left(\left\langle \phi_k \psi_{\underline{k}} \right\rangle - \delta_{k\underline{k}} \right) \right] = 0, \tag{3.22}$$

which will take us to

$$h(1)\, (r_1)\, \phi_k (r_1) + + \sum_{\underline{k}} \left[\int dr_2 \phi_k (r_1)\, r_{12}^{-1} \phi_{\underline{k}}^* (r_2)\, \phi_{\underline{k}} (r_2) - \int dr_2 \phi_{\underline{k}} (r_1)\, r_{12}^{-1} \phi_{\underline{k}}^* (r_2)\, \phi_k (r_2) \right] - \sum_{\underline{k}} \varepsilon_{\underline{k}k} \phi_{\underline{k}} (r_1) = 0, \tag{3.23}$$

from which we define

$$J_{\underline{k}} (r_1)\, \phi_k (r_1) = \left[\int dr_2 \phi_{\underline{k}}^* (r_2)\, r_{12}^{-1} \phi_{\underline{k}} (r_2) \right] \phi_k (r_1) = \left[\int dr_2 \left| \phi_{\underline{k}} (r_2) \right|^2 r_{12}^{-1} \right] \phi_k (r_1) \tag{3.24}$$

and

$$K_{\underline{k}} (r_1)\, \phi_k (r_1) = \left[\int dr_2 \phi_{\underline{k}}^* (r_2)\, r_{12}^{-1} \phi_k (r_2) \right] \phi_{\underline{k}} (r_1). \tag{3.25}$$

This allows us to rewrite equation 3.23 as

$$F(r_1) \phi_k (r_1) = \sum_{\underline{k}} \varepsilon_{\underline{k}k} \phi_{\underline{k}} (r_1), \quad k = 1, 2, \ldots, N, \tag{3.26}$$

or, considering ε is hermitian we can write it on a basis in which it is also diagonal

$$F (\mathbf{r}_i)\, \phi(r_i) = \epsilon_i \phi(r_i) = h(1)\, (\mathbf{r}_i) + J (\mathbf{r}_i) - K (\mathbf{r}_i). \tag{3.27}$$

$F (\mathbf{r}_i)$ is known as the Fock operator, it contains three terms. The $h(1)\, (\mathbf{r}_i)$ stands for the one-electron Hamiltonian for electron i. The $J (\mathbf{r}_i)$ is the Coulomb operator that defines the electron-electron repulsion term. The $K (\mathbf{r}_i)$ is the exchange operator that comes from the anti-symmetric property of the Slater determinant.

Equations 3.26 and 3.27 are known as the Hartree-Fock equation, It might not be immediately obvious but it approximates the many body problem we previously had

into N one-body problems. This is equivalent to applying Mean-field Theory [510] to the original problem by replacing all electron-electron interactions with an effective interaction.

This omission is primarily due to the method's reliance on a single Slater determinant to approximate the wavefunction of a multi-electron system. Electron correlation plays a crucial role in accurately describing the behavior of electrons in molecules, affecting properties such as bond strength, ionization energies, and electron affinity.

The Hartree-Fock method, however, includes the concept of exchange energy, which stems from the exchange operator $K(\mathbf{r}_i)$ or in other words, from the antisymmetry requirement of the Slater determinant. Exchange energy accounts for the Pauli exclusion principle, ensuring that electrons with the same spin do not occupy the same quantum state. This introduces a level of electron-electron repulsion indirectly but does not capture the dynamic correlation resulting from their interaction.

To illustrate the significance of electron correlation and exchange energy, consider the dissociation of a Hydrogen molecule. The Hartree-Fock method can predict the energy and molecular orbitals accurately at equilibrium bond lengths. However, as the bond is stretched and the molecule dissociates into two hydrogen atoms, the lack of dynamic electron correlation leads to an overestimation of the bond strength which could give an incorrect prediction of the dissociation energy.

The reader should also realize that the Hartree-Fock method will start by computing the spin-orbitals, the one-electron wave functions that construct the Fock Operator and the Slater determinant. For atomic systems these take the shape of a hydrogen-like atom, an atom with one orbiting electron. For molecular systems, the one-electron wave function will look like linear combinations of atomic orbitals where the basis functions are single-electron wavefunctions for each of the component atoms of the molecule [68].

3.1.5 The Hartree-Fock Method and the Self-Consistent Field Algorithm

We can now see five simplifications that are crucial to the Hartree-Fock method:

1. The Born-Oppenheimer approximation must be used.
2. The variational solution, the wave-function, will be assumed to be a linear combination of orthogonal basis functions.
3. All eigenstates are functions assumed to be described by a single Slater determinant.
4. Mean Field Theory must be used and electron correlations must not be considered.
5. All relativistic effects are neglected. This is the first time we are mentioning relativistic effects, but it should be evident that we do not consider them as we are working on the Schrödinger equation and not the Dirac Equation.

With these in mind, after defining the Fock operator in equation 3.27, we proceed to expand the molecular orbitals in terms of a set of basis functions. This expansion necessitates the introduction of the overlap matrix S, which ensures that the basis functions remain orthonormal.

As mentioned previously, in the framework of the Hartree-Fock method, molecular orbitals are expressed as linear combinations of a finite set of basis functions $\{\psi_i(r_i)\}$. This approach transforms the continuous problem into a discrete one, solvable by numerical methods. To maintain the orthonormality of these basis functions, we introduce the overlap matrix S, defined by its elements:

$$S_{ij} = \int \psi_i(\mathbf{r})^* \psi_j(\mathbf{r})\, d\mathbf{r}, \tag{3.28}$$

where S_{ij} is the overlap integral between basis functions i and j, and $*$ denotes complex conjugation. The diagonal elements of S (S_{ii}) are always equal to 1 for normalized basis functions, indicating perfect self-overlap. In contrast, the off-diagonal elements (S_{ij}, $i \neq j$) measure the extent of overlap between different basis functions, varying from 0 (no overlap) to 1 (complete overlap).

Another relevant matrix is the F matrix, which is just a matrix representation of the Fock operator in equation 3.27

$$F_{ij} = \int \psi_i(\mathbf{r_1})^* F(\mathbf{r_1}) \psi_j(\mathbf{r})\, d\mathbf{r}, \tag{3.29}$$

The presence of the overlap matrix S and the Fock matrix F will be crucial to solve the Schrödinger equation. First, we rewrite the spatial function ϕ_i in the ψ_i basis,

$$\phi_i = \sum_{\mu=1}^{K} C_{\mu i} \psi_\mu, \quad i = 1, 2, \ldots, K \tag{3.30}$$

and substitute back in equation 3.13,

$$F(r_1) \sum_\nu C_{\nu i} \psi_\nu(r_1) = \varepsilon_i \sum_\nu C_{\nu i} \psi_\nu(r_1) \tag{3.31}$$

and integrate while multiplying by ψ^*, resulting in

$$\sum_\nu C_{\nu i} \int dr_1 \psi_\mu^*(r_1)\, f(r_1)\, \psi_\nu(r_1) = \varepsilon_i \sum_\nu C_{\nu i} \int dr_1 \psi_\mu^*(r_1)\, \psi_\nu(r_1) \tag{3.32}$$

which we can rewrite as

$$\sum_\nu F_{\mu\nu} C_{\nu i} = \varepsilon_i \sum_\nu S_{\mu\nu} C_{\nu i}, \quad i = 1, 2, \ldots, K \tag{3.33}$$

or, in the more usual form,

$$FC = SC\epsilon, \tag{3.34}$$

where C is the matrix of coefficients that represent the expansion of molecular orbitals in terms of the basis set, and ϵ is a diagonal matrix containing the orbital energies.

These are known as the Roothan-Hall equations and are a direct consequence of applying the variational principle to the Hartree-Fock approximation of the Schrödinger equation, under the constraints imposed by a chosen basis set. They enable the determination of the best possible set of molecular orbitals (within the limits of the Hartree-Fock approximation and the chosen basis set) that minimize the energy of the system. The

Fock matrix F incorporates the kinetic energy of the electrons, their potential energy due to interactions with the nuclei, and the mean-field representation of electron-electron repulsions, including exchange interactions.

Let's also define the density matrix as well to see how it can be written in terms of C, we begin by writing the electronic density of the system

$$\begin{aligned}
\rho(r) &= 2\sum_{a}^{N/2} \phi_a^*(r)\phi_a(r) \\
\rho(r) &= 2\sum_{a}^{N/2} \sum_{\nu} C_{\nu a}^* \psi_\nu^*(r) \sum_{\mu} C_{\mu a}\psi_\mu(r) \\
\rho(r) &= \sum_{\mu\nu} \left[2\sum_{a}^{N/2} C_{\mu a} C_{\nu a}^* \right] \psi_\mu(r)\psi_\nu^*(r)
\end{aligned} \tag{3.35}$$

where

$$D_{\mu\nu} = 2\sum_{a}^{N/2} C_{\mu a} C_{\nu a}^* \tag{3.36}$$

is called the Density Matrix. The density matrix offers a concise way to represent the electron density distribution across a molecular system. This expression simplifies the calculation of various properties, including the total electronic energy, electron density, and other observables, by reducing them to integrals over the density matrix.

Let's now describe a practical implementation of the Hartree-Fock Method.

The Self-Consistent Field (SCF) algorithm is used to find an approximate solution to the electronic Schrödinger equation for multi-electron systems. The goal of the SCF process is to solve the Fock matrix equations iteratively to find a set of molecular orbitals that minimize the system's energy, under the constraint that the orbitals remain orthonormal. The key idea is that the Fock operator, which describes the effective field felt by an electron, depends on the very orbitals that are used to calculate.

Initially, a guess for the molecular orbitals is needed. These can be obtained by using atomic orbitals. The algorithm then proceeds in a loop, with each iteration consisting of the following steps:

1. Build the Fock matrix based on the current guess of the orbitals. This involves calculating the contributions from the kinetic energy of the electrons, the nuclear attraction, and the electron-electron repulsion, including exchange terms.

2. Solve the Roothan-Hall equations, to obtain updated molecular orbitals. This step involves diagonalizing the Fock matrix to obtain its eigenvalues and eigenvectors. The eigenvectors represent the new orbitals, while the eigenvalues are associated with their energies.

3. Check for convergence by comparing the energy of the current iteration with that of the previous iteration. If the difference is within a pre-defined threshold, and the difference in the density matrix between iterations is also within a certain tolerance, the algorithm has converged, and the current set of orbitals represents the solution.

4. If the algorithm has not converged, form a new guess for the orbitals and repeat the process. This can involve simply taking the orbitals obtained from diagonalizing the Fock matrix, or it might include a mix of these orbitals with those from previous iterations to improve convergence.

5. Update the total energy of the system, including contributions from electron-electron repulsion, and return to step 1 if necessary.

Algorithm 1 Self-Consistent Field (SCF) Algorithm

Initialize:
- Guess initial molecular orbitals, ψ_{initial}
- Set convergence criteria $\varepsilon_{\text{energy}}$ and $\varepsilon_{\text{density}}$
- Set maximum number of iterations, max_iter

repeat
1. Build the Fock matrix, F, based on current orbitals
2. Solve the Roothan-Hall equations: $FC = SC\epsilon$
   where C is the matrix of coefficients (orbitals),
   S is the overlap matrix, and ϵ are the orbital energies
3. Form the new density matrix, D_{new}, from C
4. Calculate total energy, E_{total}, and check for convergence:
   if $|E_{\text{total}} - E_{\text{previous}}| < \varepsilon_{\text{energy}}$ and $|D_{\text{new}} - D_{\text{previous}}| < \varepsilon_{\text{density}}$:
   Convergence achieved, exit loop
5. Update D_{previous} to D_{new} and E_{previous} to E_{total}
6. Optionally, mix D_{new} with D_{previous} to improve convergence

until Convergence or max_iter reached

Output:
- Converged molecular orbitals
- Total energy of the system

The SCF algorithm iteratively solves the Fock matrix equations to converge on a stable set of molecular orbitals that minimize the system's energy. However, this iterative process poses several computational challenges, including convergence issues. For instance, the choice of initial guess for the molecular orbitals can significantly affect the convergence rate and even the possibility of reaching the true ground state energy within a reasonable number of iterations. The procedure is described in pseudo-code in Algorithm 1.

In practice, the SCF loop is a fixed-point iteration on the density, $D_k \to F[D_k] \to C_k \to D_{k+1}$. A common strategy to accelerate convergence, is to apply Pulay's DIIS (direct inversion in the iterative subspace) between building F (Step 1) and solving the generalized eigenproblem (Step 2): we form the commutator-like error $e_k = F_k D_k S - S D_k F_k$, keep a short history of (F_i, e_i), and solve a small least-squares problem to construct an extrapolated Fock $\tilde{F}_k = \sum_i b_i F_i$ that minimizes the residual. Step 2 then uses $\tilde{F}_k$ in place of F_k. We optionally stabilize the update with density damping $D_{k+1} \leftarrow (1-\alpha)D_k + \alpha D_{k+1}$ (Step 6) and a temporary virtual-orbital level shift applied to F_k before Step 2 to prevent oscillations near degeneracies. Convergence is declared only when both the energy change and a density (or DIIS-error) norm fall below tight thresholds.

Furthermore, the Hartree-Fock method's limitations become evident when dealing with systems where electron correlation is essential for accurate predictions. While HF provides a good approximation for many properties of atoms and molecules, its neglect of correlation energy can lead to substantial errors in predicting reaction energies, excited states, and properties of ionic systems.

3.1.6 Density Functional Theory

In contrast to the explicit wavefunction-based strategy of HF, Density Functional Theory (DFT) introduces a paradigm shift by focusing on the electron density $\rho(\mathbf{r})$ as the fundamental quantity from which all ground-state properties of the system can be derived. Rooted in the Hohenberg-Kohn theorems [22], DFT asserts that the electron density alone suffices to determine the system's total energy, bypassing the need for the many-electron wavefunction. This shift not only simplifies the computational complexity but also introduces a new way of thinking about electronic interactions, especially through what became known as the exchange-correlation functional, $E_{xc}[\rho]$.

The Hohenberg-Kohn theorems are as follows:

1. The First Hohenberg-Kohn Theorem establishes the variational principle for the electron density, asserting that the ground state energy of a many-electron system is a unique functional of the electron density:

$$E[\rho] = T[\rho] + V_{\text{ext}}[\rho] + J[\rho] + E_{\text{xc}}[\rho], \tag{3.37}$$

   where $E[\rho]$ is the total energy, $T[\rho]$ is the kinetic energy functional, $V_{ext}[\rho]$ is the external potential energy (usually due to nuclei), $J[\rho]$ represents the classical Coulomb (Hartree) interaction, and $E_{xc}[\rho]$ is the exchange-correlation energy functional.

2. The Second Hohenberg-Kohn Theorem provides the basis for a variational approach, stating that the ground state energy of the system can be obtained by minimizing the energy functional $E[\rho]$ with respect to the electron density, leading to the exact ground state density and energy under the given external potential.

We won't go over the proof of these theorems, but let's at least define these functionals. We begin by writing the electronic density again,

$$\rho(\mathbf{r}) = N \int \mathrm{d}^3\mathbf{r}_2 \cdots \int \mathrm{d}^3\mathbf{r}_N \psi^* \left(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N\right) \psi \left(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N\right), \tag{3.38}$$

where N is the number of occupied orbitals and ψ is the normalized wavefunction. As stated by Hohenberg-John [22] this relation is reversible, meaning, for a given ground state density $\rho_0(\mathbf{r})$ it is possible to calculate the corresponding ground-state wavefunction $\psi_0 = \psi[\rho_0(\mathbf{r})]$. This way the ground state expectation value of the Hamiltonian would be given by

$$\begin{aligned} E_0 = H\left[n_0\right] &= \left\langle \psi\left[\rho_0\right] | T^N + T^E + V^{NN} + V^{NE} + V^{EE} | \psi\left[\rho_0\right] \right\rangle, \\ &= \left\langle \psi\left[\rho_0\right] | T + V + J | \psi\left[\rho_0\right] \right\rangle \end{aligned} \tag{3.39}$$

where we've summed the kinetic terms into T, the Coulomb potential terms into J and the external potential terms into V.

The external potential functional can be explicitly written, so we write the total energy functional as a sum of other energy functionals,

$$H[n] = T[\rho] + J[\rho] + \int V(\mathbf{r})\rho(\mathbf{r})\mathrm{d}^3\mathbf{r}. \tag{3.40}$$

This functional when minimized will give the ground state density ρ_0.

In order to solve this Walter Kohn and Lu Jeu Sham proposed solving the Schrödinger-like equations for the non-interacting electron system [23]:

$$\left[-\frac{1}{2}\nabla^2 + V_{\text{eff}}(\mathbf{r})\right]\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}), \tag{3.41}$$

where the effective potential $V_{\text{eff}}(\mathbf{r})$ is defined as:

$$V_{\text{eff}}(\mathbf{r}) = V_{\text{ext}}(\mathbf{r}) + V_{\text{H}}(\mathbf{r}) + V_{\text{xc}}(\mathbf{r}). \tag{3.42}$$

Here, $V_{\text{ext}}(\mathbf{r})$ represents the external potential (usually from the nuclei), $V_{\text{H}}(\mathbf{r})$ is the Hartree potential accounting for classical electron-electron repulsion, and $V_{\text{xc}}(\mathbf{r})$ is the exchange-correlation potential, which is the derivative of the exchange-correlation energy functional $E_{\text{xc}}[\rho]$ with respect to $\rho(\mathbf{r})$. The ground-state electron density $\rho(\mathbf{r})$ can be reconstructed from the Kohn-Sham orbitals ψ_i as follows:

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} |\psi_i(\mathbf{r})|^2, \tag{3.43}$$

where the sum runs over all occupied orbitals.

Up until this point, we've been purposefully avoiding extensive explanations of the exchange-correlation potential and functional but they are the key to the simplified DFT calculations. The hardest part of simulating many-body electronic systems is given by the exchange interactions and the correlation effects between electrons. The HF methods deal with these by having the anti-symmetry operator K we discussed in the previous sections and by considering only the average field effects. DFT simplifies this challenge by introducing the exchange-correlation energy functional, which depends only on the electron density, a much more tractable quantity than the wavefunction.

A major problem with this method though is that exact exchange-correlation functionals represent the primary unknown in DFT calculations. Various approximations have been developed to estimate them though, starting from the Local Density Approximation (LDA) and advancing to more sophisticated formulations like the Generalized Gradient Approximation (GGA) and beyond [515] [516]. A pseudocode approach based on the previous Self Consistent Field Algorithm but now using DFT is described in Algorithm 2

3.1.7 Basis sets, frozen cores, and effective core potentials

All electronic-structure calculations in this thesis employ contracted Gaussian-type orbitals (GTOs) basis sets, in which each molecular (or atomic) basis function is a fixed

Algorithm 2 Density Functional Theory (DFT) Algorithm

Initialize:
- Choose an initial guess for the electron density, $\rho_{\text{initial}}(\mathbf{r})$
- Set convergence criteria $\varepsilon_{\text{density}}$ and $\varepsilon_{\text{energy}}$
- Set maximum number of iterations, *max_iter*

repeat

1. Construct the effective potential, $V_{\text{eff}}(\mathbf{r})$, using the current electron density:
   $V_{\text{eff}}(\mathbf{r}) = V_{\text{ext}}(\mathbf{r}) + V_{\text{Hartree}}[\rho](\mathbf{r}) + V_{\text{xc}}[\rho](\mathbf{r})$
   where V_{ext} is the external potential,
   V_{Hartree} is the Hartree potential,
   and V_{xc} is the exchange-correlation potential.
2. Solve the Kohn-Sham equations for the current effective potential:
   $[-\frac{1}{2}\Delta + V_{\text{eff}}(\mathbf{r})]\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r})$
3. Update the electron density using the obtained orbitals:
   $\rho_{\text{new}}(\mathbf{r}) = \sum_{i=1}^{\text{occupied}} |\psi_i(\mathbf{r})|^2$
4. Calculate the total energy $E_{\text{total}}[\rho]$ for the current density.
5. Check for convergence:
   if $|E_{\text{total}} - E_{\text{previous}}| < \varepsilon_{\text{energy}}$ and $|\rho_{\text{new}} - \rho_{\text{previous}}| < \varepsilon_{\text{density}}$:
   Convergence achieved, exit loop.
6. If not converged, update ρ_{previous} to ρ_{new}, E_{previous} to E_{total},
   and go back to step 1 with the updated density.

until Convergence or *max_iter* reached

Output:
- Converged electron density $\rho(\mathbf{r})$
- Total energy of the system E_{total}
- Kohn-Sham orbitals $\{\psi_i\}$ and their energies $\{\epsilon_i\}$

linear combination of primitive Gaussians. This contraction affords large efficiency gains while retaining systematic improvability. Historically important families include the Pople split-valence sets (e.g., 3-21G, 6-31G and polarized variants) designed for general molecular work [517–519], the correlation-consistent Dunning sets `cc-pVXZ` (X = D,T,Q,5,...) devised for controlled convergence of post-HF correlation energies [520], their augmented variants for diffuse states (`aug-cc-pVXZ`) [521], and core–valence extensions (`cc-pCVXZ`) that enable explicit correlation of inner-shell electrons when needed [522]. We also make use of the Karlsruhe `def2` family (e.g., `def2-SV(P)`, `def2-TZVP`, `def2-QZVPP`), which provides balanced segmented-contraction sets across the periodic table and pairs naturally with modern auxiliary and effective-core-potential options [523].

To reduce the cost of four-center two-electron integrals, we can employ *density fitting/resolution-of-the-identity* (RI) approximations, which replace exact electron-repulsion integrals by expansions in an auxiliary Gaussian basis [524, 525]. This idea was shown to yield substantial speedups with minimal loss of accuracy in correlated methods such as MP2 (RI-MP2) [526], and, while MP2 is not used in our work, this has since become standard in HF/DFT and post-HF workflows as well; carefully optimized auxiliary sets further improve stability and accuracy [527]. As a complementary route, *Cholesky decomposition* of the integral tensor achieves similar savings by factorizing the Coulomb metric on the fly [528]. Both strategies also reduce memory footprints and the cost of AO $\rightarrow$ MO integral

transformations, which is valuable for large active spaces and embedded fragments.

For many applications, we can invoke the *frozen-core* approximation: chemically inert inner-shell orbitals are held doubly occupied and excluded from the excitation (correlation) space, thereby lowering cost with little impact on valence properties [68]. This is the default in high-accuracy coupled-cluster benchmarks (e.g., CCSD(T)) [529]. When core effects are non-negligible (e.g., precise thermochemistry, core-level spectroscopy, or very light elements), one can correlate them explicitly in conjunction with core–valence basis sets such as `cc-pCVXZ` [522].

For heavier elements, it's possible to replace explicit core electrons with effective core potentials (ECPs; pseudopotentials) that fold scalar-relativistic and core–valence effects into an effective one-electron operator, reducing the number of explicit electrons and basis functions. Widely used families include the LANL (Hay–Wadt) ECPs for transition metals and main-group elements [530, 531] and the Stuttgart/Cologne energy-adjusted ECPs for second- and third-row transition metals [532]. The Karlsruhe `def2` basis sets are designed to be used together with matching small-core ECPs on heavier atoms [523].

Finally, we can mitigate basis-set incompleteness by increasing the basis cardinal number and, where feasible, extrapolating to the complete-basis-set (CBS) limit. Correlation-consistent sequences permit simple two-point formulas, e.g., $E_X \approx E_\infty + AX^{-3}$, for the correlation energy, and analogous forms for the HF component [533, 534]. Classic studies demonstrate that such extrapolations recover most of the remaining basis error at a fraction of the cost of the next larger basis [533,535]. Diffuse augmentation is recommended for anions, Rydberg states, and weakly bound systems [521].

Choice of basis/ECP directly sets the size and structure of the second-quantized Hamiltonian (Sec. 3.1.1): fewer, more local orbitals ⇒ fewer qubits, sparser Pauli operators after fermion–to–qubit mapping, and lower measurement as well as Trotter/LCU costs (via reduced l_1 norms and commutation structure). A pragmatic workflow is: (i) use def2-TZVP or cc-pVTZ as a balanced default [520, 523]; (ii) add "aug"/"+" diffuse functions for anions and weak interactions [521, 536]; (iii) freeze chemically inert cores (or employ small-core ECPs) unless core-level properties are targeted [68]; (iv) apply counterpoise for noncovalent benchmarks [537]; and (v) where feasible, perform TZ/QZ two-point CBS extrapolations for post-HF comparisons [533, 534]. We adopt these settings in the coupled-cluster benchmarks in chapters 5 and 6 and in defining active spaces for quantum variational studies.

DFT replaces the many–determinant wavefunction by an interacting density functional, trading systematic improvability for favorable scaling and broad applicability. In the next sections we return to wavefunction treatments that recover correlation explicitly in the determinant space defined by the same AO/MO basis and core/ECP choices above. These choices now also control the dimension of the Hilbert space (number of determinants) and the sparsity pattern of the Hamiltonian matrix elements, thereby setting both classical cost and, eventually, the size of mapped qubit Hamiltonians.

To establish a variational baseline and a clear point of comparison for coupled-cluster theory, we begin with Configuration Interaction (CI). Full CI (FCI) provides the exact solution within the chosen basis, while truncated CI (CIS, CID, CISD, ...) offers controlled but not size-extensive approximations. This contrast will motivate the exponential ansatz of CCSD and CCSD(T) that follows.

3.1.8 Full Configuration Interaction and Truncated CI Methods

In CI theory, the many-electron wavefunction is expanded as a linear combination of Slater determinants that span the N-electron Hilbert space of a given one-particle basis. In the limit of FCI, one includes *all* possible Slater determinants constructed from a chosen finite orbital basis. This yields an exact diagonalization of the electronic Hamiltonian in that finite space, providing the exact non-relativistic solution (within the basis) via the variational principle [68]. Mathematically, an FCI wavefunction can be written as:

$$|\Psi_{\mathrm{FCI}}\rangle; =; \sum_I c_I, |\Phi_I\rangle \ , \tag{3.44}$$

where $|\Phi_I\rangle$ is the complete set of N-electron Slater determinants in the chosen spin-orbital basis, and c_I are linear expansion coefficients. Each $|\Phi_I\rangle$ is typically a determinant of N spin-orbitals (e.g. Hartree–Fock orbitals) occupied by the N electrons. The Hartree–Fock determinant is often taken as a reference $|\Phi_0\rangle$, and "excited" determinants are those obtained by replacing occupied orbitals in $|\Phi_0\rangle$ with virtual (unoccupied) orbitals. For example, $|\Phi_i^a\rangle$ denotes the Slater determinant obtained by exciting one electron from occupied orbital i into a previously empty orbital a, and $|\Phi_{ij}^{ab}\rangle$ corresponds to simultaneously exciting two electrons from i, j to a, b, etc. In second-quantized notation, one can write $|\Phi_i^a\rangle = a_a^\dagger, a_i, |\Phi_0\rangle$ and $|\Phi_{ij}^{ab}\rangle = a_a^\dagger a_b^\dagger, a_j, a_i, |\Phi_0\rangle$, where $a_p^\dagger, a_q$ are creation and annihilation operators for orbitals p, q. The FCI expansion (3.44) thus contains the reference configuration, all singles, doubles, triples, up to N-tuple excitations. For clarity, one may express a truncated CI wavefunction with explicit excitation classes:

$$|\Psi_{CI}\rangle; =; c_0, |\Phi_0\rangle; +; \sum_{a,i} c_i^a, |\Phi_i^a\rangle; +; \tfrac{1}{2!} \sum_{a<b,,i<j} c_{ij}^{ab}, |\Phi_{ij}^{ab}\rangle; +; \tfrac{1}{3!} \sum_{a<b<c,,i<j<k} c_{ijk}^{abc}, |\Phi_{ijk}^{abc}\rangle; +; \cdots \ , \tag{3.45}$$

where $i, j, k, \ldots$ index occupied orbitals in the reference and $a, b, c, \ldots$ index virtual orbitals. The first term is the HF determinant, the second term summation runs over all singly-excited determinants $|\Phi_i^a\rangle$, the third term runs over all distinct double excitations $|\Phi_{ij}^{ab}\rangle$, and so on. The factorial denominators account for the indistinguishability of excitation order (e.g. $|\Phi_{ij}^{ab}\rangle$ and $|\Phi_{ji}^{ba}\rangle$ are the same determinant). In an FCI, all excitation levels up to N-tuple are included; in a *truncated* CI, the expansion is cut off at a certain excitation rank (see below). The c_I coefficients are determined by solving the Schrödinger equation projected onto the determinant basis. Substituting expansion (3.44) into $\hat{H}|\Psi\rangle = E, |\Psi\rangle$ and projecting onto a determinant $\langle\Phi_J|$ yields the *secular equations*:

$$\sum_I H_{JI}, c_I = E, c_J, \qquad \text{for each basis determinant } J, \tag{3.46}$$

where $H_{JI} = \langle\Phi_J|\hat{H}|\Phi_I\rangle$ are the Hamiltonian matrix elements. The Hamiltonian $\hat{H}$ consists of one- and two-electron operators, $\hat{H} = \sum_{p,q} h_{pq}, a_p^\dagger a_q + \frac{1}{2}\sum_{p,q,r,s} g_{pqrs}, a_p^\dagger a_q^\dagger a_s a_r$ (with h_{pq} one-electron integrals and g_{pqrs} two-electron integrals in the orbital basis). Because $\hat{H}$ is at most two-body, H_{JI} is nonzero only if $|\Phi_J\rangle$ and $|\Phi_I\rangle$ differ by at most two orbital substitutions. Each nonzero matrix element can be evaluated from the integrals using Slater–Condon rules [19,68,507,538]. The full CI matrix is extremely large but sparse

in this determinant basis. Diagonalizing this matrix (e.g. by iterative eigensolvers like Davidson's [539]) yields the FCI eigenvalues and eigenvectors. The lowest eigenvalue from FCI is the ground-state energy E_0, and the corresponding eigenvector gives the exact (within the basis) ground-state wavefunction coefficients c_I. Higher eigenvalues/eigenvectors from the same diagonalization correspond to excited electronic states within that basis.

In practice, FCI is only tractable for very small systems, because the number of determinants grows combinatorially with system size and basis size. The dimension of the FCI Hilbert space (ignoring spatial symmetries) is given by the binomial coefficient $N_{\mathrm{det}} = \binom{M}{N}$ for N electrons in M spin-orbitals (or more generally $\binom{M_\alpha}{N_\alpha}\binom{M_\beta}{N_\beta}$ if α, β spins are considered separately) [540]. This number explodes rapidly: for example, a modest system of $N = 10$ electrons in $M = 28$ spin-orbitals (14 spatial orbitals) yields on the order of $\binom{28}{10} \approx 1.3 \times 10^7$ determinants. The storage and processing of even tens of millions of configurations is extremely demanding. Indeed, few FCI calculations with chemically interesting basis sets have been achieved; they are mostly limited to very small molecules (e.g. FCI for H_2O in a double-zeta basis, or for N_2 with minimal basis) as benchmark studies [541]. Because of this steep computational cost, one usually resorts to a *truncated CI* approach, wherein only determinants up to a certain excitation level from a reference are included in the expansion. This defines the hierarchy of CI methods:

- CIS: Configuration Interaction with Singles. Only one-electron excitations (plus the reference $|\Phi_0\rangle$) are included. CIS is primarily used for approximate excited states, since a single excitation from a closed-shell reference produces one-hole–one-particle configurations that correspond to excited electronic configurations. For a closed-shell HF reference, CIS does not improve the ground-state energy because of Brillouin's theorem (the HF determinant has no single excitation mixing if orbitals are canonical, as $\langle\Phi_0|\hat{H}|\Phi_i^a\rangle = 0$) [68]. Thus, CIS ground-state correlation energy is zero, but CIS can describe singly-excited states qualitatively.

- CID: Configuration Interaction with Doubles. Only double excitations (and the reference) are included. For a closed-shell system, double substitutions are the first that directly reduce the ground-state energy by introducing electron correlation. CID captures some correlation energy, especially the so-called dynamical correlation, but neglects any effect of single excitations (orbital relaxation) and higher (triple, quadruple, ...) excitations. CID is rarely used alone (because if singles do contribute or the reference is not optimized, omitting singles is suboptimal), but it is size-consistent for closed-shell pairs in a limited sense (see below).

- CISD: Configuration Interaction with Singles and Doubles. This includes all one- and two-electron excitations on top of $|\Phi_0\rangle$. CISD was historically a popular post-Hartree–Fock method, as it provides a significant improvement over HF by recovering a large fraction of the correlation energy. For many small molecules at their equilibrium geometries, CISD recovers roughly $\sim$ 90–96% of the total correlation energy in the given basis. The remaining correlation comes from higher excitations (triples, etc.). While CISD is a variational improvement over HF, its accuracy deteriorates for larger systems and in cases of strong correlation.

- CISDT, CISDTQ, ...: Higher-order CI including triple, quadruple, etc., excitations in addition to all lower excitations. As one includes more excitation levels, the CI

expansion systematically approaches the FCI limit. For instance, CISDT (singles, doubles, triples) can recover around 95–99% of correlation for small molecules, and CISDTQ (up to quadruples) often yields $> 99.5\%$, nearly exact. However, the number of triple and quadruple excitations is extremely large (the count grows combinatorially with excitation rank), so these methods are feasible only for very small systems (CISDTQ might be attempted for molecules with 2–4 light atoms at most). In general, including triples and beyond increases the computational cost dramatically.

Each increase in excitation level improves the variational flexibility of the wavefunction and lowers the energy (since the exact wavefunction is approached from above). However, the convergence of truncated CI can be slow, higher excitations may contribute non-negligibly, especially in multi-reference situations (e.g. bond breaking or stretched geometries). Notably, the fraction of correlation energy recovered by CISD tends to *decrease* as the molecular size increases if one uses the same excitation truncation. This is because larger systems have more room for electron correlation via higher-order excitations that CISD cannot capture. For example, CISD might recover 95% of correlation for a single small molecule, but for two non-interacting copies of that molecule, CISD would recover appreciably less than 95% of the combined correlation energy (since some correlation in each fragment would require simultaneous double excitations in each, i.e. a quadruple overall, which CISD lacks). In other words, truncated CI yields a smaller fraction of correlation for larger systems, reflecting the fundamental issue of *size-extensivity*, discussed below.

A key attractive feature of CI is that it is a *linear variational method*. The CI expansion coefficients c_I are obtained by diagonalizing the true Hamiltonian in an approximate subspace (spanned by the chosen determinants). Therefore, the CI ground-state energy E_{CI} is always an upper-bound to the exact ground-state energy (for the given Hamiltonian). In a finite incomplete one-electron basis, the FCI energy is the exact non-relativistic energy within that basis (often called the "full-CI limit" for that basis), and any truncated CI will give an energy $\geq E_{\mathrm{FCI}}$ for that basis. By the variational theorem, improving (enlarging) the CI expansion can only lower or maintain the energy, and in the limit of including all determinants, one obtains the exact energy in the basis [68]. This variational property gives CI a systematic improvability and a direct measure of error: e.g. the difference $E_{CISD} - E_{\mathrm{FCI}}$ in a given basis indicates the residual correlation missing due to triples and higher. Also, as $E_{\mathrm{CI}} \to E_0$ (exact), the approximate CI wavefunction $|\Psi_{\mathrm{CI}}\rangle$ converges towards the exact wavefunction; expectation values of other observables tend to improve as well (albeit not necessarily as rapidly as the energy).

In contrast, many other correlated methods are *non-variational*. For example, Coupled-Cluster (CC) theory with truncation does not produce energies that are guaranteed to lie above the exact energy [542]. It often yields extremely accurate results even when truncated (CCSD, CCSD(T), etc.), but its energy is not a variational upper bound – indeed, truncated CC can sometimes give an energy lower than the true FCI energy in the same basis [542]. The CI method's strict adherence to the variational principle is a double-edged sword: it guarantees no "over-correlation" (one will never get an unphysical energy below the true ground state), but it also means truncated CI must include many terms to approach the exact result, whereas CC can achieve similar accuracy with fewer parameters by using a nonlinear ansatz.

A well-known deficiency of truncated CI (except FCI) is the lack of *size-extensivity* and *size-consistency* [540, 543]. A method is said to be *size-extensive* if the energy scales linearly with the number of non-interacting identical subsystems (extensivity implies that $E(2A) = 2E(A)$ for system A duplicated, etc.). A related concept, *size-consistency*, specifically means that for two non-interacting fragments A and B, the total energy $E(A + B)$ equals the sum $E(A) + E(B)$, even if A and B have different electronic structures. Full CI is size-extensive and size-consistent by construction (since an exact treatment simply sums the energies of separate subsystems). However, a truncated CI expansion is not guaranteed to fulfill these properties. In fact, CIS, CISD, etc., are *neither* size-extensive nor size-consistent [542].

The error can be demonstrated by a simple example: consider two identical molecules A and B far apart with no interaction. A CISD calculation on the combined $A + B$ system will miss correlation energy that involves simultaneous excitations on both A and B. In particular, a double excitation on fragment A and a double excitation on fragment B together constitute a quadruple excitation in the supermolecule $A + B$. CISD, which only includes up to double excitations overall, cannot capture both fragments' correlation energy at once. As a result, $E_{CISD}(A + B) > E_{CISD}(A) + E_{CISD}(B)$. The deficiency grows with system size: the more electrons or fragments present, the greater the proportion of correlation that involves higher-than-double excitations. Consequently, truncated CI energies do not scale properly with system size—the correlation energy per electron effectively decreases as the system grows, indicating a failure of size-extensivity.

This is a serious drawback if one's aim is to apply the method to large systems, as error rapidly accumulates. By contrast, coupled cluster methods with an exponential ansatz are size-extensive by design, which has contributed to their popularity for extensive systems. Various attempts have been made to correct the size-consistency error of truncated CI a posteriori. A famous example is the Davidson correction [543], an empirical perturbative estimate of the contribution of higher excitations (particularly the unlinked quadruple excitation contribution missing in CISD). In the Davidson CISD+Q approach, one computes an approximate correction $\Delta E_Q \approx (1 - c_0^2), (E_{CISD} - E_{\mathrm{HF}})$ (where c_0^2 is the weight of the reference in the CISD wavefunction) to estimate the effect of connected quadruples [543]. Davidson-corrected CISD energies are often much closer to FCI and approximately size-consistent for small systems, and this approach was used in the past to mitigate errors. However, the correction is not exact and fails for larger systems where higher-than-quadruple excitations matter; moreover, it provides no improved wavefunction.

Another strategy introduced by Pople and coworkers is the *Quadratic Configuration Interaction* method (QCI), such as QCISD, which incorporates nonlinear terms in the CI expansion to restore size-extensivity approximately [544]. In practice, QCISD+T (and later, CCSD(T)) became widely used, but it was shown that coupled cluster methods are more robustly size-extensive than QCI [545]. In summary, while truncated CI is variationally well-behaved, its lack of size-extensivity means it is not a consistent method as the system grows or when considering separated subsystems. This significantly limits the applicability of straightforward truncated CI in computational chemistry, especially compared to size-extensive methods like CC.

The computational cost of CI grows rapidly with the excitation level and number of basis functions. Constructing and diagonalizing the Hamiltonian matrix in the FCI space formally scales exponentially with system size. Even storing the FCI wavefunction requires

$\mathcal{O}(N_{\text{det}})$ memory, which, as noted, grows combinatorially with N and M. For truncated CI, the number of determinants is more manageable but still large: for example, CISD includes the reference + all singles ($N_{occ} \times N_{virt}$ configurations) + all doubles ($\sim \frac{N_{occ}^2 N_{virt}^2}{2}$ configurations). The count scales roughly as $\mathcal{O}(M^2N^2)$ for CISD (where $N_{occ} \approx N$ and $N_{virt} \approx M - N$), which is already quite steep. The operation count to evaluate the CI eigenproblem is also high. A naive direct diagonalization of the Hamiltonian (of dimension N_{det}) would cost on the order of $\mathcal{O}(N_{\text{det}}^3)$ operations, which is prohibitive for N_{det} beyond a few thousands. Fortunately, one rarely constructs the full matrix explicitly for larger CI; instead, one typically uses iterative (such as Davidson's [539]) algorithms to converge a few lowest eigenvalues, requiring only matrix–vector products at each iteration. Each matrix–vector multiplication can be performed by exploiting the structure of H: one can loop over the determinants and sum contributions of Hamiltonian couplings (using precomputed integrals) to form the $H|\Psi\rangle$ product. For CISD (singles and doubles), this matrix–vector operation can be optimized to $\mathcal{O}(N_{occ}^2 N_{virt}^2)$ per iteration (scaling roughly as $\mathcal{O}(M^2N^2)$, similar to CCSD) [542]. Higher excitations increase the cost: e.g. CISDT (triples) would involve $\mathcal{O}(N_{occ}^3 N_{virt}^3)$ operations, akin to $\mathcal{O}(n^7$–$n^8)$ scaling, and CISDTQ (quadruples) even more ($\mathcal{O}(n^{10})$ scaling in the worst case).

These steep costs make CISD the highest-level truncated CI feasible for most molecules, and even then typically for small to medium molecules with moderate basis sets. The memory bottleneck is also severe: storing millions of c_I coefficients and addressing them efficiently is challenging. Modern approaches to manage CI spaces include factorizing the determinants (e.g. in spin-coupled basis, or using string-based algorithms) and using distributed parallel algorithms for FCI [541]. There are also specialized CI approaches like *Selected CI* (which selects only the most important determinant excitations iteratively) and *Incremental FCI* or *Quantum Monte Carlo* in FCI space, which aim to tackle larger systems by sampling the determinant space. Nonetheless, the exponential baseline of FCI looms in the background, and these methods only extend the reach somewhat. In summary, the factorial growth of the CI expansion fundamentally limits the method's scalability. This is often visualized as CI providing a systematically improvable but *brute-force* route to exactness, in contrast to coupled-cluster theory which achieves similar accuracy with a clever re-summation of many excitations implicitly via an exponential ansatz.

To illustrate the procedure of an FCI calculation, Algorithm 3 outlines the main steps one would follow to perform a full CI diagonalization for a given system:

Algorithm 3 Full CI (FCI) diagonalization

1: Choose a spin–orbital basis of size M and electron counts (N_α, N_β).
2: Generate all determinants $|\Phi_I\rangle$ with those occupations; $N_{\text{det}} = \binom{M}{N_\alpha}\binom{M}{N_\beta}$.
3: Compute one- and two-electron integrals h_{pq} and g_{pqrs} in that basis.
4: Implement $y = \hat{H}c$ without forming $\hat{H}$: for each J, accumulate $y_J = \sum_I H_{JI} c_I$ using Slater–Condon [19] rules (only I differing by ≤ 2 substitutions couple).
5: Use an iterative eigensolver (e.g., Davidson [539]) with an initial guess (HF) to converge the lowest eigenpair (E_0, c) via repeated $y = \hat{H}c$.
6: Return E_0 and c defining $|\Psi_{\text{FCI}}\rangle = \sum_I c_I |\Phi_I\rangle$ [Eq. (3.44)].

This algorithm yields the exact non-relativistic ground-state energy (within the chosen basis) by construction. In practice, steps 2 and 4 are the most computationally

intensive: step 2 involves combinatorial generation of configurations, and step 4 (within the iterative solver) is where the bulk of CPU time is spent (as each iteration requires looping over millions of determinants for large FCI). Modern codes employ improvements such as symmetry exploitation (using spin, spatial symmetry to block-diagonalize H or to generate configurations more smartly) and efficient integral-direct methods (computing needed integrals on the fly or screening tiny contributions) to push the limits of FCI. Still, the exponential scaling means that FCI becomes intractable beyond perhaps $N \sim 16$ electrons in $M \sim 30$ orbitals even on large parallel computers, which corresponds to very small molecules or minimal basis sets. Truncated CI with singles and doubles (CISD) has a more benign $O(n^6)$ scaling and has been applied to moderately larger systems, but even CISD is generally limited to perhaps tens of electrons in ∼50–100 orbitals maximum, and it still suffers from the qualitative size-consistency problem and the missing correlation from triples, etc.

Full CI provides the exact solution within a finite basis by considering all possible configurations, and truncated CI methods build up this solution by including excitations of increasing rank (CIS, CID, CISD, ...). CI expansions are variationally well-founded and systematically convergent, but suffer from prohibitive exponential cost and lack size-extensivity when truncated. These limitations motivate the introduction of Coupled Cluster theory, which offers an alternative compromise: by using an exponential wavefunction ansatz, CC achieves size-extensive results and faster convergence with excitation level, albeit at the expense of a non-linear, non-variational formulation. The transition from CI to CC can be seen as moving from a linear variational treatment of correlations to a non-linear one that effectively resums infinite subsets of excitation contributions, leading to improved accuracy and consistency for many-electron systems [545]. The strengths and weaknesses of CI discussed here set the stage for understanding why coupled cluster methods have become the preferred high-accuracy single-reference approach in electronic structure theory.

3.1.9 Coupled Cluster Theory: CCSD and CCSD(T)

Coupled–cluster (CC) theory provides a systematically improvable, size–extensive description of electron correlation for many–electron systems [19, 29, 542]. Originating in nuclear physics (Coester and Kümmel) [546, 547] and introduced to quantum chemistry by Čížek, [548] CC employs the exponential ansatz

$$|\Psi_{\mathrm{CC}}\rangle \;=\; e^{T}|\Phi_0\rangle, \qquad T \;=\; T_1 + T_2 + T_3 + \cdots\,, \tag{3.47}$$

where $|\Phi_0\rangle$ is (typically) the Hartree–Fock (HF) reference determinant and T_k creates all k-tuple excitations from $|\Phi_0\rangle$. In second quantization,

$$T_1 = \sum_{i}^{\mathrm{occ}} \sum_{a}^{\mathrm{vir}} t_i^a \, a_a^\dagger a_i, \tag{3.48}$$

$$T_2 = \tfrac{1}{4} \sum_{ij}^{\mathrm{occ}} \sum_{ab}^{\mathrm{vir}} t_{ij}^{ab} \, a_a^\dagger a_b^\dagger a_j a_i, \quad \text{etc.} \tag{3.49}$$

The connected exponential generates products such as $T_1^2/2!$, $T_1 T_2$, ..., which implicitly incorporate higher excitations; this connected–cluster structure enforces *size–extensivity*

(correct scaling with system size), in contrast to truncated CI [29, 542]. CC is not variational (the energy need not be an upper bound), but in practice yields highly accurate ground–state energies for single–reference systems [29].

Projecting the similarity–transformed Hamiltonian

$$\bar{H} \equiv e^{-T} H e^{T} \tag{3.50}$$

onto the reference and excited determinants gives the working equations:

$$E_{\mathrm{CC}} = \langle \Phi_0 | \bar{H} | \Phi_0 \rangle, \tag{3.51}$$

$$0 = \langle \Phi_i^a | \bar{H} | \Phi_0 \rangle, \qquad 0 = \langle \Phi_{ij}^{ab} | \bar{H} | \Phi_0 \rangle, \ \ldots \tag{3.52}$$

where $|\Phi_i^a\rangle$ ($|\Phi_{ij}^{ab}\rangle$) denote singly (doubly) excited determinants. For two–body H, the Baker–Campbell–Hausdorff expansion for $\bar{H}$ terminates at fourth order in nested commutators when only connected terms are retained (linked–cluster theorem) [19, 29].

In CCSD, T is truncated to $T_1 + T_2$ [29, 514]. The coupled, nonlinear amplitude equations for $\{t_i^a, t_{ij}^{ab}\}$ are solved iteratively using orbital–energy denominators and precomputed intermediates to reduce operation counts [19, 542]. With N_o occupied and N_v virtual spatial orbitals, CCSD scales as $\mathcal{O}(N_o^2 N_v^4)$ (commonly reported as $\mathcal{O}(N^6)$) and stores $\mathcal{O}(N_o N_v + N_o^2 N_v^2)$ amplitudes [19, 29]. An explicit spin–orbital expression for the CCSD correlation energy reads (in a canonical MO basis)

$$E_{\mathrm{corr}}^{\mathrm{CCSD}} = \sum_{ia} f_i^{\,a}\, t_i^a + \tfrac{1}{4} \sum_{ijab} \langle ij||ab\rangle\, t_{ij}^{ab} + \tfrac{1}{2} \sum_{ijab} \langle ij||ab\rangle\, t_i^a t_j^b, \tag{3.53}$$

with antisymmetrized two–electron integrals $\langle ij||ab\rangle$ and Fock elements $f_i^{\,a}$ [19, 542]. CCSD is size–extensive and typically recovers the majority of dynamic correlation for single–reference problems [29].

The noniterative triples correction, CCSD(T), [529, 549] evaluates the leading connected T_3 contribution to the energy using perturbation theory on top of converged CCSD amplitudes. The additional cost scales as $\mathcal{O}(N_o^3 N_v^4)$ (overall $\mathcal{O}(N^7)$) [19, 529]. For well–behaved, closed–shell molecules near equilibrium, CCSD(T) routinely achieves "chemical accuracy" and is widely regarded as a practical gold standard for thermochemistry [19, 29, 549]. Breakdown can occur in multi–reference regimes (stretched bonds, diradicals, many transition–metal complexes), for which diagnostics such as the T_1 (and D_1) measures are commonly consulted [29, 550].

Operationally, CCSD solves the nonlinear amplitude equations by fixed–point iteration using a preconditioned update, exactly as outlined in Algorithm 4. Starting from zero guesses for the single and double amplitudes $\{t_i^a, t_{ij}^{ab}\}$, each iteration first builds intermediates, Fock–like and two–electron effective tensors constructed from the current amplitudes, to avoid repeated high–cost tensor contractions. One then forms the *projected residuals* $R_i^a = \langle \Phi_i^a | \bar{H}(T_1, T_2) | \Phi_0 \rangle$ and $R_{ij}^{ab} = \langle \Phi_{ij}^{ab} | \bar{H}(T_1, T_2) | \Phi_0 \rangle$, which are the working expressions of the amplitude equations in Eq. (3.52) with the similarity–transformed Hamiltonian $\bar{H} = e^{-T} H e^{T}$. The amplitudes are updated by dividing residuals by canonical orbital–energy denominators, $(\varepsilon_i - \varepsilon_a)$ for singles and $(\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b)$ for doubles, a preconditioner that accelerates convergence. After each update, the CCSD energy is recomputed from Eq. (3.51), and Pulay's DIIS is usually applied to the amplitude vector to minimize the residual norm and suppress oscillations. Iterations continue until both

the energy change and the residual norms fall below tight thresholds; with standard integral–driven intermediates, as previously noted, the leading cost scales as $\mathcal{O}(N_o^2 N_v^4)$ (often summarized as $\mathcal{O}(N^6)$) and storage as $\mathcal{O}(N_o N_v + N_o^2 N_v^2)$. [19, 29, 29, 542]

Once the CCSD amplitudes have converged, the noniterative triples correction (T) is evaluated *in one shot* using the converged $\{t_i^a, t_{ij}^{ab}\}$ and canonical denominators to yield $E_{(T)}$, after which the final energy is reported as $E_{\mathrm{CCSD(T)}} = E_{\mathrm{CCSD}} + E_{(T)}$. This adds an $\mathcal{O}(N_o^3 N_v^4)$ step (overall $\mathcal{O}(N^7)$) but often delivers "chemical accuracy" for well–behaved single–reference systems; breakdown in strong multireference regimes can be diagnosed though. [29, 529, 549, 550]

Let t_i^a denote the CCSD singles amplitudes (exciting an electron from occupied i to virtual a). Form the rectangular amplitude matrix T_1 with elements $(T_1)_{ai} = t_i^a$. The D_1 *diagnostic* of Janssen and Nielsen is the spectral norm of T_1,

$$D_1 \;=\; \|T_1\|_2 \;=\; \sqrt{\lambda_{\max}\!\left(T_1 T_1^\dagger\right)} \;=\; \sqrt{\lambda_{\max}\!\left(T_1^\dagger T_1\right)}, \tag{3.54}$$

i.e., the square root of the largest eigenvalue of $T_1 T_1^\dagger$. Because it is based on the largest singular value, D_1 is size–intensive and more robust to orbital rotations than simple norms, and it often flags multi–reference character earlier than T_1 [551].

For comparison, the T_1 *diagnostic* of Lee and Taylor measures the root–mean–square magnitude of the singles vector,

$$T_1 \;=\; \sqrt{\frac{\sum_{ia} |t_i^a|^2}{n_{\mathrm{occ}}}}, \tag{3.55}$$

with typical "safe" thresholds $T_1 \lesssim 0.02$ for closed–shell cases [550]. In practice, many users regard $D_1 \lesssim 0.05$ as indicative of reliable single–reference behavior for CCSD(T) in a given basis, whereas larger values (or simultaneously large T_1 and D_1) suggest caution and possible multi–reference treatment [551].

Algorithm 4 CCSD fixed-point iteration

Require: One- and two-electron integrals, canonical orbital energies $\{\varepsilon_p\}$, and HF reference determinant $|\Phi_0\rangle$.

Notation: occupied indices $i, j, \ldots$; virtual indices $a, b, \ldots$; cluster operator $T = T_1 + T_2$; similarity transform $\bar{H}(T) \equiv e^{-T} H e^{T}$; singles/doubles projectors $|\Phi_i^a\rangle, |\Phi_{ij}^{ab}\rangle$.

Ensure: CCSD correlation energy and amplitudes $\{t_i^a, t_{ij}^{ab}\}$.

Initialize amplitudes $t_i^a \leftarrow 0$, $t_{ij}^{ab} \leftarrow 0$; set E_{HF}.

repeat

Build CC *intermediates* (Fock-like and two-electron effective tensors) from current $\{t_i^a, t_{ij}^{ab}\}$ to reuse contractions efficiently.

Form projected residuals (amplitude equations):

$$R_i^a \;=\; \langle \Phi_i^a \,|\, \bar{H}(T_1, T_2) \,|\, \Phi_0 \rangle, \qquad R_{ij}^{ab} \;=\; \langle \Phi_{ij}^{ab} \,|\, \bar{H}(T_1, T_2) \,|\, \Phi_0 \rangle.$$

Update amplitudes with canonical orbital-energy denominators (preconditioned fixed point):

$$\Delta_i^a = \varepsilon_i - \varepsilon_a, \quad \Delta_{ij}^{ab} = \varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b;$$

$$t_i^a \leftarrow t_i^a + \frac{R_i^a}{\Delta_i^a}, \qquad t_{ij}^{ab} \leftarrow t_{ij}^{ab} + \frac{R_{ij}^{ab}}{\Delta_{ij}^{ab}}.$$

Evaluate energy $E_{\mathrm{CCSD}} = \langle \Phi_0 | \bar{H}(T_1, T_2) | \Phi_0 \rangle$ (or equivalently $E_{\mathrm{HF}} + E_{\mathrm{corr}}^{\mathrm{CCSD}}$).

Apply DIIS to the amplitude vector using residuals to accelerate convergence (optional but recommended).

Test convergence: require $|E_{\mathrm{CCSD}}^{(k)} - E_{\mathrm{CCSD}}^{(k-1)}| < \varepsilon_E$ and $\|R\| < \varepsilon_R$.

until both energy change and residual norms fall below thresholds

Output: converged $\{t_i^a, t_{ij}^{ab}\}$ and E_{CCSD}.

(Optional) Noniterative triples: compute $E_{(T)}$ from converged amplitudes with canonical denominators; report $E_{\mathrm{CCSD(T)}} = E_{\mathrm{CCSD}} + E_{(T)}$.

The reader must now have an understanding of the pros and cons of all these methods and be able to understand what sets each of them apart. CI provides a conceptually straightforward, systematically convergent approach to electron correlation, in contrast to mean-field or density functional methods. HF theory, for instance, uses a single Slater determinant optimized by the variational principle, but it includes only the average field experienced by electrons and neglects the instantaneous electron correlation. As a result, HF yields an energy E_{HF} that is above the exact energy, missing the *correlation energy* $E_{corr} = E_0 - E_{\mathrm{HF}}$. A full CI, in principle, recovers 100% of this correlation energy (within the basis) by including all possible excited configurations. In practice, a truncated CI (like CISD) recovers a large portion of E_{corr} (e.g. 90–95%) but not all, whereas a well-chosen coupled cluster method (like CCSD(T)) can often recover a similar >95–99% of E_{corr} with far fewer parameters.

DFT, on the other hand, approaches the correlation problem from a different angle: DFT includes correlation effects through an exchange-correlation functional $E_{\mathrm{XC}}[\rho]$ rather than an explicit multi-determinant wavefunction. DFT calculations are typically as fast as HF (often scaling $O(n^3)$ with system size) and can handle very large systems, which is far

beyond the reach of CI. However, DFT's accuracy depends on the approximation chosen for $E_{\mathrm{X}C}$; there is no systematic, improvable hierarchy akin to CISD $\rightarrow$ FCI. A good DFT functional may yield energies and properties close to correlated wavefunction methods for many systems, but it may also err for others, especially where the single-determinant picture breaks down (multi-reference situations, strong static correlation). CI, by explicitly building a multi-determinant wavefunction, can in principle handle multi-reference cases well (indeed, FCI is the gold standard for any scenario, reproducing correct dissociation limits, etc., within the finite basis). Truncated CI, though, if truncated too low (like CISD), may still fail if the true wavefunction has important higher-order excitations (for example, at a bond dissociation limit, one might need essentially all N-tuple excitations to describe the separated fragments correctly, requiring FCI or a multi-reference extension).

CC theory can be viewed as an alternative approach that uses an exponential ansatz $|\Psi_{CC}\rangle = e^{T}|\Phi_0\rangle$ (with $T = T_1 + T_2 + \cdots$ a cluster operator) instead of the linear expansion of CI. Truncated CC, such as CC with singles and doubles (CCSD) and perturbative triples [CCSD(T)], manages to incorporate effects of higher excitations through the exponential form (for example, T_2 acting twice effectively generates certain quadruple excitation components, etc.). This gives CC a faster convergence with excitation rank and restores size-extensivity, at the cost of losing the strictly variational property. In practical terms, CCSD typically outperforms CISD in accuracy for ground-state energies, despite both using similar computational resources (on the order of $N_{occ}^2 N_{virt}^2$ amplitudes or coefficients) – CCSD recovers more correlation (often 98–99% for normal systems) and is size-extensive. The famous CCSD(T) method (which includes a perturbative estimate of triples) often achieves near chemical accuracy for a wide range of systems by capturing $> 99\%$ of correlation energy, whereas CISD would be several percent off. Thus, while CI is conceptually simple and systematically improvable to exactness, it has largely been supplanted by coupled-cluster methods for quantitative work on larger systems due to the latter's better size-scaling and accuracy for a given truncation level. Nonetheless, CI remains important in quantum chemistry as the foundation for understanding correlation, as a reference for benchmarking (FCI is used to calibrate other methods in small systems), and as a component in multi-reference methods (e.g. MR-CI in combination with Complete Active Space approaches).

Our survey of classical electronic-structure methods is necessarily selective. We omit several important families whose detailed treatment would require substantial additional background and are not directly used in the workflows we benchmark or in the quantum algorithms we develop. For example, Møller–Plesset perturbation theory (MPn) provides a low–cost, single-reference perturbative route that captures dispersion at second order (MP2) but is non-variational and unreliable near degeneracies or for strongly correlated systems, regimes central to some of our test cases and better handled here by CCSD(T) or by active-space approaches [513, 552]. Tensor-network methods, especially the density matrix renormalization group (DMRG), can deliver near-exact results in quasi-1D or weakly entangled orbital orderings, but they entail entanglement-aware orbital localization/ordering and algorithmic choices beyond our current scope [553–555]. Likewise, multireference methods (CASSCF/CASPT2/NEVPT2/MR-CI) are indispensable when static correlation dominates (bond breaking, open-shell transition-metal chemistry), yet our focus is on single-reference benchmarks and on quantum workflows that start from compact active spaces and symmetry-tapered second-quantized Hamiltoni-

ans [276, 556, 557]. Finally, Green's-function and quantum Monte Carlo techniques (e.g., GW/BSE; VMC/DMC/AFQMC) target excited-state spectroscopy or stochastic ground-state projection with distinct error systems and implementation requirements; they fall outside our emphasis on deterministic diagonalization (classical or sample-based) for ground and low-lying states [558, 559]. Where relevant, we might mention comparisons against literature values from these methods but defer a full exposition to the cited works.

The CC idea though is very much relevant to quantum algorithms, as we'll see it motivates the unitary coupled–cluster ansatz $U(\boldsymbol{\theta}) = \exp[T(\boldsymbol{\theta}) - T^{\dagger}(\boldsymbol{\theta})]$ used in quantum VQE [560, 561]. While conceptually attractive, straightforward Trotterizations of UCCSD generally produce circuits too deep for current hardware, motivating adaptive schemes (ADAPT / qubit–ADAPT) discussed later in this thesis [291, 560, 561]. Beyond that, it also remains important in sampling based approaches as a starting point for ansatz construction [42].

3.2 Quantum Computational Chemistry Methods

We now take this discussion to Quantum Computational Chemistry methods. Our focus here, as should be evident from the previous section, is on doing electronic structure calculations. The Variational Quantum Eigensolver (VQE) utilizes similar approaches to the classical algorithms, still requiring an initial guess of the ground state and relying on an iterative approach that leverages the variational principle to converge to the minimum. It also quickly rose to prominence as one of the only ones capable of being reliably executed on present quantum hardware, hence our interest in it.

3.2.1 The VQE algorithm

It was first proposed and showcased by Alberto Peruzzo et al. in 2014 [30] as an alternative to Quantum Phase Estimation (QPE). Both QPE and VQE were developed to solve complex eigenvalue problems, such as those encountered in quantum chemistry. QPE is recognized for its theoretical potential to significantly accelerate calculations beyond classical capabilities by focusing on the relationship between an eigenvector and its corresponding eigenvalue. However, the implementation of QPE faces practical challenges, primarily due to the extensive quantum coherence and the large number of quantum gate operations required, which exceed current technological capabilities in quantum hardware.

Alternatively, VQE represents a hybrid approach that combines quantum and classical computing techniques to estimate the eigenvalues and eigenvectors of Hamiltonians. By utilizing the variational principle, VQE minimizes the energy expectation of a quantum state through an iterative process, where quantum computations are performed to evaluate the energy, and classical algorithms adjust parameters that prepare the quantum states towards minimizing this energy. This method significantly reduces the quantum coherence time needed by leveraging shorter quantum operations. The effectiveness of VQE is partly attributed to its use of parameterized quantum states, known as ansatz states, allowing for a systematic and potentially more efficient exploration of solution spaces with a polynomial number of parameters. A simple pseudocode procedure is described in Algorithm 5

Algorithm 5 Variational Quantum Eigensolver (VQE) Algorithm

Initialize:
- Choose an initial guess for the quantum state parameters, θ_{initial}
- Set convergence criteria $\varepsilon_{\text{energy}}$
- Set maximum number of iterations, *max_iter*

repeat
1. Construct the trial quantum state, $|\psi(\theta)\rangle$, using the current parameters θ:
   $|\psi(\theta)\rangle = U(\theta)|0\rangle$ where $U(\theta)$ is the unitary transformation defined by θ
2. Measure the expectation value of the Hamiltonian:
   $E(\theta) = \langle\psi(\theta)|H|\psi(\theta)\rangle$
3. Use a classical optimizer to update the parameters:
   $\theta_{\text{new}} = \text{optimizer}(\theta, E(\theta))$
4. Check for convergence:
   if $|E(\theta_{\text{new}}) - E(\theta)| < \varepsilon_{\text{energy}}$:
   Convergence achieved, exit loop.
5. If not converged, update θ to θ_{new},
   and go back to step 1 with the updated parameters.

until Convergence or *max_iter* reached

Output:
- Optimized parameters θ
- Minimum energy estimation $E(\theta)$
- Prepared quantum state $|\psi(\theta)\rangle$

To understand how this happens in a quantum chemistry context we'll go over how a typical problem would be setup and what challenges need to be solved.

3.2.1.1 Problem mapping

The first thing we need to understand is the system we are studying. As mentioned in the previous section we'll be looking at molecular systems given by the Born-Oppenheimer approximation as given in equation 3.9. In molecular systems, electrons are fermions and must be described by fermionic operators that follow the Pauli exclusion principle, which means the Hamiltonian is written using fermionic creation and annihilation operators.

These operators do not directly correspond to the operations available on a quantum computer, which manipulates qubits, the basic units of quantum information that are typically manipulated using Pauli operators (X, Y, Z), which obey specific commutation relations and not the anti-commutation of fermionic particles. Let's write these relations, for qubits we have

$$\begin{aligned} [\sigma_x, \sigma_y] &= 2i\sigma_z, \\ [\sigma_y, \sigma_z] &= 2i\sigma_x, \\ [\sigma_z, \sigma_x] &= 2i\sigma_y. \end{aligned} \tag{3.56}$$

Where $[A, B] = AB - BA$ denotes the commutator of operators A and B, while X,Y, and

Z are the usual Pauli operators, for fermions we have

$$\begin{aligned} \{a_i, a_j^\dagger\} &= a_i a_j^\dagger + a_j^\dagger a_i = \delta_{ij}, \\ \{a_i, a_j\} &= \{a_i^\dagger, a_j^\dagger\} = 0. \end{aligned} \tag{3.57}$$

where $\{A, B\} = AB + BA$ is the anti-commutator, and δ_{ij} is the Kronecker delta.

Fortunately for us, Pascual Jordan and Eugene Wigner devised in 1928 a way to map a fermionic Hamiltonian into a Pauli Hamiltonian [192] with what became known as the Jordan-Wigner transformation. For brevity's sake, we'll simply state the mapping,

$$a_j = \left(\prod_{k=1}^{j-1} \sigma_z^{(k)}\right) \sigma_-^{(j)}, \quad a_j^\dagger = \left(\prod_{k=1}^{j-1} \sigma_z^{(k)}\right) \sigma_+^{(j)} \tag{3.58}$$

where

$$\sigma_+ = \frac{1}{2}(\sigma_x + i\sigma_y), \quad \sigma_- = \frac{1}{2}(\sigma_x - i\sigma_y). \tag{3.59}$$

This mapping can be inverted and give us a way to go back as well.

$$\sigma_-^{(j)} = \left(\prod_{k=1}^{j-1} \sigma_z^{(k)}\right) a_j, \quad \sigma_+^{(j)} = \left(\prod_{k=1}^{j-1} \sigma_z^{(k)}\right) a_j^\dagger \tag{3.60}$$

Using these relations, it's even possible to check that all commutation and anti-commutation are still respected before and after the mappings.

Beyond Jordan–Wigner, the *parity* transform encodes cumulative occupations so that qubit j stores the parity $p_j \equiv \sum_{k=0}^{j} n_k \bmod 2$, with $n_k \in \{0, 1\}$ the fermionic occupations.[1] Creation/annihilation operators act locally on qubit j but are accompanied by (i) a *parity* string on a suffix of qubits enforcing anticommutation and (ii) an *update* string that flips parities of qubits whose stored parities depend on n_j. Using set notation $P(j)$ for the parity suffix and $U(j)$ for the update set,[2] The operators take the schematic form

$$a_j^\dagger \;\mapsto\; \tfrac{1}{2}\left[\Big(\prod_{\ell\in U(j)} X_\ell\Big)(X_j - i\,Y_j)\Big(\prod_{\ell\in P(j)} Z_\ell\Big)\right], \qquad a_j \;\mapsto\; \tfrac{1}{2}\left[\Big(\prod_{\ell\in U(j)} X_\ell\Big)(X_j + i\,Y_j)\Big(\prod_{\ell\in P(j)} Z_\ell\Big)\right]. \tag{3.61}$$

Relative to JW (which uses long prefix Z strings), parity relocates the Z string to the *right* of site j and renders global number parity—and, with spin–orbital interleaving, spin parity—as single-qubit Z operators. This makes the two $\mathbb{Z}_2$ parities especially transparent and is routinely exploited to *taper* up to two qubits when N and S_z are fixed [31, 196]. Comparative studies report shorter Pauli strings and advantageous symmetry exposure for many molecular instances when using parity rather than JW [564].

Another possibility is the Bravyi–Kitaev (BK) transform, which balances occupation and parity information across qubits by organizing modes on a binary indexed (Fenwick) tree [193, 195, 563]. Each fermionic operator acts nontrivially only on $\mathcal{O}(\log K)$ qubits: a local X/Y on site j, together with *parity* and *update* strings whose supports are

[1] We follow the conventions in Refs. [195, 562].

[2] $P(j)$ and $U(j)$ are defined by the chosen spin–orbital ordering; explicit constructions are given in [195, 563].

Table 1 – Qualitative comparison of fermion–to–qubit mappings (for K spin–orbitals). "Update" refers to the number of qubits affected when flipping an occupation; "Parity" refers to the number needed to compute fermionic parity. Shorter support generally yields shorter Pauli strings and shallower two-qubit ladders.

Mapping	Update support	Parity support	Notes
Jordan–Wigner (JW)	$\mathcal{O}(1)$	$\mathcal{O}(K)$	Simple; long Z chains ⇒ long Pauli strings on distant modes. Good for linear topologies.
Parity	$\mathcal{O}(K)$	$\mathcal{O}(1)$	Global (spin) parities appear as single Z's ⇒ easy qubit tapering (often up to two). Longer X/Y strings for updates.
Bravyi–Kitaev (BK)	$\mathcal{O}(\log K)$	$\mathcal{O}(\log K)$	Balanced locality; typically shorter strings than JW, easier tapering than JW. More preprocessing/book-keeping and basis changes.

the logarithmic-size ancestor/descendant sets of j in the tree. In the same set notation as Eq. (3.61), one can write

$$a_j^\dagger \;\mapsto\; \tfrac{1}{2}\left[\Big(\prod_{\ell\in U_{\mathrm{BK}}(j)} X_\ell\Big)(X_j - i\,Y_j)\Big(\prod_{\ell\in P_{\mathrm{BK}}(j)} Z_\ell\Big)\right], \tag{3.62}$$

with $|P_{\mathrm{BK}}(j)|, |U_{\mathrm{BK}}(j)| = \mathcal{O}(\log K)$ by construction. Compared with JW's $\mathcal{O}(K)$-length strings, BK reduces operator locality and often the two-qubit gate depth of chemistry circuits, while still exposing $\mathbb{Z}_2$ symmetries amenable to tapering [195,563,565]. Empirically, BK tends to yield measurement groupings with smaller commuting cliques and shallower entangling ladders for electronic-structure Hamiltonians on near-term devices [564].

In practice, each mapping has distinct advantages that are summarized in Table 1. JW is the simplest to implement and reason about, and often works well with chemically ordered orbitals, but its long Z strings can inflate two-qubit depth for excitations and for Trotterized evolution. Parity is especially attractive when exploiting $\mathbb{Z}_2$ symmetries: for typical spin-conserving problems, it places the spin-up and spin-down number parities on single Z operators, enabling deterministic removal (tapering) of up to two qubits when N_α and N_β are fixed. BK provides shorter Pauli strings on average and can reduce CNOT counts in Trotter steps; however, the encoded operators are more intricate (extra basis changes and bookkeeping on the BK tree), which can increase single-qubit overhead and complicate low-level circuit optimizations and some symmetry handlings compared with parity. Ultimately, performance is problem- and hardware-specific, and orbital reordering/localization remains important regardless of mapping [195, 563].

After considering the fermion–to–qubit transform (e.g., Jordan–Wigner [192], parity or Bravyi–Kitaev mappings [193, 195, 562]), we can exploit abelian $\mathbb{Z}_2$ symmetries of the mapped Hamiltonian—notably total electron *parity* and *spin-parity*—to *taper off* qubits. Following Bravyi *et al.* [196], if a set of commuting Pauli symmetry generators $\{S_j\}$ commute with $\hat{H}$ and among themselves, there exists a Clifford unitary U such that $US_jU^\dagger = Z_{q_j}$. Fixing the symmetry eigenvalues $s_j \in \{\pm 1\}$ (determined, e.g., from the Hartree–Fock bitstring) allows the substitutions $Z_{q_j} \mapsto s_j$ and removal of the corresponding qubits q_j. Each independent $\mathbb{Z}_2$ symmetry reduces the qubit count by one, block-diagonalizes $\hat{H}$ into symmetry sectors, and typically shortens the Pauli operator list used for measurement [196].

In addition to these parities, we normally restrict to fixed (N, S_z) sectors (number and spin projection) and, when applicable, project onto irreducible representations of the molecular point group. Such non-abelian symmetry information can further curtail the relevant Hilbert space and reduce qubit requirements or excitation pools for variational algorithms [566]. In practice, we select a mapping that exposes desired symmetries (parity/BK often diagonalize parity operators in shorter strings than JW [195]), compute symmetry eigenvalues from the mean-field reference, apply the Clifford tapering transform [196], and carry the resulting constraints forward into ansatz design and measurement grouping.

3.2.1.2 The ansatz

With the problem of mapping the system to a quantum computer solved let's focus on the trial state $|\psi(\theta)\rangle$. In a quantum computer, this state will be described as a quantum circuit, a sequence of quantum gates and measurements applied to a set of qubits. These are unitary operators that manipulate qubits and enable the execution of quantum algorithms.

Parametrizing quantum circuits involves defining quantum gates or sequences of gates that depend on adjustable parameters. To do this we can use a variety of parametric quantum gates, such as $Rx(\theta)$, $Ry(\theta)$, and $Rz(\theta)$, where θ is the rotation angle, as well as their controlled counterparts. (A controlled gate is a gate that only applies the operation if the control qubit is in state $|1\rangle$). It's also common to include entangling operations in the circuit to generate a broader class of quantum states by introducing non-local correlations between qubits. This ability is essential for VQE. The ground states of many-body quantum systems, such as molecules, often exhibit entanglement. In Figure 10 we can see a 3-qubit parameterized quantum circuit.

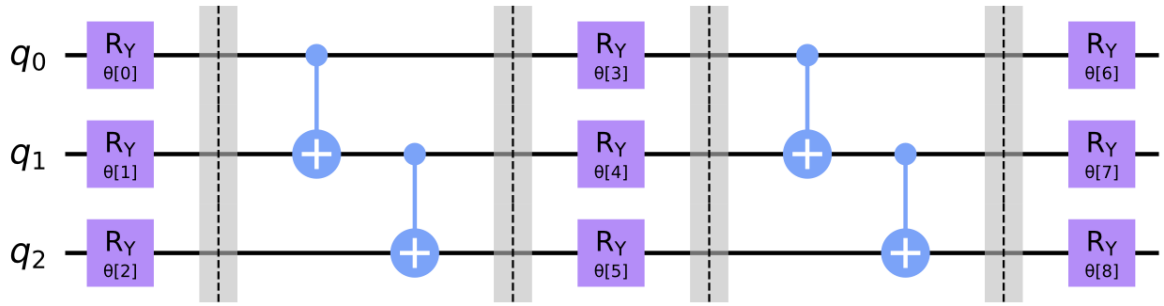


Figure 10 – **A 3-qubit circuit with eight θ parameters which parameterized the Ry rotations.** The circuit also includes CNOTs as the entangling operations.

For chemistry applications there is a commonly used ansatz that is application-inspired, the Unitary Coupled Cluster (UCC) ansatz [561]. The wave function is parametrized as

$$|\Psi(\vec{\theta})\rangle = e^{\hat{T}(\vec{\theta}) - \hat{T}^\dagger(\vec{\theta})}|\Phi_0\rangle \tag{3.63}$$

where the excitation operator to order n is defined as

$$\hat{T}(\vec{\theta}) = \hat{T}_1(\vec{\theta}) + \hat{T}_2(\vec{\theta}) + \cdots + \hat{T}_n(\vec{\theta}),$$

with the first and second-order terms given by

$$\hat{T}_1(\vec{\theta}) = \sum_{i,m} \theta_i^m \hat{a}_m^\dagger \hat{a}_i, \tag{3.64}$$

$$\hat{T}_2(\vec{\theta}) = \frac{1}{2} \sum_{i,j;m,n} \theta_{i,j}^{m,n} \hat{a}_n^\dagger \hat{a}_m^\dagger \hat{a}_j \hat{a}_i. \tag{3.65}$$

Here, the parameter vector $\vec{\theta} = \{\{\theta_i^m\}, \{\theta_{i,j}^{m,n}\}, \ldots\}$ encompasses sets of expansion coefficients. In practical applications, emphasis is typically placed on the Single-Doubles (UCCSD) version, incorporating $\hat{T}_1(\vec{\theta})$ and $\hat{T}_2(\vec{\theta})$; Expanding to include terms beyond these involves the construction of more extensive circuits, which, as of now, are not feasible due to limitations in both experimental setups and computational simulations.

Implementing the UCCSD expansion poses significant challenges, especially concerning the circuit complexity required for the exponential operations detailed in equation 3.63. The depth (circuit depth refers to the number of layers of quantum gates that are executed sequentially in a quantum circuit, representing the time required to perform the computation) of such circuits typically scales with $\mathcal{O}((N_{\text{occ}}^2) \times (N_{\text{virt}}^2) \times N_{\text{qubits}})$, where N_{occ} and N_{virt} denote the counts of occupied and virtual orbitals involved in the excitations, respectively. This makes this ansatz prohibitive for current quantum hardware.

3.2.1.3 Measurement cost and optimization pathologies

VQE runtime is often dominated by measurement. Molecular Hamiltonians decompose into many Pauli strings; naive estimation requires shots scaling with the sum of coefficient variances.

Write the (qubit) molecular Hamiltonian as a weighted sum of Pauli strings,

$$H \;=\; \sum_{k=1}^{M} c_k P_k, \qquad P_k \in \{\mathbb{I}, X, Y, Z\}^{\otimes n}. \tag{3.66}$$

Estimating $E = \langle H \rangle$ by sampling each term in its eigenbasis yields the unbiased estimator $\hat{E} = \sum_k c_k \hat{p}_k$, where $\hat{p}_k$ is the sample mean of ± 1 outcomes for P_k from S_k shots. Since $\mathrm{Var}(\hat{p}_k) = \big(1 - \langle P_k \rangle^2\big)/S_k \le 1/S_k$, the total variance is

$$\mathrm{Var}(\hat{E}) \;=\; \sum_k c_k^2 \, \frac{1 - \langle P_k \rangle^2}{S_k}. \tag{3.67}$$

Minimizing this variance subject to a fixed total shot budget $S = \sum_k S_k$ gives the optimal allocation $S_k \propto |c_k| \sqrt{1 - \langle P_k \rangle^2}$ and the bound

$$S \;\gtrsim\; \frac{1}{\varepsilon^2} \left(\sum_{k=1}^{M} |c_k| \sqrt{1 - \langle P_k \rangle^2} \right)^2, \tag{3.68}$$

to reach root-mean-square error ε on E [39, 40]. For realistic chemical Hamiltonians, the sum in (3.68) can be large, making measurement the dominant VQE runtime cost.

Two widely used approaches compress measurement cost by reusing each shot more effectively:

(i) Commuting-group measurement. Partition $\{P_k\}$ into G mutually commuting groups $\{\mathcal{G}_g\}_{g=1}^{G}$ (e.g., qubit-wise commuting or graph-coloring/minimum-clique-cover groupings) so that all $P_k \in \mathcal{G}_g$ are jointly diagonalizable and measurable in a single rotated basis [39, 567]. If S_g shots are devoted to group g, one obtains

$$\mathrm{Var}(\hat{E}) \;=\; \sum_{g=1}^{G} \frac{1}{S_g} \sum_{k\in\mathcal{G}_g} c_k^2\Big(1 - \langle P_k\rangle^2\Big), \tag{3.69}$$

and the same optimization yields

$$S \;\gtrsim\; \frac{1}{\varepsilon^2}\left(\sum_{g=1}^{G} \sqrt{\sum_{k\in\mathcal{G}_g} c_k^2\Big(1 - \langle P_k\rangle^2\Big)}\;\right)^2. \tag{3.70}$$

The square-root aggregation inside each group lowers the effective prefactor compared with (3.68).

(ii) Shadow-based and IC-POVM estimators. Instead of measuring each P_k directly, *classical shadows* use random (typically local Clifford) measurements and classical post-processing to predict many observables simultaneously. To estimate M bounded observables with additive error ε and failure probability δ, the number of random measurements scales as

$$S \;=\; \mathcal{O}\!\left(\frac{\lambda^2\,\log(M/\delta)}{\varepsilon^2}\right), \tag{3.71}$$

where λ is the (problem-dependent) shadow norm that captures estimator variance [223]. Alternatively, informationally complete POVMs (IC-POVMs) implemented via ancilla-assisted measurements provide a frame allowing joint estimation of all Pauli terms from each shot; properly designed adaptive IC-POVMs can reduce the constant-factor variance relative to Pauli sampling, at the expense of deeper measurement circuits and heavier post-processing [235].

Even with efficient measurement, VQE performance can be limited by the classical optimization landscape. Hardware-efficient ansätze may exhibit barren plateaus, where gradient variances decay exponentially with system size, rendering local optimization ineffective [36, 37]. Chemically motivated ansätze (e.g., UCCSD) often improve trainability but at the cost of deeper circuits (and thus noise). Adaptive strategies such as ADAPT-VQE and qubit-ADAPT-VQE build the ansatz iteratively by adding the operator with the largest energy gradient, achieving compact circuits that mitigate barren-plateau behavior [291, 568]. The trade-off is measurement intensity: each iteration requires estimating many operator gradients (or energy differences) with sufficient precision, so adaptive methods exchange circuit depth for additional sampling.

Measurement requirements scale inversely with ε^2 and with problem-dependent variance prefactors; exploiting commutativity structure, shadow tomography, or IC-POVMs can substantially reduce constants, while ansatz choices balance expressivity, depth, and trainability. These considerations motivate the sampling-based and short-time evolution techniques developed in the next chapters, which retain systematic accuracy control while avoiding fragile global optimizations.

This should be enough to understand how we can assemble all the relevant parts to initiate a VQE calculation. Next, we'll take a look at how to simplify these calculations to be able to use current hardware.

3.2.2 Reducing the problem

In the previous section, we mentioned how circuit depth is closely tied to the number of orbitals involved in the calculation, so one obvious simplification would be to reduce the number of orbitals being considered. This is also a strategy common in classical methods known as Complete Active Space (CAS). It involves selecting a subset of molecular orbitals, known as the active space, which includes all the electrons in these orbitals for the FCI calculations. This selection is critical for capturing important electron correlation effects in complex molecules, especially those involving transition metals or with potential multi-reference character. The choice of orbitals and electrons defines the "active space," and computations are restricted within this space, balancing computational feasibility with accuracy [569].

The same approach can be used for VQE where only a small subset of molecular orbitals will be used for the quantum calculation [405]. This allows us to "embed" HF or even DFT results in the VQE calculation for the inactive space, which in turn ensures that the limited quantum resources are applied to the most relevant part of the calculation. The next problem then becomes choosing which molecular orbitals to include in the active space.

Active Space Selection

As we mentioned before, CAS is a common strategy in classical calculations, which means we can reuse strategies used in classical workflows for VQE. Two common methods are Density Matrix Embedding Theory (DMET) [53] and Atomic Valence Active Space (AVAS) [570].

DMET works by dividing the entire system into fragments or subsystems. Each fragment typically consists of a few orbitals and their associated electrons. For each fragment, a high-level quantum mechanical calculation is performed. This part of the calculation treats the electrons within the fragment explicitly, using accurate quantum chemical methods to account for electron correlation effects robustly. The effects of the rest of the system (the environment) on the fragment are included by constructing an effective Hamiltonian. This Hamiltonian is used to simulate the interactions between the fragment and its environment, crucial for capturing the physics of systems where interactions are not localized just within the fragments but extend into the surrounding material. The core of DMET is the use of density matrices to encapsulate the quantum state of the entire system, even though only fragments are treated explicitly. The local density matrix of a fragment, derived from the wave function of the whole system, informs the boundary conditions or the embedding potential used in the fragment calculation. Usually, when used in combination with VQE, only one small fragment is left for the quantum system [571]. Chapter 4 has a more in-depth explanation of DMET, but considering the relevance to this thesis we do not spend too much time on its description.

AVAS, on the other hand, focuses on constructing a targeted active space for quantum chemical calculations that specifically includes atomic orbitals contributing to bonding or valence states in molecules. AVAS works by projecting the molecular orbitals onto a pre-selected set of atomic orbitals, typically those considered to be involved

in chemical bonding or reactions. This projection results in a subset of orbitals that predominantly reflects the valence character, simplifying the complexity of electronic structure calculations while maintaining the focus on chemically relevant interactions. This approach is particularly useful in transition metal chemistry and other scenarios where understanding the electronic structure around specific atoms is crucial. The AVAS method helps isolate the effect of these valence orbitals, making it easier to analyze and interpret the role of these electrons in chemical properties and reactivity.

Further on Chapter 5 another method that is particularly suited for studying chemical reactions called Local Embedding will be described an utilized in our work [6,250].

3.2.3 VQE variations

Beyond trying to simplify the problem, there is active research in modifying the ansatz construction to have better starting points for the optimization and facilitate convergence by limiting the search space for the ansatz. Below, we describe two such methods that try to iteratively build the ansatz so as not to let it increase to unnecessarily large depths for given convergence criteria. Note that this is not an exhaustive list, and despite the sheer amount of available VQE variations, we want to keep this section brief by tackling only the ones relevant to this work.

3.2.3.1 ADAPT-VQE

The idea behind the Adaptive Derivative-Assembled Pseudo-Trotter ansatz Variational Quantum Eigensolver (ADAPT-VQE) is to start with a simple initial ansatz and iteratively enhance it by incorporating additional operators based on their ability to decrease the system's energy, thus systematically approaching the true ground state energy [568].

The process begins with the selection of a minimal set of fermionic operators that significantly impact the correlation energy of the molecular system. These operators are then added one at a time to the existing ansatz. After each addition, the parameters of the new ansatz are optimized to minimize the Hamiltonian's expectation value. This stepwise refinement continues until the addition of new operators no longer yields significant energy improvements, or until predefined accuracy requirements are met.

What sets ADAPT-VQE apart is its dynamic building of the ansatz which is based directly on the system's response, rather than relying on a predetermined set of excitations like UCCSD. Each iteration involves calculating the gradient of the energy with respect to each operator not yet included in the ansatz, selecting the operator that most reduces the energy, and then recalculating the energy. This method allows ADAPT-VQE to achieve high accuracy with potentially fewer parameters and shallower circuits. The pseudocode description is given in Algorithm 6

Algorithm 6 Adaptive Derivative-Assembled Pseudo-Trotter VQE (ADAPT-VQE)

Initialize:
- Choose a reference state $|\Psi_0\rangle$
- Define a threshold ϵ for acceptable energy improvement
- Initialize an empty ansatz $\mathcal{A}$
- Set maximum number of iterations, *max_iter*

repeat
1. Compute energy gradients:
   - For each operator $\tau \notin \mathcal{A}$, calculate:
     $\nabla E(\tau) = \frac{\partial \langle \Psi(\theta)|H|\Psi(\theta)\rangle}{\partial \tau}$
2. Select operator τ^* with the highest negative gradient:
   - $\tau^* = \arg\min_\tau \nabla E(\tau)$
3. Add τ^* to ansatz $\mathcal{A}$
4. Optimize ansatz parameters to minimize $E = \langle \Psi(\theta)|H|\Psi(\theta)\rangle$ (Usual VQE procedure)
5. Calculate change in energy ΔE
6. Check for convergence:
   if $|\Delta E| < \epsilon$:
   Convergence achieved, exit loop.
7. If not converged, update $\mathcal{A}$ and continue

until Convergence or *max_iter* reached

Output:
- Optimized parameters and the minimal energy
- The refined quantum state $|\Psi(\theta)\rangle$

3.2.3.2 qubit-ADAPT-VQE

There is one more alternative based on the previous variation, the qubit-Adaptive Derivative Assembled Problem-Tailored (qubit-ADAPT) VQE is an evolution on the ADAPT-VQE strategy, specifically tailored to leverage the advantages of present quantum devices more effectively. Unlike the original ADAPT-VQE, which used a pool of fermionic operators requiring complex mappings and potentially deep circuits, qubit-ADAPT employs a pool of simpler Pauli string operators. This shift dramatically reduces the complexity of the quantum circuits involved, particularly in terms of the number of required CNOT gates, which are a common bottleneck in terms of coherence time and error rates on current quantum hardware.

Qubit-ADAPT maintains the systematic, dynamic construction of the wavefunction ansatz characteristic of ADAPT-VQE but does so with a more hardware-efficient approach. Each operator in the pool is evaluated for its effect on the system's energy, and the most impactful operator is iteratively added to the ansatz. This method not only retains the adaptability and precision in targeting the ground state energy but also significantly lowers the circuit depth and the operational overhead, making it better suited for the limited capabilities of current devices. Due to the similarity to ADAPT-VQE, we'll omit the pseudocode for this variation.

Furthermore, a new completeness condition introduced with qubit-ADAPT ensures

that the pool of operators is sufficient to potentially converge to the exact ground state. The proof provided with qubit-ADAPT shows that the minimal number of operators needed in the pool grows linearly with the number of qubits [291].

3.3 Adapted Article: Exploring qubit-ADAPT-VQE for materials discovery

Previous sections established the second-quantized setting, our spin/orbital labeling conventions, and the measurement-aware mindset that motivates sample-based solvers later in the thesis. In this chapter I revisit a published case study through that common lens: qubit-ADAPT-VQE as a compact, hardware-attentive route to adsorption energetics in porous materials. The presentation below adopts the notation and assumptions used elsewhere in the thesis and foreshadows the sampling-centric strategies developed in Chapters 4–6.

Provenance and reuse. This chapter reworks material originally published as: M. A. Barroca, R. N. B. Ferreira, and M. Steiner, *Exploring qubit-ADAPT-VQE for materials discovery*, APL Quantum **1**, 046103 (2024), DOI: 10.1063/5.0219500. The article is open access under a Creative Commons Attribution (CC BY 4.0) license (link). Text is independently rewritten but figures, tables and numerical results are reproduced as in the original. This use complies with the CC BY terms, with full attribution to the version of record.

Abstract. Direct air capture (DAC) of carbon dioxide is a promising method for mitigating climate change. Solid sorbents, such as metal-organic frameworks, are currently being tested for DAC application. However, their potential for deployment at scale has not been fully realized. The computational discovery of solid sorbents is challenging, given the vast chemical search space and the DAC requirements for molecular selectivity. Quantum computing can potentially accelerate the discovery of solid sorbents for DAC by predicting molecular binding energies. In this work, we explore algorithms for predicting gas adsorption in metal-organic frameworks using a quantum computer. Specifically, we simulate the potential energy surfaces of CO_2, N_2, and H_2O molecules at the Mg^{+2} metal center that represents the binding sites of typical metal-organic frameworks. We apply the qubit-ADAPT-VQE technique to run simulations on both classical and quantum computing hardware, and achieve reasonable accuracy while maintaining hardware efficiency.

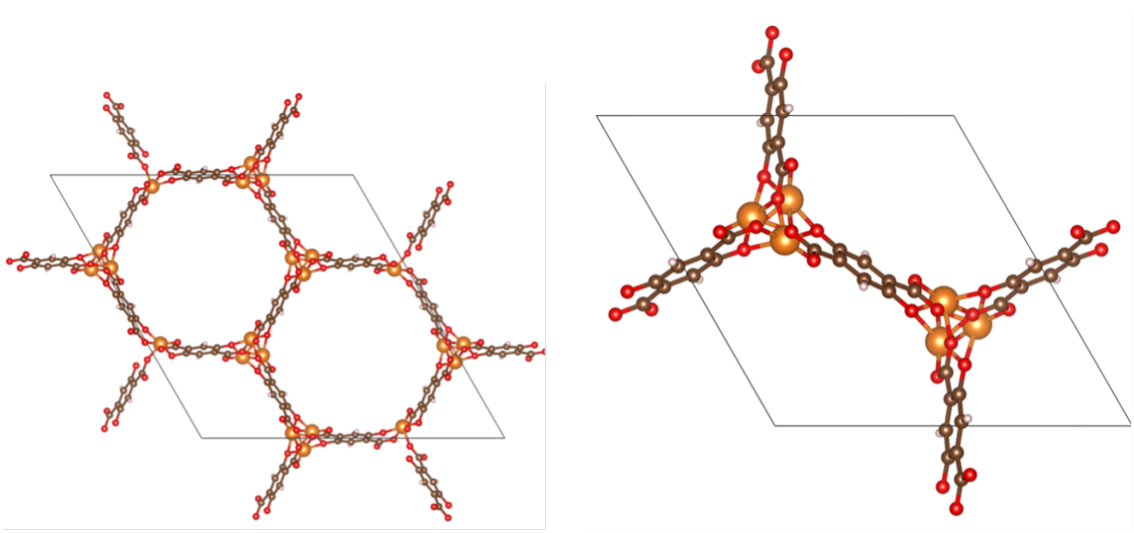
3.3.1 Problem Statement

Technologies that capture greenhouse gases (GHGs) are a central pillar of climate-mitigation portfolios because they directly lower atmospheric concentrations and help constrain warming [572]. One pragmatic pathway is to retrofit existing fossil-based power infrastructure for post-combustion CO_2 capture. Doing so extends the useful lifetime of installed assets while reducing their contribution to the global carbon budget. Yet eliminating ongoing emissions alone will not suffice to meet temperature targets: active

removal of CO_2 from ambient air is required [573]. DAC addresses this need by using sorbent materials that selectively bind CO_2 directly from the atmosphere.

Liquid sorbents, especially amine solutions, are mature and widely deployed for CO_2 capture [574]. They offer rapid uptake kinetics and high saturation capacities, but their drawbacks are well known: corrosion and a large regeneration heat duty that typically accounts for 60–80% of total CCS operating costs [575, 576]. Porous solid sorbents provide an alternative route that can mitigate these issues. This materials class spans zeolites, metal–organic frameworks (MOFs), covalent–organic frameworks, zeolitic imidazolate frameworks, and porous polymer networks, among others. Because adsorption can proceed via physisorption (i.e., without forming chemical bonds) regeneration energies (and therefore operating costs) can be markedly reduced [577]. On the other hand, solid sorbents often exhibit sensitivity to water vapor: competitive H_2O adsorption can lower CO_2 selectivity, and water can passivate sites, hindering full regeneration [578].

Among candidate solids, MOFs stand out for their tunable pore architectures and vast chemical diversity [579]. Built from metal nodes and organic linkers, MOFs assemble into nanoporous networks with channel diameters ranging from a few to tens of ångströms. Crucially, geometric and chemical features can often be varied semi-independently to optimize both absolute capacity and relative selectivity for a target separation [580]. This flexibility creates a computational challenge: the combined space of hypothetical [581, 582] and experimentally realized [583] MOFs now comprises on the order of 10^5–10^6 candidates, complicating down-selection.



(a) (b)

Figure 11 – **Mg-MOF-74 structure:** (a) conventional unit cell (162 atoms) and (b) primitive unit cell (54 atoms). *Reproduced from* M.A. Barroca, et al., [5] (DOI) CC BY 4.0 (link).

Although interest in using quantum computers for carbon-capture modeling is rising, the published literature remains sparse. Most prior work targets methods capable of treating large MOF fragments. A prominent example is DMET, which, as discussed previoulsy, partitions the crystal into fragments handled at high fidelity while the remainder is represented approximately. This construction enables accurate treatment of strong interactions and entanglement at reduced cost [265, 266]. DMET has been combined with

quantum algorithms by assigning the fragment (high-level) step to a quantum backend while the environment remains classical, thereby focusing quantum resources on chemically relevant regions [584, 585]. In contrast, the present study examines how the high-level step itself can be improved by adopting a modern variant of the VQE. [30, 560, 586]

Here we carry out quantum-chemistry calculations on both classical simulators and superconducting-qubit processors available via the IBM Quantum platform to estimate binding potentials for representative atmospheric molecules—CO_2, H_2O, and N_2—in a simplified model of adsorption at a Mg^{2+} metal center that abstracts common MOF binding sites. We explore VQE variants, with emphasis on qubit-ADAPT-VQE, to simulate gas adsorption in MOFs on contemporary hardware and to assess the approach's promise for accelerating DAC-relevant materials discovery. In addition, we provide a systematic comparison across VQE methods, outlining their respective advantages and limitations, and we highlight avenues for improvement and future investigation. To constrain chemical scope while retaining relevance, we focus on the X-MOF-74 family (X = Mg^{2+}, Ni^{2+}, Fe^{2+}, Co^{2+}, Zn^{2+}, Mn^{2+}, Cu^{2+}) as a template for probing MOF diversity. MOF-74 is characterized by coordinatively unsaturated (open) metal sites that can form bonds with adsorbates; such chemisorption strongly affects adsorption strength and selectivity, especially at low pressures, and largely sets the energetic cost of desorption [577, 578].

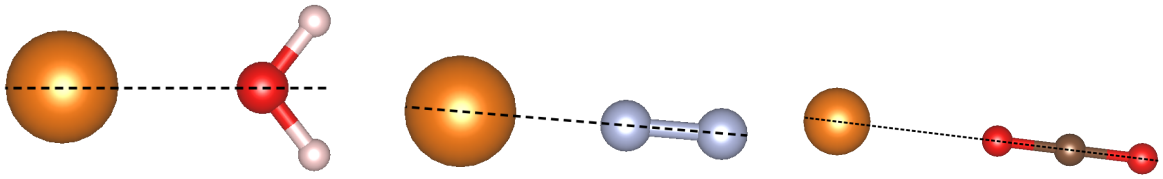
3.3.2 Methods

We use three complementary modalities: classical simulations, quantum simulations, and quantum experiments to generate and analyze adsorption energetics. Classical simulations are standard electronic-structure calculations on a conventional computer, driven by PySCF [587, 588]. Quantum simulations evaluate VQE circuits in noiseless emulation on classical hardware using Qiskit [589, 590]. Quantum experiments execute the same VQE workflows and circuits on superconducting devices accessed via IBM Quantum; as such, they are inherently subject to hardware noise.

Because the Mg-MOF-74 primitive cell in Fig. 11 contains more than 50 atoms, we adopt a further reduction: a single metal ion that represents an open metal site interacting with the adsorbate. This choice reflects current hardware limits for reliable circuit execution given width and depth. A more faithful MOF treatment would include additional atoms around the metal center, but that pushes quantum resource requirements beyond the reach of available devices [585]. Within this simplified model, we compute potential-energy surfaces (PES) for multiple metal/molecule pairs and extract equilibrium separations and binding energies for the systems shown in Fig. 12. The aim is a proof-of-concept for how quantum computing can contribute to DAC materials discovery.

Species selection

Given the central role of the metal node in DAC-relevant MOFs, we focus on the Mg^{2+} ion paired with each of the three probe molecules (H_2O, CO_2, N_2), yielding six distinct systems. In brief, Mg^{2+}+H_2O is expected to be the simplest case because it has the fewest orbitals.



(a) (b) (c)

Figure 12 – **Molecular geometries and bond axes used in the PES calculations for (a) H_2O, (b) N_2 and (c) CO_2 molecules in the vicinity of a metal ion M.** For H_2O, the ion is moved along the C_2 rotation axis. For N_2, the ion is moved along the N-N axis. For CO_2, the ion is moved along the O-C-O axis. *Reproduced from* M.A. Barroca, et al., [5] (DOI) CC BY 4.0 (link).

Classical simulation methods

Two levels of theory are employed in PySCF: Hartree–Fock (HF) and density functional theory (DFT). All calculations, classical and quantum, use the same 6-31G* basis set; DFT adopts the B3LYP functional. These choices mirror the methodology used to construct the 3D geometries (SD files) for H_2O, CO_2, and N_2 obtained from the NIST Chemistry WebBook, which also supplies the molecular structures used herein [591]. The classical calculations serve as references and as inputs for quantum workflows, since the one- and two-electron integrals required for the second-quantized fermionic Hamiltonian are derived from them.

Quantum simulations and experiments

We examine three VQE variants:

- **UCCSD benchmark.** A standard VQE using an application-specific, compute-intensive UCCSD ansatz [592], treated as the quantum benchmark because it typically yields the best accuracy for this setting. It is, however, difficult to deploy on today's hardware due to circuit complexity. Operator ordering matters; we apply double excitations before singles, following Ref. [593].
- **Hardware-efficient heuristic.** A standard VQE with a general-purpose, hardware-efficient ansatz [594] that is tractable on current devices. In our case, layers of R_Y and R_Z rotations (denoted $R_Y R_Z$) converge to the target energies, albeit typically requiring more iterations.
- **qubit-ADAPT-VQE.** A modified ADAPT-VQE [568] that builds the ansatz from a pool of qubit-space operators. This qubit-ADAPT-VQE achieves UCCSD-like accuracy with substantially reduced hardware demands [291].

For each variant we perform noiseless Qiskit simulations [589, 590] and execute the same circuits on IBM Quantum hardware to assess noise impact under otherwise identical workflows.

Table 2 – The qubit-ADAPT-VQE ansatz converges in 3x and 116x fewer iterations, and is 5x and 25x faster than the UCCSD-VQE and $R_Y R_Z$-VQE ansatze, respectively. The number of iterations for qubit-ADAPT-VQE is the sum of all VQE iterations. *Reproduced from* M.A. Barroca, et al., [5] (DOI) CC BY 4.0 (link).

	UCCSD-VQE	$R_Y R_Z$-VQE	qubit-ADAPT-VQE
Run time (s)	25	151	6
Total # of iterations	133	4892	42
Energy Value (Ha)	-56.93440792	-56.93383614	-56.93411960

Active space selection

To contain VQE cost, we construct compact active spaces using PySCF's atomic valence active space (AVAS) procedure [570]. AVAS automatically identifies chemically relevant valence orbitals(e.g., transition-metal d functions or bond-dissociation orbitals) as the seed for an active space suitable for strongly correlated problems where single determinants fail. The approach leverages DMET ideas to determine a minimal set of orbitals sufficient to capture essential physics, with a target size no more than roughly twice the initial set. Unlike fragment–bath partitions, AVAS keeps the occupied/virtual character of active orbitals, enabling natural truncation while preserving key determinants.

We recompute the active space at each metal–molecule separation. At large separations the molecule remains near its equilibrium configuration with essentially unperturbed orbitals; as the ion approaches, the active space transitions to those orbitals participating in bond formation, reflecting the evolving metal–adsorbate interaction.

3.3.3 Results and discussion

Convergence studies

We first benchmark qubit-ADAPT-VQE against two standard VQE baselines. Figure 13 compares all three approaches in terms of total VQE iterations, convergence behavior, and accuracy relative to the ground-state energy obtained by direct diagonalization of the second-quantized Hamiltonian in NumPy (formally equivalent to FCI) [18]. The test case is the $Mg^{2+} + H_2O$ model at a separation of 1.9 Å with an active space of six electrons in seven molecular orbitals, i.e., $(6, 7)$. Table 2 lists wall-clock runtime and iteration counts to convergence for each ansatz. In this setting, qubit-ADAPT-VQE requires 3× and 64× fewer iterations than UCCSD-VQE and $R_Y R_Z$-VQE, respectively, and is 5× and 12× faster in runtime. Qubit-ADAPT-VQE resets the iteration count each time the ansatz is enlarged.

Beyond faster convergence, qubit-ADAPT-VQE is markedly more hardware-efficient. Table 3 reports circuit sizes—expressed in native gate counts and transpiled depth for IBM backends—evaluated at the largest circuit depth reached by qubit-ADAPT-VQE. The most consequential metric is the CNOT count, since two-qubit operations typically have error rates orders of magnitude higher than single-qubit gates.

The observed advantage arises because qubit-ADAPT-VQE grows the ansatz on

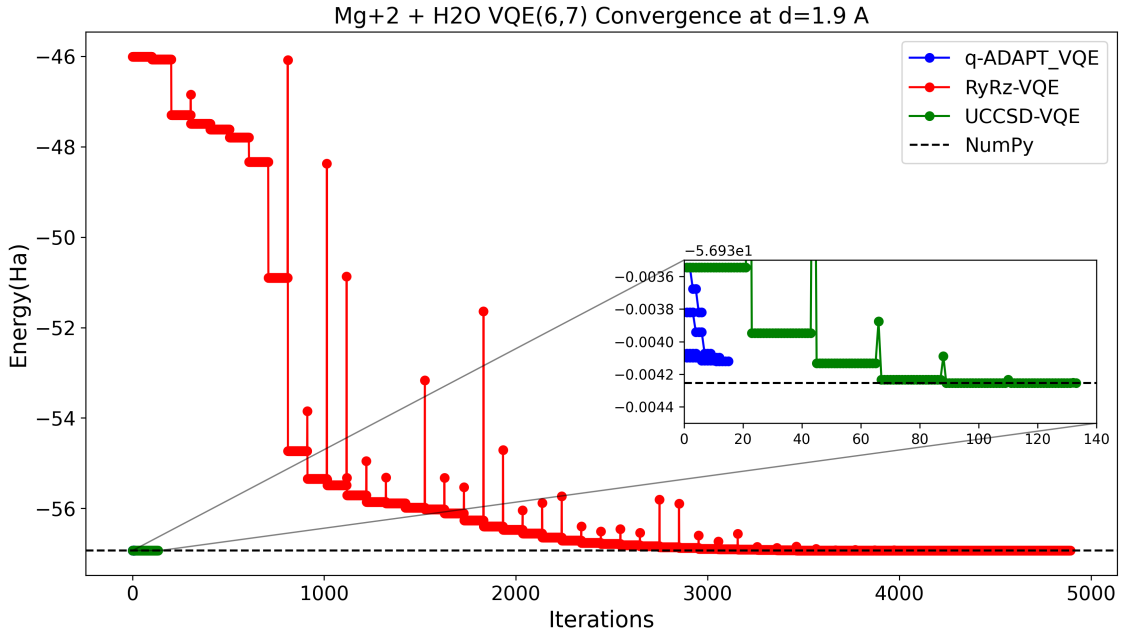


Figure 13 – **Convergence to the ground-state energy from quantum simulations using UCCSD-VQE (green), $R_Y R_Z$-VQE (red), and qubit-ADAPT-VQE (blue) as a function of the total number of VQE iterations.** The black line shows the target from direct diagonalization of the fermionic Hamiltonian in NumPy (FCI-equivalent). All ansätze converge to within chemical accuracy ($\sim$ 1 mHa) of the target. Notably, qubit-ADAPT-VQE reaches this tolerance in far fewer iterations than UCCSD-VQE and $R_Y R_Z$-VQE. Energies correspond to the $(6, 7)$ active-space calculation. *Reproduced from* M.A. Barroca, et al., [5] (DOI) CC BY 4.0 (link).

Table 3 – **Comparing the circuit sizes generated by different ansätze.** qubit-ADAPT-VQE manages to get satisfactory results while using the shallowest circuits. Circuits are described in terms of the number of gate operations and depth required to run on IBM quantum hardware. *Reproduced from* M.A. Barroca, et al., [5] (DOI) CC BY 4.0 (link).

	UCCSD-VQE	$R_Y R_Z$-VQE	qubit-ADAPT-VQE
# of CNOT gates	1610	27	20
# of Rz gates	1191	125	37
# of Sx gates	878	80	19
# of X gates	44	5	1
Depth	3255	35	30

demand from an operator pool, adding operators only when convergence thresholds warrant expansion. Our pool is generated by (i) constructing a UCCSD ansatz, (ii) mapping the fermionic operators to qubit operators via the parity transform, and (iii) discarding Pauli strings with an even number of Y operators, retaining only the "odd" strings that contribute to the energy [291]. We further remove Z-chains longer than two after verifying they yield negligible improvement. The parity mapping is selected over Jordan–Wigner because it preserves particle-number parity for each spin sector and enables the removal of

two redundant qubits for these problems [195]. After testing several optimizers (BFGS, COBYLA), we adopt SLSQP due to faster convergence in our experiments. Convergence thresholds are set to: maximum outer iterations = 4, maximum operator gradient = 10^{-4}, and maximum change in the energy expectation value between consecutive iterations = 10^{-3}. These settings achieve chemical accuracy ($\sim 10^{-3}$ Ha) relative to UCCSD-VQE and NumPy references, and are used for all qubit-ADAPT-VQE runs.

No noise model is included in the simulations: our primary goal is to compare ansätze and workflows against the NumPy diagonalization benchmark on a level playing field. Only once convergence/accuracy are established do we repeat the best-performing procedure on quantum hardware.

Active-space size is chosen to keep quantum runs tractable. When feasible we use the AVAS procedure; otherwise, if the implied active space would exceed 15 qubits, we default to a $(6, 7)$ space centered on the HOMO–LUMO region with orbitals ordered by energetic contribution. This choice uses 10–11 qubits and balances fidelity with runtime. All VQE and PySCF calculations are executed on a laptop (Apple MacBook Air M1 2020) in a slim, reproducible software environment. Please refer to the "Code availability" section for further details.

Mg^{+2} systems

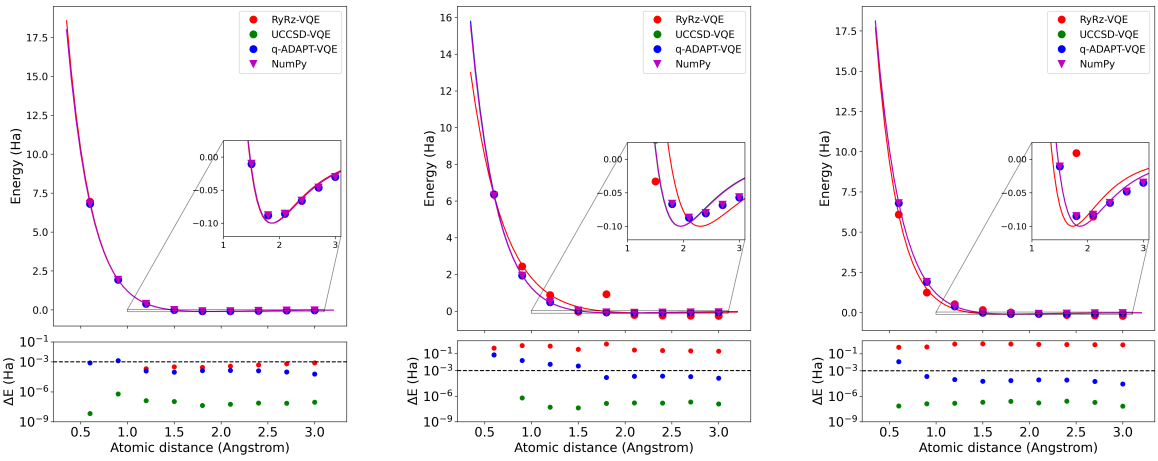
Given the strong cost–benefit demonstrated by qubit-ADAPT-VQE in simulation, we selected it as the primary candidate for execution on quantum hardware. To make these runs feasible on currently available devices, we further reduced the active space to three electrons in four molecular orbitals $(3, 4)$, which maps to just 4 qubits. Although the hardware nominally supports larger qubit counts, the circuit depths required for this application, absent fault tolerance, arere not reliably executable. As before, the molecular orbitals are chosen around the HOMO–LUMO region and ordered by energy. We cap each execution at 1024 shots; depending on the metal–molecule separation, a single point requires five to eight hours. With an average of nine points per PES, a full surface demands roughly 40–72 hours of compute time.

This 4-qubit restriction led us to confine hardware VQE calculations –hereafter 'hardware experiments' – to Mg^{2+} systems. In this configuration, the active space contributes only $\approx 5\%$ of the absolute ground-state energy, so extending to heavier ions such as Zn^{2+} would reduce the active-space share even further and yield diminishing returns. All hardware experiments use Qiskit Runtime on `ibmq_montreal` with Twirled Readout Error Extinction (T-REx) for error mitigation [595].

We construct PES curves by evaluating the energy along the bond axis for metal–molecule separations from 0.3 to 3.6 Å in 0.3 Å increments. After sampling, we fit a Morse potential [596] to extract equilibrium distances and binding energies; plotted energies are shifted so that the asymptotic (infinite-separation) limit is set to zero, simplifying the readout of binding energies.

For Mg^{2+} paired with each of the three probe molecules, we simulate the three VQE variants: $R_Y R_Z$-VQE, UCCSD-VQE, and qubit-ADAPT-VQE in a $(6, 7)$ active space and assess accuracy against direct diagonalization (NumPy). In Fig. 14, both qubit-ADAPT-VQE and UCCSD-VQE reproduce the NumPy benchmark to within chemical

accuracy, with UCCSD-VQE reaching slightly tighter agreement ($\Delta E \lesssim 10^{-6}$ Ha). By contrast, $R_Y R_Z$-VQE tracks the benchmark only for the simplest case, $Mg^{2+} + H_2O$; for N_2 and CO_2 it overestimates energies left of the minimum and underestimates to the right, degrading the Morse fit and the derived observables.



(a) (b) (c)

Figure 14 – **Potential energy surfaces and energy differences with respect to direct diagonalization with NumPy, for (a) $Mg^{+2} + H_2O$, (b) $Mg^{+2} + N_2$ and (c) $Mg^{+2} + CO_2$ systems.** The black dashed line in the ΔE plot represents the chemical accuracy threshold. Noiseless quantum simulations of UCCSD-VQE (green circles) and qubit-ADAPT-VQE (blue circles) match the NumPy (magenta triangles) benchmark in all three cases, however, $R_Y R_Z$-VQE (red circles) only matches the benchmark for the simplest system, namely, $Mg^{+2} + H_2O$. The black dashed line in the ΔE plot represents the chemical accuracy threshold. *Reproduced from* M.A. Barroca, et al., [5] (DOI) CC BY 4.0 (link).

On the basis of Fig. 14 (accuracy) together with Table 2, Table 3, and Fig. 13 (efficiency), we proceed with qubit-ADAPT-VQE on hardware. Consistent with the device constraints, we restrict the hardware demonstration to the smallest system, $Mg^{2+} + H_2O$, and use the $(3, 4)$ active space.

The results in Fig. 15 include a direct comparison to hardware. While the raw hardware energies remain within 10^{-1} Ha of the benchmark, the Morse fit to those data overlays closely with the NumPy and simulated-VQE fits, despite the noise intrinsic to present-day devices. Each hardware point uses 1024 shots with T-REx mitigation; increasing the shot count did not improve accuracy, and more aggressive techniques such as zero-noise extrapolation increased variance beyond acceptable levels.

In these plots, solid markers denote sampled points and solid lines denote Morse fits. Circles represent VQE simulations, triangles the NumPy diagonalization, squares PySCF calculations, and diamonds the hardware experiments.

From the fitted Morse curves for $Mg^{2+} + H_2O$, we extract observables (equilibrium distance and binding energy) summarized in Table 4. Although the hardware results do not achieve chemical-accuracy agreement, qubit-ADAPT-VQE simulations and NumPy match closely, and the `ibmq_montreal` values agree reasonably well given the larger uncertainties

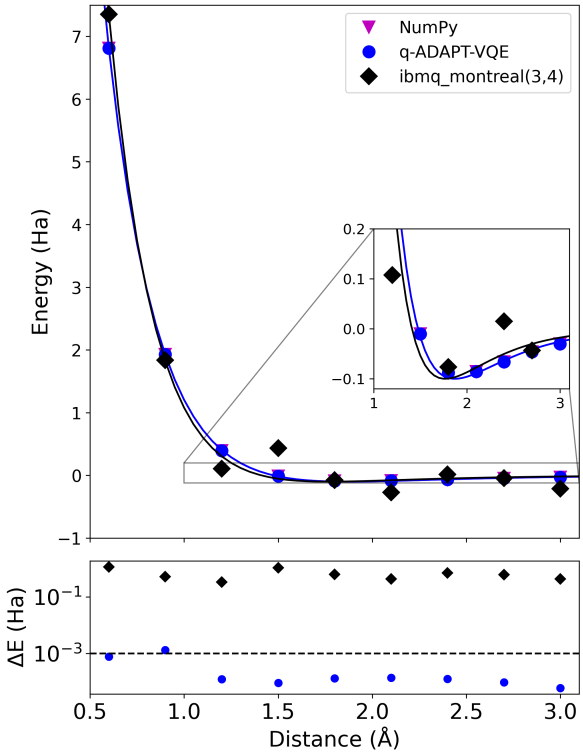


Figure 15 – **Potential Energy Surfaces and energy differences with respect to direct diagonalization with NumPy (purple triangles) for VQE quantum experiments on hardware (black diamonds) and qubit-ADAPT-VQE simulation (blue circles).** The black dashed line in the ΔE plot represents the chemical accuracy threshold. The adjusted Morse potentials exhibit a considerable overlap, despite the level of noise present in the hardware data points, which leads to a reasonable agreement between the observables shown in Table 4. *Reproduced from* M.A. Barroca, et al., [5] (DOI) CC BY 4.0 (link).

induced by hardware noise.

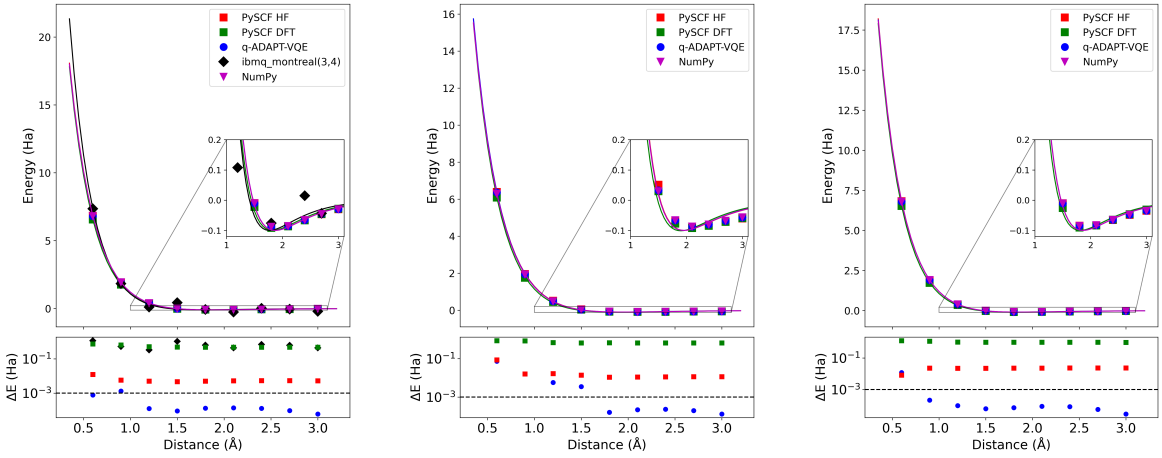
We next compare PES from classical PySCF methods (HF, DFT) with the best simulated quantum approach (qubit-ADAPT-VQE, $(6, 7)$). For $Mg^{2+} + H_2O$, Fig. 16a also includes the hardware experiment (qubit-ADAPT-VQE on `ibmq_montreal`, $(3, 4)$). HF does not achieve $\Delta E \lesssim 10^{-3}$ Ha, whereas qubit-ADAPT-VQE is the only method to reach chemical accuracy. DFT deviates by 10^{-1} Ha across these systems.

Despite noise, the hardware PES aligns reasonably with classical references. This is evident visually in Fig. 16 and quantitatively in Table 5. Binding energies agree within the quoted uncertainties across methods; equilibrium distances from qubit-ADAPT-VQE and HF coincide, and both are 2–3% larger than the DFT predictions.

Tightening qubit-ADAPT-VQE convergence thresholds would likely reduce residual errors further, but since chemical accuracy is already achieved we prioritize shallow circuits and hardware efficiency. Finally, we note that our error-mitigation choices trade some precision for accuracy; given the observed uncertainty levels, we avoided more advanced strategies that would have increased variance.

Table 4 – **Comparison of Equilibrium Distance and Binding Energy for Mg^{+2} + H_2O obtained from the fitted Morse potential to results obtained with different methods.** The uncertainties represent the error bars of the curve fit parameters. Taking the fit uncertainties and the inherent noise of the quantum hardware data points, we consider all methods to be in agreement. *Reproduced from* M.A. Barroca, et al., [5] (DOI) CC BY 4.0 (link).

	`ibmq_montreal`	qubit-ADAPT-VQE	NumPy
Equilibrium Distance (Å)	1.8 ± 0.3	1.87 ± 0.01	1.87 ± 0.01
Binding energy (Ha)	-0.1 ± 0.4	-0.1 ± 0.02	-0.1 ± 0.02



(a) (b) (c)

Figure 16 – **Potential energy surfaces and energy differences with respect to direct diagonalization with NumPy, for (a) Mg^{+2} + H_2O, (b) Mg^{+2} + N_2 and (c) Mg^{+2} + CO_2 systems.** The black dashed line in the ΔE plot represents the chemical accuracy threshold. Results obtained from quantum experiments on hardware using (3,4) active space (black diamonds) and quantum simulations in a (6,7) active space (blue circles) compared to PySCF-based, all-electron classical simulations using the HF (red squares), DFT (green squares) and NumPy direct diagonalization (purple triangles). The simulated qubit-ADAPT-VQE result improves upon the HF reference for all three systems. The result from the hardware experiments, despite their inherent noise and worse energy values, also shows an accurate equilibrium distance within 0.1 Å of both the simulated VQE and the classical methods. *Reproduced from* M.A. Barroca, et al., [5] (DOI) CC BY 4.0 (link).

Finally, we note that experimental binding energies for CO_2 on Mg-MOF-74 fall in the 39–47 kJ/mol range [597]. This corresponds to 15–18 mHa per molecule, about six times smaller than our values, highlighting the limitations of reducing the Mg-MOF-74 structure to a single-ion model.

Table 5 – **Comparison of Equilibrium Distances and Binding Energies for Mg^{+2}-based systems obtained from the fitted Morse potential across different methods, both quantum (simulations and hardware experiments) and classical.** The Binding Energy values all agree to within the numeric uncertainty. The equilibrium distances agree between qubit-ADAPT-VQE and HF, but are 2-3% higher than what is predicted by DFT. *Reproduced from* M.A. Barroca, et al., [5] (DOI) CC BY 4.0 (link).

$Mg^{+2} + H_2O$				
	montreal	q-ADAPT-VQE	HF	DFT
Equilibrium Distance (Å)	1.8 ± 0.3	1.87 ± 0.01	1.868 ± 0.004	1.83 ± 0.02
Binding energy (Ha)	-0.1 ± 0.4	-0.10 ± 0.02	-0.10 ± 0.01	-0.1 ± 0.02
$Mg^{+2} + N_2$				
	montreal	q-ADAPT-VQE	HF	DFT
Equilibrium Distance (Å)	-	1.95 ± 0.04	1.96 ± 0.05	1.90 ± 0.07
Binding energy (Ha)	-	-0.10 ± 0.04	-0.10 ± 0.06	-0.1 ± 0.08
$Mg^{+2} + CO_2$				
	montreal	q-ADAPT-VQE	HF	DFT
Equilibrium Distance (Å)	-	1.86 ± 0.02	1.86 ± 0.02	1.81 ± 0.02
Binding energy (Ha)	-	-0.10 ± 0.02	-0.10 ± 0.02	-0.1 ± 0.02

3.3.4 Conclusion

We've leveraged the well-established qubit-ADAPT-VQE technique to explore a comparatively underexamined application: materials discovery for DAC. While the method's efficiency and accuracy are known, our results show that, in this context, qubit-ADAPT-VQE attains chemical accuracy relative to UCCSD-VQE while remaining hardware-efficient. In practice, this yields an advantage on present-day processors by pairing a problem-tailored ansatz with shallow circuits. We further demonstrate feasibility on current hardware despite noise, and we show that fitting the hardware data with a Morse potential brings it into agreement with Hartree–Fock predictions and with noiseless simulations.

There is headroom for improvement. More robust error-mitigation strategies (e.g., Zero Noise Extrapolation (ZNE)) may reduce bias, and larger systems with expanded active spaces should sharpen the preliminary trends reported here; doing so, however, will first require improved active-space selection techniques.

Across Mg^{2+}-based systems we consistently observed less satisfactory behavior for N_2 adsorption: both accuracy and precision deteriorate relative to the other probe molecules. A plausible explanation is the triple bond, which involves six electrons and is perturbed en bloc by the approaching cation; when calculations are restricted to an active space, the $(6, 7)$ choice is essentially consumed by those six electrons, leaving other relevant degrees of freedom outside the treatment.

We also find a systematic overestimation of Hartree–Fock equilibrium distances rel-

ative to DFT by under 5%. The discrepancy is plausibly attributable to exchange–correlation contributions absent in HF and merits further investigation.

Present hardware limits are primarily twofold: reliability (noise resilience) and throughput (the rate at which circuits can be executed and measured). Continued growth in available quantum volume and circuit-layer operations per unit time should in the future enable larger-scale computations of CO_2 capture in MOFs for DAC.

Overall, our work provides a framework that highlights both practical gains and current constraints. It offers a viable path to applying quantum computing to DAC-oriented materials discovery that can scale as devices improve and as techniques for handling larger systems and noisy runs mature. The approach is also compatible with established multiscale workflows such as DMET; qubit-ADAPT-VQE can be slotted in as a high-level solver, though integrating it in that fashion lies beyond the scope of the present study.

3.4 Post-publishing extension: Zn^{2+} adsorption complexes

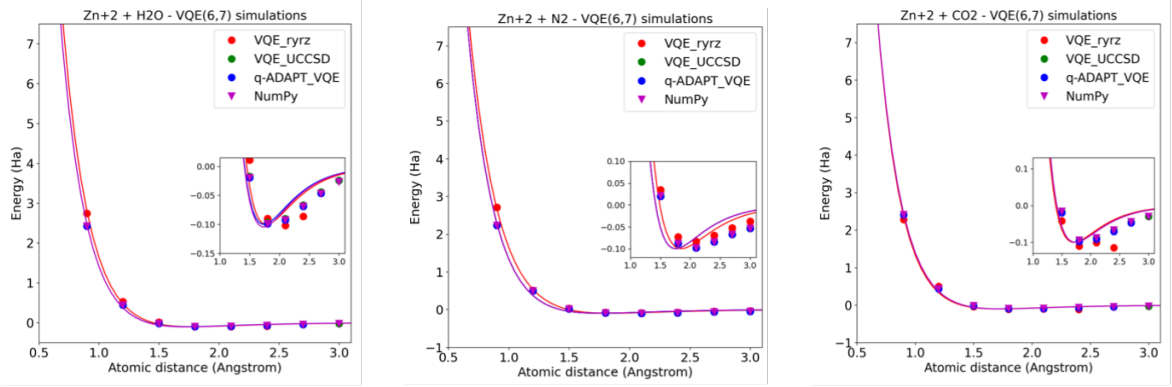
To assess the transferability of the qubit-ADAPT-VQE workflow demonstrated in the adapted article (Sec. 3.3) from Mg^{2+} (closed-shell $3s^0$) to a heavier, closed-shell d^{10} center, we repeated the study for Zn^{2+} adducts with H_2O, N_2, and CO_2. Unless otherwise noted, we adhered to the *exact same protocol* as in the reprint: identical one- and two-electron integral settings and basis/ECP choices; a (6,7) active space; Jordan–Wigner mapping with symmetry tapering; three ansätze (VQE_RyRz, VQE_UCCSD, and qubit-ADAPT-VQE); the same optimizer and convergence thresholds; geometry scans over the metal–ligand separation; Morse fits to extract (r_e, D_e); and uncertainty estimates from the fit residuals. NumPy direct diagonalization on the same active-space Hamiltonians served as a reference, and PySCF HF/DFT curves were computed for classical context.

(i) VQE_UCCSD and qubit-ADAPT-VQE track the NumPy benchmark across all Zn^{2+} adducts within the fitting uncertainties; (ii) VQE_RyRz behaves well for H_2O and CO_2 but deviates for N_2; (iii) compared to PySCF benchmarks, equilibrium distances from qubit-ADAPT-VQE agree closely with HF and are 2–4% longer than DFT; (iv) fitted binding energies cluster near -0.10 Ha in this single-ion model, substantially larger (by $\sim \times 10$) than experimental heats of adsorption in MOF-74, reflecting the oversimplified, fragment-only representation.

Figure 17 shows the Zn^{2+} PES obtained with the three ansätze alongside the NumPy reference. VQE_UCCSD and qubit-ADAPT-VQE maintain excellent overlap with NumPy for all three ligands. The hardware-efficient VQE_RyRz ansatz performs comparably on H_2O and CO_2 but diverges for N_2, indicating an expressivity/optimization limitation for this system.

Table 6 summarizes Morse-fit observables for Zn^{2+}+H_2O as a representative case: VQE_UCCSD and qubit-ADAPT-VQE agree with NumPy within the quoted uncertainties; VQE_RyRz shows higher uncertainty, failing to reach the benchmarks.

For completeness, Fig. 18 compares qubit-ADAPT-VQE PES (blue) with PySCF HF (red) and DFT (green). The qubit-ADAPT-VQE curves largely overlap with HF



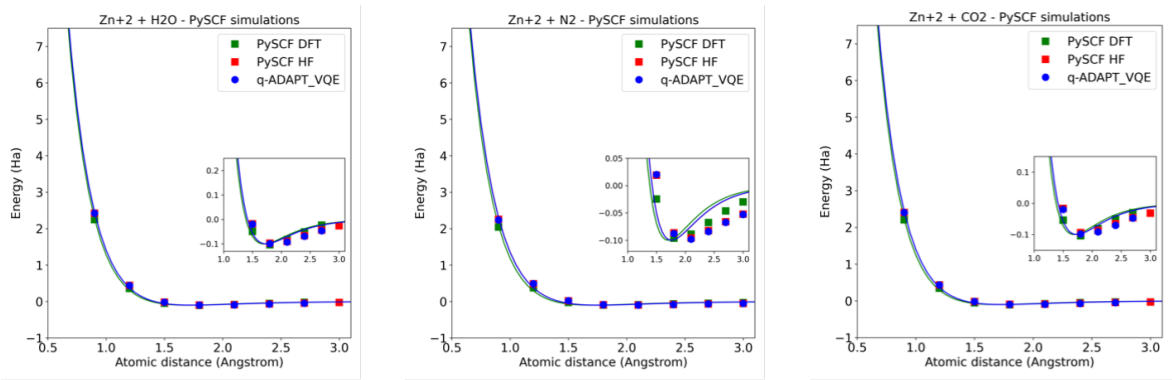
(a) $Zn^{2+} + H_2O$ (active space (6,7)). (b) $Zn^{2+} + N_2$ (active space (6,7)). (c) $Zn^{2+} + CO_2$ (active space (6,7)).

Figure 17 – **VQE PES for Zn^{2+} adducts: VQE_UCCSD (green), qubit-ADAPT-VQE (blue), and NumPy exact diagonalization (magenta) agree across all ligands**; VQE_RyRz (red) departs from the benchmark for N_2.

Table 6 – **Morse-fit equilibrium distance r_e and binding energy D_e for $Zn^{2+} + H_2O$ from the three VQE ansätze versus NumPy.** Uncertainties reflect fit residuals.

	VQE_RyRz	VQE_UCCSD	qubit-ADAPT-VQE	NumPy
r_e (Å)	1.80 ± 0.10	1.74 ± 0.05	1.74 ± 0.04	1.73 ± 0.05
D_e (Ha)	-0.10 ± 0.10	-0.10 ± 0.05	-0.10 ± 0.05	-0.10 ± 0.06

across all three ligands, while DFT yields slightly shorter r_e, most visibly for N_2.



(a) $Zn^{2+} + H_2O$. (b) $Zn^{2+} + N_2$. (c) $Zn^{2+} + CO_2$.

Figure 18 – **qubit-ADAPT-VQE (blue) compared to PySCF HF (red) and DFT (green) for Zn^{2+} adducts.** HF and qubit-ADAPT-VQE agree closely; DFT predicts slightly shorter r_e, especially for N_2.

Table 7 collects the (r_e, D_e) values from the Morse fits. While D_e agrees within numerical uncertainties across methods, r_e from qubit-ADAPT-VQE/HF is consistently 2–4% longer than DFT.

Two observations mirror the Mg^{2+} study and are instructive for the chapters ahead. First, ansatz *expressivity versus robustness*: chemically motivated ansätze (UCCSD,

Table 7 – **Morse-fit equilibrium distances r_e and binding energies D_e for Zn^{2+} adducts from qubit-ADAPT-VQE and PySCF HF/DFT.**

$Zn^{2+} + H_2O$			
	qubit-ADAPT-VQE	PySCF HF	PySCF DFT
r_e (Å)	1.74 ± 0.04	1.73 ± 0.05	1.70 ± 0.04
D_e (Ha)	-0.10 ± 0.05	-0.10 ± 0.02	-0.10 ± 0.05
$Zn^{2+} + N_2$			
	qubit-ADAPT-VQE	PySCF HF	PySCF DFT
r_e (Å)	1.76 ± 0.09	1.76 ± 0.08	1.70 ± 0.05
D_e (Ha)	-0.10 ± 0.10	-0.10 ± 0.09	-0.10 ± 0.06
$Zn^{2+} + CO_2$			
	qubit-ADAPT-VQE	PySCF HF	PySCF DFT
r_e (Å)	1.73 ± 0.06	1.73 ± 0.02	1.68 ± 0.04
D_e (Ha)	-0.10 ± 0.07	-0.10 ± 0.03	-0.10 ± 0.05

qubit-ADAPT-VQE) remain reliable across ligands, whereas a shallow hardware-efficient circuit (RyRz) may fail on systems like Zn^{2+}+N_2. Second, *model fidelity*: all fragment calculations yield $D_e \approx -0.10$ Ha, far larger than experimental heats of adsorption in Zn-/Mg-MOF-74 (27–31 kJ/mol ≈ 10–12 mHa [597]). This overbinding is expected from a single-ion cluster lacking framework polarization and many-body screening; it motivates the embedded/periodic Hamiltonians and sampling-based solvers developed in the subsequent chapters.

Note on formatting. The Zn^{2+} results were generated with the same code and settings as the reprinted study; they were omitted from the article due to space constraints and scope, and are included here to document transferability of the methodology.

Note on reproducibility. Code and releavant data is available on <https://github.com/MarcoBarroca/q4c-team>

3.5 Summary and motivation for sampling-based methods

This chapter established the ingredients needed for quantum simulations of fermionic systems. On the classical side, we reviewed the Born–Oppenheimer separation, the second-quantized electronic Hamiltonian of Eq. (3.1), and the Hartree–Fock framework leading to the Roothaan–Hall equations and practical SCF acceleration (overlap/Fock matrices, density matrix updates, and DIIS). We discussed how basis choices (Gaussian families, diffuse/polarization functions), frozen-core/ECP approximations, and integral factorization shape accuracy and cost, and how these modeling choices directly determine the size and sparsity of the Hamiltonian that downstream quantum methods must handle. We also summarized post-HF references used for benchmarking (e.g., CCSD, FCI) and highlighted where DFT fits in conceptually and practically.

On the quantum side, we outlined VQE (Algorithm 5), mapped fermions to qubits (Jordan–Wigner, parity, and Bravyi–Kitaev), explained how to exploit $\mathbb{Z}_2$ and point-group symmetries to taper qubits and shrink operator lists before measurement. We then discussed the role of ansätze ranging from hardware-efficient circuits to chemically motivated UCCSD and adaptive variants (ADAPT / qubit-ADAPT). The accompanying reprinted article and the Zn^{2+} follow-on results illustrated that, in modest active spaces with symmetry reductions, variational approaches can reproduce reference trends. At the same time, we saw the practical bottlenecks that emerge in larger or harder instances: measurement cost dominated by many Pauli terms, optimizer fragility, and circuit depth constraints that limit how far one can push expressivity on near-term devices.

These observations motivate a pivot to sampling-based eigensolvers that avoid global, high-dimensional, nonconvex optimization. Instead of training a deep parametric circuit, one prepares a *small set* of physically informed states using short circuits, estimates the reduced Hamiltonian and overlap matrices from measurements, and solves a generalized eigenvalue problem classically. This strategy turns depth into shots: depth stays low and robust, while accuracy is controlled statistically through sampling. Symmetry constraints, measurement grouping, and reuse of the same measurement primitives across matrix elements further improve efficiency.

The next chapter develops these sampling-based methods in detail. We introduce the subspace construction strategies and samplers we use, derive the resulting generalized eigenvalue problems, analyze statistical and systematic errors, and give resource estimates that explicitly reference measurement budgets. We then apply these techniques to new use cases in battery reactions and to embedded/periodic Hamiltonians, demonstrating how short-depth sampling can deliver systematic accuracy improvements without the optimization pathologies of large variational ansätze and allowing for the study of much larger systems.

4 Sample-based diagonalization: concepts, methods, and applications

4.1 Motivation and problem setting

Accurately computing ground-state and low-lying excited-state energies of quantum many-body systems is a central goal in quantum chemistry and physics. Traditional quantum algorithms like QPE offer rigorous convergence guarantees but require deep, coherent circuits beyond the reach of today's noisy hardware [46, 48]. The more NISQ-friendly alternative, VQE, has been widely explored, but its global parameter optimization lacks guaranteed convergence and becomes impractical for larger systems due to barren plateaus (exponentially vanishing gradients) and the prohibitive number of measurements needed for its cost function [36, 37, 46]. In fact, the parametric optimization bottleneck has so far prevented VQE from scaling beyond small molecule demonstrations [598]. Clearly, new approaches are needed that bypass long coherent evolutions and costly optimization routines, while still leveraging quantum devices for solving eigenvalue problems on classically intractable systems.

Sample-based (measurement-driven) diagonalization has emerged as a promising third route to tackle this challenge [42, 48]. Rather than encoding the solution in a deep quantum circuit or minimizing a global cost function, these methods use quantum hardware as a sampling engine to construct an approximate eigensubspace of the Hamiltonian, which is then solved via classical linear algebra [42, 45, 599]. In essence, the quantum processor prepares and measures a set of relevant many-body basis states (e.g. Slater determinants or short-time evolved states), and the Hamiltonian is projected into the span of those states and diagonalized classically [600]. This yields estimates of ground or excited eigenvalues and eigenstates within that effective subspace. Because the heavy diagonalization step is done classically, the quantum circuit depth can be kept shallow – often just enough to prepare some initial correlated states or implement short Trotterized evolutions – making the approach well-suited for pre-fault-tolerant devices [45, 48].

This idea was pioneered in Quantum-Selected Configuration Interaction (QSCI) where the quantum device is used to stochastically sample important determinants (configurations) from a reference wavefunction, which are then used to build a CI Hamiltonian matrix that is solved for the lowest eigenvalues [601]. Its extension, Sample-based Quantum Diagonalization (SQD), generalizes this idea and has demonstrated ground-state electronic structure calculations on systems far beyond classical exact diagonalization by combining quantum sampling with high-performance classical computing. Notably, Robledo-Moreno et al. used SQD on a quantum–classical workflow (IBM's Heron processor coupled with the Fugaku supercomputer) to simulate molecular systems up to 77 qubits and ∼10,000 gates, producing variational ground-state energy upper bounds and sparse wavefunction representations for N_2 and iron-sulfur clusters [42]. These were the largest electronic structure simulations at the time on quantum hardware, opening the door to chemistry applications close to the hundred-qubit scale.

In the SQD procedure, one prepares a shallow parametrized circuit or trial state on the quantum device (for instance, a Hartree–Fock state or a short-depth entangler) and then performs repeated measurements in the computational basis to sample the output wavefunction. The collected bitstrings (configurations) with highest frequencies are taken as a basis for a low-dimensional subspace [42]. The Hamiltonian matrix elements $H_{ij} = \langle\phi_i|H|\phi_j\rangle$ and overlap elements $S_{ij} = \langle\phi_i|\phi_j\rangle$ on this basis are obtained from classical calculation when possible, and then the generalized eigenvalue problem $Hc = ESc$ is solved on a classical computer. The lowest eigenvalue E provides an approximation to the ground-state energy, variational by construction (i.e., an upper bound to the true ground energy, absent noise) because the subspace is a subset of the full Hilbert space. Importantly, this approach bypasses the need for a variational optimization loop, instead leveraging the natural importance sampling of the quantum state to guide the subspace choice. As a result, sample-based diagonalization can be more noise-robust and measurement-efficient than VQE: one does not repeatedly reprepare states to compute gradients, and the measurement data directly contributes to building the Hamiltonian of a small effective problem rather than estimating a global expectation value. Recent analyses indeed report that SQD exhibits significantly flatter error landscapes and lower shot requirements compared to energy minimization methods, making it attractive for near-term devices [45]

The choice of subspace is flexible, and different flavors of sample-based diagonalization tailor this choice to balance accuracy and circuit cost. In the original SQD (and QSCI) framework, the subspace consists of important Slater determinants selected by measuring a reference state's amplitudes [42, 601]. This effectively performs a measurement-driven selected CI, where the algorithm "learns" the dominant configurations of the ground state. Another class of approaches builds the subspace through quantum Krylov vectors, i.e. by applying powers of the Hamiltonian or short-time evolution to an initial state. This leads to *Sample-Based Krylov Quantum Diagonalization* (SKQD) [46, 48]. Here one uses shallow Trotter or randomized product formula circuits to approximate states like $|\psi\rangle, H|\psi\rangle, H^2|\psi\rangle, \ldots$ on hardware, measures each to collect basis samples, and then diagonalizes H in the span of these Krylov basis states. Krylov subspaces are known to rapidly converge to the extreme eigenvalues. The SKQD variant can be shown to converge in polynomial time under assumptions similar to those of QPE if the ground state has a low effective dimension (is supported on a small subset of the basis). Yu et al. demonstrated this approach by combining quantum sampling with a Lanczos/Krylov strategy, achieving accurate ground-state energies for challenging lattice models and even executing an 85-qubit simulation of a single-impurity Anderson model (with 41 bath sites) on hardware [48]. In the presence of device noise, their sample-based Krylov algorithm outperformed purely quantum Krylov methods that lack measurement post-processing, underscoring the value of the classical diagonalization step in mitigating noise. Ongoing work is further improving this paradigm: for example, Piccinelli et al. introduced a randomized Trotter technique (dubbed "SqDRIFT") to generate Krylov basis states with lower circuit depth, thereby enabling SKQD calculations on molecules beyond the reach of exact diagonalization while still preserving rigorous convergence guarantees [599].

A key aspect of sample-based diagonalization methods is that they offer systematic accuracy control absent in black-box VQE. By increasing the number of sampled basis states (or the Krylov order, or number of measurements), one can systematically improve

the approximation and approach the exact eigenvalues from above. In practice, the subspace size is expanded until the desired accuracy (e.g. chemical accuracy of 1 kcal/mol $\approx$ 1.6 mHa) is reached. Because the classical solve scales polynomially with subspace dimension (which is kept much smaller than the full Hilbert space), this approach remains tractable even as more basis states are included, so long as the subspace expansion is efficient at capturing the wavefunction. In several benchmark studies, the ground-state energy error dropped rapidly (often exponentially) with each additional sampled basis state, indicating that relatively compact subspaces can capture the essential correlation for chemical systems [42, 46]. Moreover, the variational principle provides a built-in indicator of accuracy: if the estimated energy is still too high compared to known bounds or a higher-level method, the subspace can be enlarged further. This ability to converge from above contrasts favorably with VQE, where there is no a priori guarantee that increasing ansatz complexity or circuit depth will variationally improve the result (especially on noisy devices where the VQE energy may drift unpredictably due to noise-induced non-variational errors).

Equally important in reaching high accuracy are the variance-reduction and error-mitigation techniques employed alongside sampling. Since these algorithms rely on repeated measurements, statistical noise and hardware errors can contaminate the subspace construction if left unchecked. A variety of strategies have been developed to counteract this. For instance, one can enforce known symmetries on each measured sample to discard unphysical results: particle-number, spin, or point-group symmetries can be used as selection rules so that only bitstrings corresponding to the correct symmetry sector are admitted into the subspace [600]. This markedly reduces the impact of state-preparation errors and decoherence events which often violate conserved quantities. Robledo-Moreno et al. applied symmetry verification to filter sampled determinants and reported that it substantially improved the quality of the SQD wavefunction by removing spurious high-energy components [42]. Another technique is to perform a classical post-selection or quasi-purification on the sampled subspace: for example, the S-CORE method used in recent SQD work optimizes linear combinations of sampled configurations to restore correct electron counts and spin quantum numbers, effectively denoising the quantum outputs [45].

On the measurement side, Pauli Twirling and other error mitigation techniques can be applied to the short circuits used in subspace generation, further reducing measurement errors ensuring accurate sampling [42, 48]. Thanks to these measures, sample-based approaches have been able to achieve near chemical accuracy in many cases of interest. For example, a recent study incorporating an implicit solvent model into SQD achieved energy estimates for solvated molecules (water, alcohols, etc.) that matched high-level classical benchmarks within 1–2 mHa, i.e. within chemical accuracy [45]. These results underscore that with careful error mitigation, even today's noisy devices can deliver quantitatively reliable energies for non-trivial systems with the help of sampling=based algorithms.

Measurement-driven subspace algorithms like QSCI, SQD and SKQD offer a compelling alternative to VQE for finding eigenvalues on NISQ hardware. They leverage shallow circuits and repeated measurements to directly extract spectral information, rather than contending with deep circuits or difficult global optimizations. This approach directly addresses the core pain points of variational methods – barren plateaus and enormous measurement overhead – by trading quantum depth for classical post-processing. The

quantum computer is used only for what it does best (preparing and sampling modestly entangled states), while classical computation handles the exact diagonalization in a reduced space. This hybrid strategy has already enabled the largest electronic structure simulations on quantum processors to date [6, 42, 43, 48], and offers a clear path to systematically improve accuracy by expanding subspaces and incorporating advanced error mitigation. In the following sections, we detail the core algorithms and variants in this family – from the basic SQD procedure to extensions for excited states and Krylov-based improvements – and discuss their applications to quantum chemistry problems. We also examine how these methods can be scaled up through embedding techniques and parallelization, and we provide practical guidelines for their implementation on real hardware. These sample-oriented diagonalization methods not only evade the scalability challenges of global VQE optimizations, but also retain a firm handle on systematic errors, positioning them as a crucial bridge toward the error-corrected era of quantum simulation.

4.2 Core algorithms

In this section, we describe in depth the current core hybrid quantum–classical sampling based algorithms. These include Quantum Selected Configuration Interaction (QSCI) and the Sampled Quantum Diagonalization (SQD) family of methods, along with two advanced extensions: Extended SQD (Ext-SQD) for accessing ground and excited states, and Sample-Based Krylov Quantum Diagonalization (SKQD). For each algorithm, we provide a theoretical explanation with relevant equations and a pseudo-code style outline of the procedure. We also include a brief refresher on Krylov subspace methods where appropriate, and conclude with a summary comparing these approaches.

4.2.1 Quantum Selected Configuration Interaction (QSCI)

QSCI is a hybrid quantum–classical algorithm that uses a quantum device to select an important subset of basis configurations (determinants) for a configuration interaction (CI) calculation, which is then solved on a classical computer. The goal is to leverage a quantum state as an oracle for where the weight of the wavefunction lies, thereby truncating the CI basis to those configurations that are most important for describing the eigenstates of a many-electron Hamiltonian. This yields a variational approximation to ground- or excited-state energies that is noise-resilient and does not require expensive variational circuit optimization [601].

QSCI begins by preparing an input state on the quantum processor that has a non-negligible overlap with the true ground state of the target Hamiltonian H. In practice, such a trial $|\Psi_{in}\rangle$ can be obtained from a short VQE pre-optimization or any other heuristic state-preparation routine. We assume it admits an expansion in the computational (Slater-determinant) basis of a chosen spin–orbital set,

$$|\Psi_{in}\rangle \;\approx\; \sum_i c_i \, |D_i\rangle,$$

where each $|D_i\rangle$ is a basis determinant.

Next, we sample the state in the computational Z basis. Repeated measurements (many shots) of $|\Psi_{\text{in}}\rangle$ yield bitstrings corresponding to specific determinants $|D_i\rangle$. From these outcomes we estimate empirical probabilities $P(D_i) = n(D_i)/N_{shots}$. While drawing from this (potentially exponentially supported) distribution is classically non-trivial, the quantum device samples directly from the Born probabilities $P(D_i) = |c_i|^2$ up to finite-shot noise.

We then select important configurations by identifying the most probable determinants in the measurement record. A convenient choice is top-K selection, i.e. retain the K most frequent outcomes, to define a selected subspace

$$\Omega_K \;=\; \{\, |D_{i_1}\rangle,\; |D_{i_2}\rangle,\; \ldots,\; |D_{i_K}\rangle \,\}.$$

(Alternately, a probability threshold may be used, or one can include all observed determinants; fixing K simplifies exposition.) The span of Ω_K is a truncated selected-CI subspace of the full Hilbert space.

With Ω_K fixed, we assemble the subspace Hamiltonian on a classical computer by evaluating all matrix elements

$$H_{ij} \;=\; \langle D_i|\, H\, |D_j\rangle, \qquad |D_i\rangle, |D_j\rangle \in \Omega_K,$$

using standard techniques and algorithms (e.g., Slater–Condon rules and Davidson's algorithm in quantum chemistry), and build the $K \times K$ matrix H_Ω.

We then *diagonalize in the selected subspace* by solving the CI eigenproblem

$$H_\Omega\, \vec{c} \;=\; E\, \vec{c}.$$

The lowest eigenpair $(E_0, \vec{c}_0)$ provides the QSCI ground-state estimate. The corresponding subspace eigenvector defines the approximate state explicitly,

$$|\Psi_{out}\rangle \;=\; \sum_{|D_j\rangle \in \Omega_K} c_{0,j}\, |D_j\rangle,$$

i.e., a linear combination of the selected configurations, known classically from $\vec{c}_0$.

Finally, we *report energy and (optionally) other observables.* The energy $E_0 = \langle \Psi_{out}|H|\Psi_{out}\rangle$ is a variational upper bound to the exact ground energy E_{exact} because it results from exact diagonalization within a subspace:

$$E_0 \;=\; \langle \Psi_{out}|H|\Psi_{out}\rangle \;\geq\; E_{exact}.$$

As K (or, more generally, the span of Ω_K) increases, E_0 decreases monotonically toward E_{exact}, furnishing a built-in convergence check: enlarging the subspace can only improve (or leave unchanged) the energy if the exact eigenstate is already contained. Moreover, since QSCI yields the explicit coefficient vector $\vec{c}_0$, any observable O expressible in the same determinant basis can be evaluated purely classically as $\langle \Psi_{out}|O|\Psi_{out}\rangle$ after constructing the corresponding subspace matrix elements. This completes the workflow summarized in Algorithm 7, without removing any of its steps or assumptions.

Algorithm 7 Quantum Selected Configuration Interaction (QSCI) for ground state

Require: Hamiltonian H; shots N_{shots}; subspace size K (or probability threshold τ).
Ensure: Approximate ground-state energy E_0 and coefficients $\{c_{0,j}\}$ over selected determinants Ω_K.

1: **Prepare trial state.** Prepare on the QPU an input state $|\Psi_{\text{in}}\rangle$ with non-negligible ground-state overlap, e.g. from a short VQE or heuristic ansatz. Write $|\Psi_{\text{in}}\rangle \approx \sum_i c_i |D_i\rangle$ in the computational (Slater-determinant) basis.
2: **Sample in the computational basis.** Measure $|\Psi_{\text{in}}\rangle$ in the Z basis N_{shots} times; tally counts $n(D_i)$ and form empirical probabilities $P(D_i) = n(D_i)/N_{\text{shots}}$.
3: **(Optional) Symmetry filtering / readout mitigation.** Discard bitstrings violating known constraints (e.g. fixed N, S_z); optionally apply readout-error mitigation to $P(D_i)$.
4: **Select important configurations.** Choose the most probable determinants by either:

- top-K selection: take the K largest $P(D_i)$, or
- thresholding: take all D_i with $P(D_i) \geq \tau$.

Denote the selected set by $\Omega_K = \{\, |D_{i_1}\rangle, \ldots, |D_{i_K}\rangle \,\}$.
5: **Build Hamiltonian submatrix.** For all $|D_i\rangle, |D_j\rangle \in \Omega_K$, compute

$$H_{ij} \;=\; \langle D_i|\, H\, |D_j\rangle$$

(e.g. via Slater–Condon rules); assemble the $K \times K$ matrix H_Ω.
6: **Diagonalize in the subspace.** Solve the eigenproblem

$$H_\Omega\, \vec{c} \;=\; E\, \vec{c},$$

take the lowest eigenpair $(E_0, \vec{c}_0)$, and form the subspace eigenstate

$$|\Psi_{\text{out}}\rangle \;=\; \sum_{|D_j\rangle \in \Omega_K} c_{0,j}\, |D_j\rangle.$$

7: **Report energy and observables.** Output $E_0 = \langle \Psi_{\text{out}}|H|\Psi_{\text{out}}\rangle$ (a variational upper bound to E_{exact}). For any operator O available in this basis, compute $\langle O\rangle = \langle \Psi_{\text{out}}|O|\Psi_{\text{out}}\rangle$ classically using $\vec{c}_0$ and the matrix elements of O.
8: **Convergence monitor (optional).** Increase K (or lower τ) and repeat steps 2–7 until the energy change $|E_0^{(t)} - E_0^{(t-1)}|$ is below a chosen tolerance.

In essence, QSCI is a form of Selected CI where the selection of determinants is guided by a quantum sample rather than by classical heuristics or perturbation theory. Historically, selected CI methods (like CIPSI, HCI, etc.) choose important configurations iteratively based on estimated energy gains or coefficient magnitudes, which can be computationally expensive but have seen renewed interest due to modern hardware. QSCI provides a systematic way to perform selected CI: it uses the quantum-prepared state as a probability oracle to directly pick out the Slater determinants with the largest coefficients in the wavefunction. Under ideal conditions (negligible device noise and infinite sampling), the selected set Ω_K will contain the determinants with the largest true CI coefficients in

the input state, which, if the input state is a decent approximation, are likely also among the most important determinants of the exact ground state [601].

Notably, QSCI only requires measuring in the computational basis, so no additional quantum gates are needed for measurements beyond state preparation. This is advantageous for NISQ devices: measuring in the Z-basis is straightforward and avoids the accumulated errors of performing, say, controlled rotations or extensive circuits for expectation estimation. Moreover, because all Hamiltonian matrix elements in the selected subspace are computed exactly on the classical side, the final energy is not polluted by quantum measurement noise – any error comes only from having an insufficient or slightly wrong subspace. The variational principle $E_0 \geq E_{exact}$ holds even in the presence of sampling noise or device noise that may cause a suboptimal subspace selection. In contrast, algorithms like VQE do not guarantee an upper bound when noise is present (the measured energy can variationally "leak" below the true minimum due to errors).

The above QSCI procedure was described for the ground state, but it can be extended to excited states as well. The simplest approach is to reuse the same selected subspace Ω_K and take not just the lowest eigenpair but also higher eigenvalues/eigenvectors of H_Ω as approximations to excited states. In other words, if H_Ω is diagonalized, one obtains a set of eigenvalues $E_0 \leq E_1 \leq \ldots$ and eigenvectors $\vec{c}^{(n)}$; these can serve as approximations for the ground state and excited states (first excited, second excited, etc.) within the chosen subspace. However, this approach works well only if the subspace Ω_K is sufficiently rich to represent those excited states. A subspace selected solely from the ground-state wavefunction might miss configurations important for excited states, leading to poor excited-state accuracy [601].

To address this, two main strategies have been proposed by Kanno et al.:

- Single subspace (multi-state) QSCI: Use multiple input states (each targeting a different eigenstate) to build a combined subspace, then diagonalize once. For example, one could prepare an approximate ground state $|\Psi_{in}^{(0)}>$ and an approximate first-excited state $|\Psi_{in}^{(1)}>$ (e.g. by orthogonalizing a guess to the ground state), sample both, and take the union of important configurations from both samples as Ω_{multi}. Diagonalizing H in this union subspace yields approximations for both states simultaneously. This "single diagonalization scheme" is analogous to a quantum subspace expansion for excited states, generalizing the idea of including singly- or doubly-excited determinants but here including those suggested by the quantum states. It maintains the variational upper-bound property for each targeted eigenstate.

- Sequential QSCI (state-by-state): Apply the ground-state QSCI algorithm to get $|\Psi_{out}^{(0)}>$ and E_0. Then for the first excited state, prepare a new input that is orthogonal to the obtained ground state (to avoid rediscovering the ground state). For instance, one can modify the quantum circuit or the Hamiltonian (as in variational quantum deflation, VQD) to penalize overlap with $|\Psi_{out}^{(0)}>$. Sample this new state to get a subspace for the first excited state and diagonalize to obtain E_1. Repeat for higher states as needed. This "sequential diagonalization scheme" is analogous to iterative classical algorithms that solve for one eigenstate at a time, ensuring orthogonality (like orthogonal CI or deflation methods). The QSCI advantage is that

overlap penalties can be implemented without extra quantum overlap measurements by working in the subspace directly (one can enforce orthogonality by projecting out the previously found state in the classical subspace).

QSCI leverages the strength of quantum sampling to adaptively choose a compact CI basis. It may be viewed as a "quantum-assisted selected CI" method, executing the CI diagonalization only once per state (as opposed to classical adaptive CI algorithms that iterate many times adding one determinant at a time). This one-shot diagonalization can significantly reduce classical cost relative to iterative approaches, provided the quantum device can provide a good sampling of the wavefunction. Furthermore, unlike other quantum subspace methods like the Quantum Subspace Expansion (QSE) which require measuring Hamiltonian matrix elements on quantum hardware, QSCI performs all Hamiltonian evaluations classically, sidestepping hardware noise in the most critical part of the computation. The trade-off is that QSCI restricts the subspace to computational-basis states (Slater determinants), rather than more general correlated basis states used in QSE. Even so, QSCI has been shown to achieve high accuracy with moderate subspace sizes on both simulated and real quantum devices. An experimental demonstration by Kanno et al. prepared an 8-qubit molecular Hamiltonian state and successfully applied QSCI, showing the method's viability on current noisy devices.

4.2.2 Sample-based Quantum Diagonalization (SQD)

The Sample-based Quantum Diagonalization (SQD) algorithm refers to essentially the same class of methods as QSCI, and the terms "QSCI" and "SQD" are often used interchangeably in recent literature [602]. Both involve using quantum samples from a state to define a subspace in which the Hamiltonian is diagonalized classically. In practice, however, SQD implementations incorporate a self-consistent configuration recovery step that repairs noise-corrupted bitstrings with the wrong particle number: one estimates the average spin-orbital occupancies from the observed data and then probabilistically flips bits to restore the target N_α and N_β, iterating this recovery and re-estimation to stabilize the pool of valid configurations used for projection and diagonalization. The recovered configurations are then sub-sampled into multiple batches to build projected Hamiltonians that are diagonalized classically, where the lowest energy value yields a variational upper bound [42]. Occupancies are then recalculated and the process iterates again. This recovery mechanism is the sample-based analogue of selection in classical selected-CI schemes such as HCI [285], and it is made explicit in Algorithm 8. In what follows we use "SQD" when referring broadly to sample-based subspace diagonalization (including the recovery step) across chemistry and lattice models, and reserve "QSCI" for the chemistry-specific instantiation that selects Slater determinants as configurations.

Another central attraction of SQD is that it does not require coherent time evolution. Unlike phase–estimation or real/imaginary–time algorithms, SQD avoids implementing e^{-iHt}, Trotter/LCU or qubitization blocks; the only quantum primitives are (i) preparing a shallow trial state $|\Psi_{\text{in}}\rangle$ with non-negligible ground-state overlap and (ii) measuring in the computational (Z) basis to obtain bitstrings. This is particularly appealing for electronic–structure Hamiltonians with thousands to millions of Pauli terms,

where accurate Trotterization is both deep and compilation–intensive; in SQD the circuit depth is essentially the ansatz depth used to prepare $|\Psi_{in}\rangle$ [42].

A second feature is scalability via classical HPC: SQD pushes the heavy linear–algebra workload onto a classical post-processing stage that assembles and diagonalizes the Hamiltonian restricted to a selected subspace inferred from the measurement record. Recent demonstrations highlight this quantum–centric split, using the quantum processor only to sample bitstrings from $|\Psi_{in}\rangle$ (on tens to $\mathcal{O}(100)$ qubits), while a large CPU backend performs subspace construction, denoising, and diagonalization; for example, IBM reported a 77-qubit molecular active-space experiment where SQD enabled classical diagonalization in a measurement-defined subspace far larger than what a direct classical FCI would allow [42]. This synergy extends the reach beyond exact classical diagonalization while keeping quantum circuits shallow.

Because hardware bitstrings are noisy, SQD workflows routinely incorporate error mitigation at the sample level. Common techniques include symmetry post-selection (discarding outcomes with the wrong particle number, spin parity, or point-group irrep), filtering out implausible low-probability outliers, and aggregating independent batches to stabilize the empirical distribution $P(D_i)$ before selecting the subspace Ω_K [42]. Such denoising especially suppresses single-qubit flips that spuriously change N or S_z, yielding cleaner subspaces and more stable eigenvalues without additional quantum depth.

Finally, SQD has shown competitive accuracy across small and intermediate benchmarks. When $|\Psi_{in}\rangle$ has adequate overlap, the SQD ground-state energies are variational within the selected subspace and have been reported within (or near) chemical accuracy relative to FCI/selected-CI references on small molecules, while larger active spaces track high-level classical baselines [42]. The reconstructed wavefunctions agree best in their dominant coefficients, with most discrepancies confined to the long tail of very small amplitudescomponents that contribute little to the total energy but may influence particularly sensitive observables; this informs practical choices of subspace size K and observable-specific validation [42].

The flow described in Algorithm 8 is exactly the same as QSCI, save for the configuration recovery step and indeed SQD and QSCI are the same algorithm in spirit. Recent works explicitly identify QSCI and SQD, referring to these methods collectively as sample-based quantum diagonalization [602]. For consistency, we will treat them under the unified framework of sample-based subspace diagonalization.

The SQD approach was inspired by the success of Quantum Subspace Expansion (QSE) and Quantum Krylov methods (discussed below) in early quantum algorithms. However, QSE typically requires evaluating matrix elements $< \phi_i|H|\phi_j >$ on quantum hardware for a set of non-orthogonal basis states $|\phi_i >$ generated by certain excitations on a reference state . Those matrix elements come with statistical and device noise. In contrast, SQD's philosophy is to only use the quantum device for sampling, and do all Hamiltonian matrix element calculations classically (which is possible since the basis is a set of computational basis states). This trades away some flexibility in choosing "fancier" basis states (which QSE might allow) in exchange for robustness and exactness in the subspace diagonalization step. The success of a sample-based method hinges on the sparsity of the target eigenstate in the computational basis – i.e., that relatively few basis states have significant weight. Fortunately, many chemical ground states are approximately

Algorithm 8 Sample-based Quantum Diagonalization (SQD)

Require: Hamiltonian H; target electron counts (N_α, N_β); shot budget N_{shots}; iterations T; batch size M; number of batches B

Ensure: Approximate ground-state energy $E_\star$ (upper bound) and coefficients $\vec{c}_\star$

1: **Prepare and sample once:** Prepare $|\Psi_{\text{ref}}\rangle$ on the QPU and measure in the computational (Z) basis N_{shots} times to obtain a multiset $\widetilde{\mathcal{X}} = \{(b, f(b))\}$.

2: **Init occupancies $\bar{n}^{(0)}$:** From the subset with correct (N_α, N_β), estimate per–spin-orbital averages $\bar{n}_p^{(0)}$ (fallback: HF if empty).

3: **for** $t = 1$ **to** T **do**

4: **Configuration recovery (using $\bar{n}^{(t-1)}$):**

- Split $\widetilde{\mathcal{X}}$ into correct-N and wrong-N.
- For each wrong-N bitstring b, *probabilistically flip bits* guided by $\bar{n}^{(t-1)}$ until (N_α, N_β) is restored, yielding $\tilde{b}$.
- Form the recovered pool $\widetilde{\mathcal{X}}_R^{(t)}$ by combining correct-N samples with the recovered $\tilde{b}$'s; renormalize weights to get $\tilde{f}^{(t)}$.

5: **Subsample batches:** Define a sampling distribution over $\widetilde{\mathcal{X}}_R^{(t)}$ (e.g., proportional to $\tilde{f}^{(t)}$). Draw B batches $\mathcal{S}^{(t,1)}, \ldots, \mathcal{S}^{(t,B)}$, each with M distinct Slater determinants.

6: **Solve batches (classical/HPC):** For each $k = 1, \ldots, B$:

- Build $H^{(t,k)}$ over $\mathcal{S}^{(t,k)} = \{|D_i\rangle\}_{i=1}^{M}$ with $H_{ij}^{(t,k)} = \langle D_i|H|D_j\rangle$.
- Diagonalize to obtain the lowest eigenpair $\left(E_{\min}^{(t,k)}, \vec{c}_{\min}^{(t,k)}\right)$.

7: **Pick lowest-energy batch:** $k^\star = \arg\min_k E_{\min}^{(t,k)}$; set $E_\star^{(t)} = E_{\min}^{(t,k^\star)}$ and $\vec{c}_\star^{(t)} = \vec{c}_{\min}^{(t,k^\star)}$.

8: **Update occupancies from lowest-energy batch only:**

$$\bar{n}_p^{(t)} \;=\; \sum_{i=1}^{M} \left|c_{\star,i}^{(t)}\right|^2 n_p(D_i),$$

where $n_p(D_i) \in \{0, 1\}$ is the occupation of orbital p in determinant D_i.

9: **end for**

10: **Return best across iterations:** $E_\star = \min_t E_\star^{(t)}$ with $\vec{c}_\star = \vec{c}_\star^{(t^\star)}$ from the iteration $t^\star$ attaining $E_\star$.

11: **Note:** Omitting the recovery step reduces the workflow to QSCI.

sparse in a chemically meaningful single-particle basis, they might be dominated by a few configurations like a reference and some excitations. This sparsity underlies the empirical success of SQD: as long as the ground-state wavefunction is not too delocalized over the Hilbert space, a moderate number of samples will identify the most important configurations with high probability. If the state were completely delocalized (no sparsity), then sample-based selection would require exponentially many samples to converge – a scenario where other methods would be needed.

4.2.3 Extended SQD (Ext-SQD) for Ground and Excited States

While basic SQD/QSCI usually focus on finding the ground-state energy, many chemical and physical properties of interest involve excited states (e.g. absorption spectra, reaction pathways involving excited surfaces, etc.). Thus, extending the sample-based diagonalization framework to obtain excited-state information is crucial. The Extended SQD (Ext-SQD) algorithm, proposed by Barison et al., is a method to systematically improve the description of excited states (and the ground state) by enlarging the subspace in a physically motivated way. Ext-SQD builds upon an initial SQD result and adds a second stage of basis expansion and diagonalization. It was demonstrated to yield significantly more accurate excited-state energies than standard SQD or a simple QSE approach, with relatively modest extra cost [43].

In SQD, if the trial state $|\Psi_{ref}\rangle$ is optimized for the ground state, the selected determinants will bias heavily toward ground-state configurations. Excited states often have qualitatively different configurations — for example, a singly-excited configuration where an electron is promoted to a higher orbital. A ground-state sampling might miss such determinants, meaning the excited states obtained by diagonalizing the ground-state subspace are poor. One approach to handle this is to include by construction some excited configurations into the subspace. Quantum Subspace Expansion does this by taking a reference state and manually adding all single (and maybe double) excitations of that reference to the basis. Ext-SQD takes a similar idea but in a more adaptive, multi-reference manner: after obtaining an approximate ground state via SQD, it generates excitations from all important configurations found in the SQD wavefunction, thereby enriching the subspace with basis states that can represent excited-states or even other relevant pieces of the ground state.

Ext-SQD augments a basic sampled diagonalization by (i) first obtaining a compact, data-driven ground-state subspace from measurements and (ii) then enriching that subspace with physically motivated excitations to capture both additional correlation and low-lying excited states. In the first step, one runs a standard SQD on the quantum device to prepare and sample a trial state and then classically diagonalizes the Hamiltonian projected onto the selected configurations. This yields an approximate ground-state eigenvector $\left|\Psi_{\text{out}}^{(\text{SQD})}\right\rangle = \sum_{i\in\Omega_{\text{SQD}}} c_i \left|D_i\right\rangle$ in a small subspace Ω_{SQD} (Alg. 9, line 1).

Before extending the basis, one may prune numerically insignificant components by discarding determinants whose amplitudes fall below a threshold τ; the resulting reduced set Ω'_{SQD} retains the dominant configurations while lowering memory footprint and improving numerical conditioning (Alg. 9, line 2). The second step then generates excited configurations: for each determinant that survives pruning, apply a chosen pool of excitation operators (typically all single and double substitutions in a quantum-chemistry context) to produce candidate determinants. Each candidate is screened for physical constraints (fixed particle number, spin projection, and point-group symmetry), and duplicates, whether already present in Ω_{SQD} or created multiple times by different parents, are removed (Alg. 9, line 3). In practice, these excitations are produced in batches to balance memory and parallelism: processing B seed determinants at a time enables efficient parallel expansion and deduplication without materializing the full candidate set at once (Alg. 9, batching note).

The extended basis is then formed as a union $\Omega_{\text{ext}} = \Omega_{\text{SQD}} \cup \{\text{excited determinants}\}$ (Alg. 9, line 6). Although each retained determinant can, in principle, spawn $\mathcal{O}(n^2)$ singles and $\mathcal{O}(n^4)$ doubles for n spin-orbitals, extensive deduplication across parents and symmetry filtering can keep $|\Omega_{\text{ext}}|$ polynomial and typically much smaller than the naive product $K' \times E$. In some cases, after pruning tiny components, the net extended set can even be *smaller* than the original if the discarded tails would have generated largely redundant or irrelevant excitations.

Finally, Ext-SQD diagonalizes the Hamiltonian projected into Ω_{ext} to obtain an improved ground-state energy E_0^{ext} and a set of low-lying excited energies $E_1^{\text{ext}}, E_2^{\text{ext}}, \ldots$ together with their eigenvectors (Alg. 9, line 7). The enriched basis systematically captures more correlation for the ground state and, crucially, includes the configurations necessary to describe target excitations with high fidelity, leading to smoother potential-energy curves and more accurate symmetry diagnostics (e.g., spin expectation values) than a multi-shot SQD subspace. The full procedure preserves the attractive workflow of sample-based methods (only the initial subspace discovery uses quantum measurements; all subsequent pruning, excitation generation, and subspace diagonalization are classical) while providing a principled path to both ground and excited states with controllable accuracy.

Algorithm 9 Extended SQD (Ext-SQD): two-step approach for ground and excited states

1: **Initial SQD (ground state).** Run the standard SQD procedure to obtain an approximate ground eigenvector in a selected subspace Ω_{SQD} of size K:

$$\left|\Psi_{\mathrm{out}}^{(\mathrm{SQD})}\right\rangle \;=\; \sum_{i\in\Omega_{\mathrm{SQD}}} c_i\,|D_i\rangle\,,$$

with coefficients $\{c_i\}$ returned by the first diagonalization.

2: **Prune small components (optional).** Remove configurations with amplitudes below a threshold τ: set

$$\Omega'_{\mathrm{SQD}} \;=\; \{\,|D_i\rangle \in \Omega_{\mathrm{SQD}} \;:\; |c_i| \geq \tau\,\}\,,$$

yielding a reduced set of size $K' \leq K$. Pruning reduces cost and can improve numerical stability without appreciable loss of accuracy.

3: **Generate excited configurations.** For each determinant retained, apply a predefined excitation set (e.g., all single and double excitations in quantum chemistry):

4: **for all** $|D_i\rangle \in \Omega'_{\mathrm{SQD}}$ **do**

5: **for all** excitation operators $\hat{E}$ in the chosen pool (S, D, and optionally higher) **do**

6: $|D'\rangle \leftarrow \hat{E}\,|D_i\rangle$

7: **if** $|D'\rangle$ is **physical** (respects particle number, spin, point-group symmetry, etc.) **then**

8: **if** $|D'\rangle \notin \Omega_{\mathrm{SQD}}$ **and** $|D'\rangle \notin \Omega_{\mathrm{ext}}$ **then**

9: add $|D'\rangle$ to Ω_{ext} {new configuration}

10: **else**

11: **skip** duplicate (already present) {avoid redundancy}

12: **end if**

13: **else**

14: **discard** unphysical outcome

15: **end if**

16: **end for**

17: **end for**

18: **Combine subspaces.** Form the extended basis by union:

$$\Omega_{\mathrm{ext}} \;\leftarrow\; \Omega_{\mathrm{SQD}} \;\cup\; \Omega_{\mathrm{ext}}.$$

19: **Diagonalize in the extended subspace.** Build the projected Hamiltonian $H_{\Omega_{\mathrm{ext}}}$ and solve

$$H_{\Omega_{\mathrm{ext}}}\,\vec{c}^{\,(m)} \;=\; E_m^{\mathrm{ext}}\,\vec{c}^{\,(m)}$$

for the lowest few eigenpairs.

20: **Output and analysis.** Report $\{E_m^{\mathrm{ext}}, \vec{c}^{\,(m)}\}$ for $m = 0, 1, \ldots$.

In simpler terms: Ext-SQD is a two-step CI approach. First, do a coarse but inexpensive CI (via SQD) to get a starting wavefunction. Second, use that wavefunction as a guide to include a second layer of configurations (those reachable by single/double excitations from its dominant components), then do a refined CI in the larger space. This is conceptually similar to performing a multi-reference configuration interaction

(MRCI) where the "reference space" is given by the SQD wavefunction's supports, and all excitations out of that reference up to a certain level are included. The benefit is that one can capture not just dynamic correlation for the ground state (improving its energy) but also the correct manifold of excited configurations for low-lying excited states. Indeed, Barison et al. report that Ext-SQD improved ground-state energies by an order of magnitude compared to SQD alone (bringing them much closer to benchmark methods like Heat-Bath CI) [43], and at the same time delivered accurate excitation energies.

Comparing to the alternative (Quantum Subspace Expansion with singles and doubles, QSE(SD)): Ext-SQD was found to be far more efficient [43]. In an example, QSE(SD) for N_2 required diagonalizing a 267×267 generalized eigenvalue problem (all singles and doubles off a reference), which took 5 hours on a 120-core classical node, whereas Ext-SQD (after SQD selection) involved subspaces of size perhaps a few hundred but took only 20 seconds on a single core [43]. The stark difference is because QSE(SD) must evaluate many Hamiltonian and overlap integrals for all excitation pairs – a costly step – and solve a potentially ill-conditioned generalized eigenproblem (overlap matrix issues at large bond lengths necessitating regularization). Ext-SQD, by targeting a tailored set of excitations (often fewer than the full set of singles/doubles) and working in an orthonormal determinant basis, avoids the overlap matrix complications and reduces the computational burden dramaticallyȦs a result, in the strong correlation regime (e.g. stretched N_2), QSE(SD) can even produce worse energies than basic SQD due to numerical issues, whereas Ext-SQD remains stable and accurate [43].

Ext-SQD demonstrates how one can iteratively refine a sample-selected subspace to include more physics for excited states. It inherits the appealing features of SQD (using the quantum device only for state preparation and sampling) and augments the classical post-processing to handle excited-state correlation. The cost trade-off is an additional diagonalization step (and the generation of new determinants), but this can be polynomial and very manageable. Ext-SQD thus serves as a blueprint for obtaining spectral information or simply improve ground state calculations from a NISQ-centric workflow and has been successfully applied not just to small molecules but also to challenging systems like periodic materials and battery reactions [6, 7].

4.2.4 Sample-Based Krylov Quantum Diagonalization (SKQD)

The Sample-Based Krylov Quantum Diagonalization (SKQD) algorithm is a recent advance that unifies the SQD approach with Krylov subspace methods [46, 48]. Krylov subspace methods are a cornerstone of classical eigenproblem algorithms (e.g. the Lanczos algorithm [120]) and also form the basis of some quantum algorithms for ground states. Here we provide a brief refresher on Krylov methods, then explain how SKQD merges that idea with sampling to yield a hybrid algorithm with improved convergence guarantees.

Given a Hermitian operator H and a reference state $|\psi_0>$, the Krylov subspace of order m is defined as: $\mathcal{K}_m(H, |\psi_0\rangle) = \text{span}\{\, |\psi_0\rangle, H|\psi_0\rangle, H^2|\psi_0\rangle, \ldots, H^{m-1}|\psi_0\rangle \,\}$. Intuitively, this is the subspace one gets by repeatedly applying the Hamiltonian to the initial vector (up to $m-1$ times). The hope is that for sufficiently large m, $\mathcal{K}_m$ will contain a good approximation to the true ground state (or a few lowest eigenstates), because powers of H acting on any state will amplify the ground-state component (if $|\psi_0>$ has

overlap with the ground state). The Lanczos algorithm is a classical iterative method that builds an orthonormal basis of $\mathcal{K}_m$ and tridiagonalizes H on it, yielding approximate eigenvalues that converge (often rapidly) to extremal eigenvalues like the lowest energy.

A variant in quantum algorithms is Krylov quantum diagonalization (KQD). Instead of $H^k|\psi_0>$ (which is not straightforward to get on quantum hardware), one uses time-evolved states as a basis: $|\psi_k = e^{-iHk\Delta t}|\psi_0 \quad (k = 0, 1, \ldots, m-1)$. If Δt is small, these states can be seen as a discretized approximation to the power series (since $e^{-iH\Delta t} \approx I - i\Delta t, H + \cdots$). In practice, KQD prepares states at different times and then uses measurements to estimate the Hamiltonian matrix in this time-evolved basis and solves a generalized eigenvalue problem $\mathbf{Hc} = E, \mathbf{Sc}$ (where $\mathbf{S}$ is the overlap matrix of the $|\psi_k>$). KQD has a key theoretical appeal: if $|\psi_0>$ has a reasonable overlap with the ground state, then as m increases, the Krylov subspace is guaranteed to converge to the ground state (in exact arithmetic), much like phase estimation but using shorter circuits. Even with noise, it has rigorous error bounds and has been experimentally demonstrated (e.g. Up to 56-spin systems were diagonlaized with KQD [46]). The downside of KQD is that one must perform fairly complex measurements to get the matrix elements of H and S between non-orthogonal states $|\psi_k>$ and $|\psi_j>$, which can be costly in terms of circuit depth and sample count.

SKQD marries the convergence intuition of Krylov subspaces with the hardware–friendly, sample–selected subspaces of SQD. Introduced by Yu *et al.* [48], the method chooses a time step Δt and an order m, then prepares on quantum hardware the ladder of short-time–evolved snapshots

$$|\psi_0\rangle, \quad |\psi_1\rangle = e^{-iH\Delta t}|\psi_0\rangle, \quad |\psi_2\rangle = e^{-iH(2\Delta t)}|\psi_0\rangle, \ \ldots, \ |\psi_{m-1}\rangle = e^{-iH(m-1)\Delta t}|\psi_0\rangle,$$

which (for suitable Δt) provide a time-evolution Krylov family $\{|\psi_k\rangle\}_{k=0}^{m-1}$. Each state is realized by a short Hamiltonian-simulation circuit of duration $k\Delta t$ (e.g., Trotter, qDRIFT, or related short-time simulators), typically reusing the same template with a time parameter [48].

Each $|\psi_k\rangle$ is then *sampled* in the computational (Z) basis to produce a multiset $\mathcal{S}_k$ drawn from $|\langle x|\psi_k\rangle|^2$. The central insight is that, even if a target ground state $|\psi_{\mathrm{g}}\rangle$ is sparse in the computational basis, its important basis states need not be prominent in the *initial* $|\psi_0\rangle$. Short-time evolution "spreads" amplitudes so that, under the paper's assumptions on overlap and choice of Δt, every important basis element of $|\psi_{\mathrm{g}}\rangle$ becomes appreciable in *at least one* snapshot $|\psi_k\rangle$ with high probability [48]. Intuitively, the m snapshots provide m distinct linear combinations of eigencomponents, increasing the chance that any given important configuration is revealed in some $\mathcal{S}_k$.

From each $\mathcal{S}_k$ one retains dominant bitstrings (e.g., top–K_k by empirical frequency or all above a threshold), and then forms the union

$$\Omega_{\mathrm{Krylov}} = \bigcup_{k=0}^{m-1} \Omega_k.$$

Optionally, the experiments in [48] apply the SQD configuration-recovery step to repair particle-number errors before selection, and use a simple carry-over heuristic across passes; this is an implementation choice, not required by the core SKQD definition.

Under a sparsity model in which $|\psi_{\mathrm{g}}\rangle$ carries (nearly) all its weight on k basis states and $|\langle\psi_{\mathrm{g}}|\psi_0\rangle|$ is not exponentially small, Yu *et al.* show that the probability of *missing*

any important configuration decays *exponentially* with the number of shots per snapshot; correspondingly, the shots required per snapshot scale *logarithmically* in the number of important configurations (and in m via a union bound), with constants that depend on overlap and the chosen Δt [48]. This yields a polynomial-time recovery guarantee for the selected support.

Once Ω_{Krylov} is fixed, SKQD proceeds like sample-selected CI/SQD: build the projected Hamiltonian $H_{ij} = \langle x_i|H|x_j\rangle$ for $x_i, x_j \in \Omega_{\text{Krylov}}$ (in electronic Hamiltonians, Slater–Condon rules imply nonzero couplings only between determinants differing by at most two orbital occupations), and solve

$$H_{\Omega_{\text{Krylov}}}\,\vec{c} = E\,\vec{c}.$$

The lowest root E_{SKQD} is the ground-state estimate, and the associated eigenvector gives an explicit wavefunction in the selected basis. Because Ω_{Krylov} pools support discovered across multiple snapshots, the chance of capturing the important components *improves with* m (for appropriate Δt and overlap), while still requiring only Z-basis measurements [48]. These steps are summarized in Algorithm 10.

Theoretically, SKQD achieves polynomial–sample, polynomial–m convergence to within a target accuracy ϵ under the stated sparsity and overlap assumptions; the residual energy error can be bounded by the weight of any "missed" important configurations, and the probability of such misses is exponentially suppressed in the per–snapshot shot count [48]. Practically, SKQD has two notable advantages over standard (matrix–element) quantum Krylov diagonalization: (i) it avoids delicate estimation of many off–diagonal overlaps and Hamiltonian matrix elements on hardware, and (ii) it concentrates measurements on discovering basis support rather than on high–precision correlators, often reducing total shots for a given energy error. Moreover, Yu *et al.* report a large-scale demonstration computing the ground state of a single–impurity Anderson model with 42 orbitals (85 qubits) on an IBM Heron device, using $m = 5$ Krylov states and circuits with up to $\sim 6 \times 10^3$ two–qubit gates, achieving excellent agreement with DMRG for a system far beyond exact diagonalization [48]. SKQD retains the "measure–and–diagonalize" simplicity of SQD while adding a Krylov sweep that provides both empirical and provable coverage of the ground–state support.

A subtlety: In classical Lanczos, one would orthogonalize the Krylov vectors and do a tridiagonal matrix. In SKQD, we bypass the need to explicitly handle overlaps by projecting into the computational basis (which is orthonormal). Essentially, by taking the union of basis states, we are constructing an orthonormal basis (the bitstrings) that is a superset or equal to the span of the Krylov vectors' support. This avoids solving a generalized eigenproblem; we solve a standard eigenproblem in the subspace spanned by those basis states. It's approximate because we likely need many basis states to fully represent the continuous space of $|\psi_k>$, but sparsity assures us that we don't need all of them, only the important ones.

The proven theorem (in simplified form) is: if the ground state of H is supported heavily on at most k basis states (sparsity k), and the initial state has an overlap $\Omega(1/\text{poly}(n))$ with the ground state, then by choosing $m = O(k)$ Krylov steps and $N = \tilde{O}(k)$ samples per state, SKQD will with high probability find all those k important basis states and obtain an energy E_{SKQD} within ϵ of the true ground energy [48] (with ϵ depending on how large m and N are polynomially). In essence, one requires the ground

Algorithm 10 Sample-Based Krylov Quantum Diagonalization (SKQD)

Require: Hamiltonian H; time step Δt; Krylov order m; shots per snapshot $\{N_k\}$; per-snapshot selection rule (top-K_k or threshold τ); optional configuration recovery
Ensure: Approx. ground-state energy E_0 and coefficients $\vec{c}_0$
1: **Initialize:** Prepare a reference state $|\psi_0\rangle$ on the QPU.
2: **for** $k = 0$ **to** $m - 1$ **do**
3: **Krylov snapshot:** Prepare $|\psi_k\rangle = e^{-iH\,k\Delta t}\,|\psi_0\rangle$ via short-time simulation.
4: **Sample:** Measure $|\psi_k\rangle$ in the Z basis N_k times to obtain multiset $\mathcal{S}_k$.
5: **(Optional) configuration recovery:** Repair particle-number violations in $\mathcal{S}_k$ using the SQD recovery procedure; denote the recovered set $\widetilde{\mathcal{S}}_k$ (otherwise $\widetilde{\mathcal{S}}_k = \mathcal{S}_k$).
6: **Select dominant bitstrings:** From $\widetilde{\mathcal{S}}_k$, keep either the top-$K_k$ by empirical frequency or those above threshold τ; call the result Ω_k.
7: **end for**
8: **Union basis:** $\Omega \leftarrow \bigcup_{k=0}^{m-1} \Omega_k$ (deduplicate).
9: **Project (classical):** Build H_Ω with $H_{ij} = \langle x_i|H|x_j\rangle$ for $x_i, x_j \in \Omega$.
10: **Diagonalize:** Solve $H_\Omega \vec{c} = E\vec{c}$; take the lowest root $(E_0, \vec{c}_0)$.
11: **Return:** $E_0,\ \vec{c}_0$.

state to be sparse in the computational basis (which might be true in some cleverly chosen basis, but not in all problems – translationally invariant systems might require a momentum basis to be sparse, for example) [48]. Even if the ground state is not strictly sparse, if it has a dominant subspace, SKQD can be effective. For cases where the chosen basis isn't sparse, the performance will degrade (as more and more basis states carry weight, one may need prohibitively many samples). In the SIAM experiment mentioned, they actually performed a basis rotation (to momentum space for the non-interacting bath) to improve sparsity before applying SKQD [48]. This highlights that basis choice can be important for success.

SKQD can be viewed as a generalization of SQD: if one sets $m = 1$ (only the $|\psi_0>$ state), SKQD reduces to plain SQD (sample from a single state and diagonalize). If one increases m, SKQD systematically expands the search space by incorporating dynamics. It can also be seen as an alternative to KQD: rather than measuring overlaps $\langle\psi_i|H|\psi_j\rangle$ on quantum hardware, SKQD directly samples the states ψ_i to build a representation of them in the computational basis. This avoids having to handle overlap matrices and potentially yields better noise resilience. In numerical tests, SKQD outperformed a standard KQD approach in finding ground state energies when both had similar numbers of total samples, particularly because KQD's need to estimate many matrix elements can amplify the effect of noise [48].

SKQD is a powerful hybrid algorithm that leverages short-depth time evolution circuits and quantum sampling to construct a Krylov subspace in an implicit way. It offers polynomial-time convergence guarantees (under sparsity assumptions), bridging the gap between heuristic methods like SQD and rigorous algorithms like phase estimation or full KQD. The only quantum operations required are state preparations and measurements, making it suitable for near-term devices. SKQD represents the current state-of-the-art in quantum-classical algorithms for ground-state problems, having already demonstrated the ability to treat problems of unprecedented scale (dozens of qubits well beyond classical

exact limits) with accuracy matching trusted methods.

All four algorithms discussed – QSCI, SQD, Ext-SQD, and SKQD – are part of a growing family of quantum-enhanced subspace algorithms for eigenvalue problems. They share a common framework: use the quantum processor to explore the Hilbert space in ways difficult for a classical computer (either by preparing a complex wavefunction or by evolving it), then project the problem to a smaller subspace and solve it exactly (or to high precision) on a classical computer. We conclude this section by highlighting the key differences, advantages, and appropriate use-cases for each method:

- QSCI vs SQD: These are essentially the same base algorithm. QSCI was coined in the context of using a quantum state to select CI configurations in quantum chemistry [601]. SQD is a broader term emphasizing sample-based diagonalization. Both assume one has a reasonably good initial state to sample. They are heuristic but variational: if the initial state is good and the ground state is sparse, they can give very accurate energies and obey $E_{\text{approx}} \geq E_{\text{true}}$ always [42]. They require minimal quantum resources (only state preparation and bitstring measurements) and have been experimentally validated [42, 44, 45, 601, 603, 604]. However, without extension, they target primarily the ground state (excited states require either taking additional eigenvectors from the same subspace, which may be insufficient, or re-running with different inputs).

- Extended SQD (Ext-SQD): This method extends SQD to handle excited states and to improve ground-state accuracy by a second-round expansion of the basis using physical insight (excitations). It is particularly useful in quantum chemistry when one wants not just the ground energy but also excitation energies (e.g. singlet–triplet gaps, absorption spectra). Ext-SQD, as demonstrated by Barison et al., showed that a relatively inexpensive augmentation of the SQD subspace can yield accuracy comparable to much more expensive classical methods, and surpasses the straightforward QSE approach both in accuracy and efficiency [43]. The cost is an extra classical diagonalization and the need to generate determinants by applying excitations, which is feasible for singles/doubles (polynomial growth). Ext-SQD still keeps quantum costs low (no extra quantum circuits beyond the initial SQD sampling). This approach is ideal when one has a decent ground-state ansatz already and wants a systematic, non-variational way to obtain excited states and better ground state with error bars. It does rely on the idea that low-lying excited states can be constructed by excitations of the ground-state configurations – which is true for many cases (like single excitations for the first excited state in molecules), but might miss completely different states (e.g. states in a different symmetry sector could require a different initial ansatz or symmetry adaptation as in some recent works [6, 7]).

- SKQD: Sample-Based Krylov Quantum Diagonalization is the most general and theoretically grounded of these algorithms. It reduces to SQD in the simplest case, but by incorporating multiple Krylov-propagated states, it gains the ability to systematically improve and eventually converge to the true ground state (given sufficient resources) [48]. It stands out by providing provable convergence under reasonable assumptions, something the others lack. The quantum resource needed is more than SQD: one must be able to perform short-time Hamiltonian simulations

(which may involve gates scaling with system size and desired time). However, the time steps can be short (thus low-depth circuits), and one doesn't need an extremely accurate time evolution – Trotter errors just reflect in slightly different subspace which still converges with m. SKQD is well-suited for systems where one either does not have a very accurate variational ansatz readily available, or where one suspects that a single reference state might not have all the support needed. By sweeping with Krylov states, SKQD can capture physics that an ansatz misses. The cost is potentially a larger union of basis states (since each Krylov state will contribute some new configurations), but if the ground state is sparse, this remains manageable. SKQD's success in a problem with 85 qubits [48] is a testament to its scalability when combined with powerful classical computing for the final diagonalization. It essentially pushes the frontier of near term quantum computing in chemistry/physics by using the quantum device to generate states that classical methods cannot handle directly (like an 85-qubit correlated state), yet still solving the problem with classical post-processing by capitalizing on sparsity.

A few extra considerations:

- Simplicity and near-term feasibility: QSCI/SQD are simplest – they can be run on today's small quantum computers with minimal overhead. Ext-SQD adds only classical overhead. SKQD requires more complex circuits (for time evolution), which might be challenging on very noisy hardware unless error rates improve or error mitigation is used. So in the very near term, SQD and maybe Ext-SQD are more practical; SKQD is a bit more demanding but still NISQ-friendly relative to full phase estimation.

- When initial state is good vs. when it's not: If a high-quality guess state is available (e.g. a chemically inspired ansatz for a ground state), SQD can capitalize on that and likely needs only a small subspace. If one is unsure of the state or if the state has entangled components that are hard to capture in one ansatz, SKQD's Krylov exploration is more robust. Ext-SQD assumes the ground state is captured well enough by SQD and focuses on adding excited configurations – it shines in scenarios where excited states are like "promotions" of ground-state electrons.

- Computational scaling: All methods ultimately must diagonalize a matrix of size equal to the number of selected basis states. In SQD, that might be on the order of a few hundred to a few thousand in current applications [42]. Ext-SQD might double or triple that size by adding excitations, but also can sometimes prune the space to keep it in check [6, 43]. SKQD might build a subspace whose size is roughly m times the typical SQD size (since each Krylov state could contribute similar number of new basis states if they're substantially different). But SKQD could also choose to keep only the union of top configurations, so its subspace might not grow drastically if many Krylov states share the same important configurations. In any case, all these sizes are far below the full Hilbert space dimension, which is the whole point – they are feasible to diagonalize on classical machines (dense diagonalization of a few thousand dimension matrix is routine).

- Error mitigation and noise resilience: A big advantage of the sample-based approach (QSCI/SQD and SKQD) is that measuring in the computational basis is relatively

less sensitive to coherent errors – an error just might cause a wrong bitstring which can be filtered out via symmetry checks in configuration recovery [42]. There is no accumulation of Trotter error in SQD (since no time evolution), whereas SKQD does have Trotter or analog evolution error to manage. However, SKQD's theoretical analysis shows it tolerates certain noise in the Krylov state preparation without completely breaking the variational guarantee. In practice, each approach will require some error mitigation: SQD/Ext-SQD may require readout error mitigation and symmetry filtering; SKQD may require short-depth but many repetitions, possibly benefiting from error mitigation on those circuits.

These algorithms form a hierarchy: QSCI is the baseline sampling method, SQD includes a configuration recovery routine inspired by classical CI methods, Ext-SQD is a targeted improvement for excited states (and better ground state) using classical augmentation, and SKQD is a systematic extension that can asymptotically reach exact results by using quantum dynamics. Together, they illustrate the paradigm of using NISQ devices not to directly solve the eigenvalue problem, but to gather just enough quantum information (samples or short evolutions) so that a classical solver can finish the task. This hybrid strategy has already enabled record-breaking simulations [6,7,42–45,47,48,603,604], suggesting a promising route toward quantum advantage in scientifically relevant problems by pairing quantum sampling with powerful classical computation.

Sampling base methods replace long coherent evolutions with "measure–and–diagonalize" pipelines that are shallow on quantum hardware yet variational or provably convergent under mild sparsity assumptions. In realistic molecular and materials problems, however, even selected subspaces grow quickly: delocalized orbitals, near degeneracies, and dynamical correlation can demand thousands to millions of configurations before energies and response properties stabilize, and the projected Hamiltonian H_Ω can become a classical bottleneck. Moreover, sampling enough shots to reliably discover all chemically relevant configurations across multiple geometries or k-points remains nontrivial. These pressures motivate embedding and scale-out strategies that (i) partition the problem into weakly coupled fragments or impurities plus baths, (ii) downfold high-energy spaces into effective interactions, and (iii) distribute sampling and subspace assembly across classical/quantum resources. In the next section, we detail such embedding and scale-out workflows—how localized orbitals, frozen/active space choices, and fragment–bath formulations shrink the quantum footprint.

4.3 Reducing required qubit count and scaling-out

Sampling-based diagonalization excels when the relevant eigenstates have compact support in a judicious basis and when the selected subspace remains modest in size. Realistic chemistry and materials Hamiltonians, however, often exhibit extensive long-range couplings and large active spaces, so a direct sample-and-diagonalize pass can still produce thousands to millions of configurations. Embedding partitions the system into a chemically or physically motivated active region and an environment, delivering a reduced Hamiltonian in which strong correlations are localized. This can shrink both (i) the number of orbitals/qubits entering sampling and (ii) the selected basis size produced by QSCI/SQD.

A widely used method is DMET, which builds an *exact* one-shot quantum bath at the mean-field level and fits an impurity correlation potential self-consistently, [53,265,266,270].

Let P project onto an active-space Slater-determinant (or spin-orbital Fock) subspace and $Q = 1 - P$ onto its complement. A standard resolvent identity gives an energy-dependent effective Hamiltonian on P,

$$H_{\text{eff}}(E) \;=\; PHP \;+\; PHQ\Big(E - QHQ\Big)^{-1}QHP, \tag{4.1}$$

whose low-lying roots coincide with those of H. Embedding schemes approximate the self-energy term $PHQ(E - QHQ)^{-1}QHP$ by constructing bath orbitals that represent Q's entanglement with P.

DMET starts out from a mean-field Slater determinant $|\Phi\rangle$ on the full system. Pick the orbitals you care about (the fragment, A). From $|\Phi\rangle$, identify how A is entangled with the rest of the system by a simple singular-value (Schmidt) analysis of the occupied space. This produces one *bath* orbital for every fragment occupied orbital that is entangled with the environment. The fragment plus its baths define a small *impurity* space $\mathcal{I} = A \cup \{b_\mu\}$. Now project the *ab initio* Hamiltonian into $\mathcal{I}$ and solve this reduced problem with a correlated solver to obtain fragment densities (1-RDMs). Finally, adjust a static one-body "correlation potential" acting on A so that the impurity density matches the mean-field density on A; iterate these two steps to self-consistency. In our workflow the correlated solver inside the impurity is sample-based diagonalization (QSCI/SQD, Ext-SQD, or SKQD), which keeps the quantum circuits shallow while retaining systematic accuracy [53, 265, 266, 270].

A recent milestone by Shajan et al. was the combination of DMET with SQD as a subsystem solver [47]. In this DMET-SQD approach, each fragment's ground state is obtained via SQD on the quantum device, instead of a classical solver. This enabled simulations of chemically interesting systems with substantially fewer qubits than a full, unfragmented calculation. They used DMET+SQD to study a ring of 18 hydrogen atoms and various conformations of cyclohexane. The full molecules (which would require 41 and 89 qubits if treated monolithically) were decomposed into fragments needing only 27 and 32 qubits respectively. These active-region quantum simulations achieved ground-state energies in good agreement with high-level classical benchmarks. In other words, embedding allowed them to scale out to an 89-qubit problem by solving 32-qubit subproblems with manageable error. This marks tangible progress in the size of electronic structure problems addressable on current quantum hardware. The DMET-SQD study demonstrated the potential to accurately treat large molecules (e.g. cyclohexane and even prospective peptides/proteins) by splitting them into tractable fragments for near-term quantum computers. Of course, fragmentation is an approximation – e.g. it can miss some inter-fragment correlation like long-range dispersion – but expanding the fragment sizes via better quantum algorithms mitigates those errors. The takeaway is that embedding transforms a daunting large-scale simulation into a series of smaller quantum tasks, making optimal use of limited qubit counts.

In quantum chemistry, a related idea is the active space selection – effectively an embedding where one chooses a subset of important orbitals to treat with an exact method while keeping the rest fixed at a mean-field level. Sampling-based algorithms can leverage active spaces to focus on strongly correlated electrons. For example, we applied a local embedding approach to study a surface chemical reaction (oxygen reduction on a battery electrode). We first used classical analysis (density difference and natural orbitals)

to identify the key orbitals involved in the reaction, then defined those as the active space. This active region was then solved using SQD (and the extended variant), while the remaining bulk was treated at a simpler level. The result was improved ground-state energies for the reaction complex, surpassing the accuracy of conventional CCSD and agreeing well with complete active space CI (CASCI) references where those could be obtained. This example, detailed further in Chapter 5 highlights how embedding plus sampling can tackle complex material systems: a localized reaction site is embedded in a larger material, and only that site's correlated orbitals are handled by the quantum algorithm, dramatically reducing the problem size without sacrificing much accuracy. Similarly, Kaliakin et al. showed that SQD on supramolecular dimers (water–water and methane–methane complexes) yields accurate noncovalent interaction energies when an appropriate active space is chosen [44]. They executed 27- and 36-qubit circuits for these dimers – achieving binding energies within 1 kcal/mol of high-level CCSD(T) results in the equilibrium geometries. Even a 54-qubit SQD experiment was attempted to push the limits for a larger hydrophobic dimer. These cases demonstrate how choosing a subset of orbitals (e.g. valence orbitals in a dimer or active sites in an enzyme/battery interface) for the quantum treatment allows scaling to larger molecules than would otherwise be possible.

Another innovative technique for scale-out is entanglement forging (EF). EF effectively splits a quantum system into two classically-coupled halves, cutting qubit requirements roughly in half by representing two qubits' worth of information on one physical qubit [603, 605]. In the context of molecular simulations, EF typically exploits a partition (such as separating spin-up and spin-down orbitals) where the entanglement between the two partitions is managed via classical post-processing. Smith et al. recently combined entanglement forging with SQD to simulate a hydrogen-abstraction reaction in an organic molecule [603]. In a standard encoding, each spin-orbital is a qubit; under EF, each spatial orbital became one qubit (with the two spin states handled via classical combinations), halving the number of qubits needed [605]. They executed the combined EF+SQD algorithm on an IBM Heron superconducting processor to calculate reaction energies for 2,2-diphenyldipropane (a model for radical chain reactions). The study provided a detailed derivation of how to merge EF with SQD and assessed its accuracy across different active space sizes. By using EF, the required qubits were dramatically reduced (a crucial advantage for near-term devices), while SQD ensured that electron correlation in the reduced space was still captured by sampling and diagonalizing in a subspace. This approach illustrates an embedding in Hilbert space: part of the wavefunction is implicitly handled classically, allowing the remaining part to fit on current quantum hardware. The successful application to a realistic chemistry problem (with results validated against known values) underscores how hybrid quantum-classical encoding methods can scale out the reach of sampling algorithms when hardware resources are limited.

In addition to algorithmic embedding, distributed computing plays a pivotal role in scaling up quantum simulations. The concept of quantum-centric supercomputing has emerged, wherein a quantum processor is tightly integrated with classical HPC infrastructure [1, 42]. In this architecture, the workload is split so that classical nodes handle data-intensive or easily parallelizable steps, while the quantum device focuses on the intrinsically quantum part of the computation. Classical HPC nodes compute molecular integrals, set up quantum circuits, and post-process measurement results (e.g. projecting

sampled configurations and diagonalizing matrices), whereas the quantum processor is invoked to generate the raw sample data from circuit executions. By distributing these tasks, we can scale computational capacity far beyond what the quantum chip alone could do, essentially by leveraging the massive parallelism and memory of classical clusters in tandem with the quantum advantage on the fragment it handles. A landmark demonstration of this scale-out strategy was the SQD work of Robledo-Moreno et al., who coupled an IBM quantum processor with the 6400-node Fugaku supercomputer [42]. This hybrid setup produced accurate energy estimates for molecules that are far too large for classical brute force – as we mentioned earlier, the [4Fe–4S] cluster active space was mapped to 77 qubits.

Recently, working on top of this paradigm, this same work has been extended to incorporate optimization routines with the SQD loop [606]. Shirakawa et al. have shown that tighter integration between the QPU and classical compute clusters now enable optimization routines to leverage SQD and iteratively update the ansatz parameters. This effectively allows SQD to drive a closed-loop variational optimizer. In practice, they run a differential-evolution loop in which multiple "walkers" propose new ansatz parameters; for each proposal the QPU only supplies Z-basis samples, while the HPC side performs configuration recovery, projects/diagonalizes the SQD subspace, and returns an energy used to update parameters. The workflow is asynchronous, while one population is sampling on the QPU, the other is being diagonalized on the cluster, so both accelerators stay busy. With a highly parallel Davidson solver and a 3-D communicator + load-balancing scheme, they diagonalize very large SQD subspaces in minutes and scale up to 152,064 Fugaku nodes, executing on Heron with 72 primary qubits (+5 ancilla). On the same challenging Fe–S clusters (36-orbital active spaces) from the previous work [42], the closed-loop SQD optimization improves over older results for [4Fe–4S] and, after energy–variance extrapolation, attains energies within 0.1 Ha of DMRG for both systems, the state of the ar method for "classical" quantum chemistry

These results show that such distributed quantum-classical computing can achieve chemical accuracy for systems beyond the reach of exact methods, by offloading the majority of computation to classical processors and only calling the quantum device for its core competency (evaluating quantum amplitudes). In practical terms, this means current noisy quantum processors – when embedded in a supercomputing workflow – can start solving useful problems that neither quantum nor classical alone could tackle. This example started the notion that hybrid architectures are key to scaling up: by parallelizing measurements and classical computations, one can increase the effective problem size solvable with a given quantum device. The scale-out approach harnesses HPC to coordinate many sampling tasks, handle large data outputs, and perform heavy classical linear algebra, thereby extending the reach of sampling-based quantum algorithms to regimes previously thought intractable on near-term hardware.

Alongside hardware and architecture solutions, algorithmic innovations within sampling-based methods themselves contribute to scaling. A further advance has been incorporating randomized compiling techniques to push SKQD to larger problem sizes. Piccinelli et al. introduced an algorithm called SqDRIFT, which merges SKQD with the qDRIFT random Trotterization approach [599]. qDRIFT replaces a long deterministic Trotter sequence with a series of short random Hamiltonian exponentials that approximate the same evolution on average, dramatically reducing circuit depth. By using SqDRIFT,

the team could carry out SQD-based simulations for molecular Hamiltonians that would otherwise require prohibitively deep circuits. Crucially, this randomized SQD preserved the convergence guarantees of SKQD while enabling utility-scale simulations on current hardware. They demonstrated ground-state energy calculations for large polycyclic aromatic hydrocarbons beyond the size of any exact diagonalization, made feasible by these depth reductions. Algorithmic refinements like this contribute to scaling out the effectiveness of sampling methods. They either expand the size of the effective basis (thus needing fewer samples or qubits to reach the same accuracy), or reduce the resource cost per sample (thus enabling deeper physics to be captured on shallow circuits). In combination with embedding and HPC parallelism, these improvements help push the frontier of what near-term quantum algorithms can solve.

4.4 Hardware and chemistry case studies

We've already mentioned multiple use cases of Sampling based algorithms including our own which will be discussed in the next chapters. Still, it's worthwhile noting a few more and consolidating it all in this section.

- **Quantum-Centric Supercomputing for Molecules (2025).** Robledo-Moreno et al. introduced SQD and combined IBM's 133-qubit Heron R1 devices with the 6,400 classical nodes (Fugaku supercomputer) to simulate challenging molecular systems. They computed the N2 triple-bond dissociation (in cc-pVDZ basis) and electronic structures of [2Fe–2S] and [4Fe–4S] iron-sulfur clusters using 58, 45, and 77 qubits respectively. Circuit sizes reached up to 10,570 total gates (3,590 two-qubit gates). Despite noisy hardware, their hybrid sampling+classical estimator produced ground-state energies beyond the reach of exact diagonalization, providing chemical accuracy for systems around the hundred-qubit scale. This marked the largest electronic structure simulation on quantum hardware at the time (77-qubit active space) [42].

- **Excited States via Extended SQD (2025).** Barison, Robledo Moreno, and Motta extended SQD to handle excited states ("Ext-SQD"). They applied Ext-SQD to calculate the lowest singlet (S_1) and triplet (T_1) excitation energies of N_2, as well as ground and excited states of a [2Fe–2S] cluster on an IBM's 133-qubit Heron R1. The extended algorithm adds a modest computational step but significantly improves accuracy over standard ground-state SQD, and also surpasses the more common Quantum Subspace Expansion approach in efficiency. Using a correlation-consistent basis for N_2 (requiring on the order of the 58-qubit active space used in the earlier N_2 study) and the 45-qubit [2Fe–2S] space, they obtained excited-state energies that closely matched experimental and high-level reference values. This study, highlights that Ext-SQD can reliably access excited electronic states on current hardware, an essential capability for simulating spectra and photochemistry [43].

- **DMET-SQD for Extended Molecules (2024).** Shajan et al. performed density matrix embedding theory with SQD to study larger molecules. They simulated a ring of 18 hydrogen atoms and four conformers of cyclohexane by splitting each full

system (41 and 89 qubits) into active fragments of 27 and 32 qubits, which they ran on an IBM 127 qubit Eagle-family processor. Only 3–6 quantum circuits were needed per calculation due to the efficient sampling approach. The quantum energies (with configuration recovery post-processing) agreed with high-level classical benchmarks (FCI-based DMET and Heat-Bath CI), demonstrating accurate fragment energies for systems far beyond previous hardware studies. This was a notable scale-up to 32-qubit chemical simulations on NISQ hardware, showing that fragment embedding plus SQD can tackle larger molecular systems [47].

- **Noncovalent Interactions (2024).** Kaliakin et al. reported the first quantum-centric simulations of supramolecular (noncovalent) interactions. They used SQD to compute potential energy surfaces of the water dimer (hydrogen-bonded, hydrophilic) and methane dimer (dispersion-bound, hydrophobic). Using 27-qubit and 36-qubit circuits on an IBM 127 qubit Eagle-family processor, they obtained binding energies within 1 kcal/mol of high-level classical results (CASCI and CCSD(T)) in the equilibrium region. They further pushed to a 54-qubit SQD experiment to test the method's limits for the methane dimer's weak dispersion force. The SQD energies remained within chemical accuracy (1.6 mH) of the gold-standard values in those cases, a significant achievement for quantum hardware simulation of van der Waals complexes [44].

- **Methylene Singlet–Triplet Gap (2024).** Liepuoniute et al. studied the singlet and triplet states of the CH2 molecule, a classic multi-reference problem. They modeled CH2 in a (6e, 23o) active space and ran a 52-qubit experiment on an IBM 127 qubit Eagle-family processor. Using SQD within a quantum-centric workflow, they computed the dissociation energies of CH2 in its ground-state triplet and first excited singlet. The SQD results for the singlet state were only a few mHa off from the reference Selected CI values, and the singlet–triplet gap predicted by SQD aligned well with experimental and high-level theoretical values. The triplet state showed somewhat larger variance (attributed to the greater complexity of an open-shell wavefunction and bitstring sampling issues). Nevertheless, this study demonstrated that SQD can treat an open-shell molecule on hardware with 50 qubits, yielding a correct excitation energy gap – an important milestone in quantum computing for computational spectroscopy [604].

- **Quantum Krylov for 2D Materials (2024).** Yoshioka et al. demonstrated Krylov Quantum Diagonalization (KQD) on a 2D lattice with up to 56 sites/qubits. They used an IBM superconducting processor to implement trotterized time evolutions on a 2D Hubbard-like model, constructing subspaces in which they diagonalized the Hamiltonian. The experiment successfully obtained low-energy eigenvalues for interacting 2D systems beyond the size reachable by exact diagonalization (56 qubits). This study showed that quantum subspace algorithms (like KQD) scale to two-dimensional many-body systems on real hardware, complementing classical Krylov methods. It established a "world record" system size (56 qubits) for quantum eigenenergy computations on a NISQ device, with results indicating that quantum diagonalization can indeed tackle models intractable classically, given sufficient qubit count and error mitigation [46].

- **Krylov SQD for Lattice Models (2025).** Yu et al. introduced Sample-Based Krylov Quantum Diagonalization (SKQD), combining Krylov subspace expansion with sampling. They validated SKQD on quantum hardware by simulating the Single Impurity Anderson Model (SIAM), a strongly-correlated lattice problem. Using IBM's 133-qubit Heron R1, they ran two cases: one with 29 bath sites (61 qubits) and one with 41 bath sites (85 qubits). Circuits were constructed via trotterized time evolution (Krylov vectors), reaching up to 6,153 two-qubit gates in the largest 85-qubit case. Despite the depth, the SKQD+configuration-recovery approach yielded ground-state energies in excellent agreement with DMRG (matrix-product-state) benchmarks. This represented the largest quantum simulation of an impurity model to date (85 qubits), and showed that Krylov-based sampling can outperform naive Krylov methods under hardware noise. Notably, SKQD bypassed VQE optimization and heavy measurement overhead by leveraging short-time evolution samples, proving to be a robust approach for lattice quantum chemistry on NISQ devices [48].

- **Randomized Compiling for SQD (2025).** Piccinelli *et al.* introduced SqDRIFT, a sample–based Krylov diagonalization variant that employs randomized Trotter compilation qDRIFT to reduce circuit depth while retaining convergence guarantees of Krylov–style phase methods. By randomizing the ordering of Hamiltonian terms in short–time evolutions, qDRIFT induces average–error cancellation, enabling much shallower circuits for preparing the Krylov ladder. SqDRIFT was demonstrated on polycyclic aromatic hydrocarbons: naphthalene ($C_{10}H_8$, 48 spin–orbitals) and coronene ($C_{24}H_{12}$, $\sim$108 spin–orbitals). The 48–qubit naphthalene case was classically tractable for validation, whereas the coronene active space ($\sim$108 qubits) was executed on IBM's 156-qubit Heron R2. On hardware for coronene, a Local Unitary Cluster Jastrow (LUCJ) reference ansatz required $\sim$ 2,600 two–qubit depth with thousands of CNOTs, while SqDRIFT attained comparable subspace coverage at substantially lower depth by interleaving randomized short–time Trotter steps. The resulting ground–state energy for coronene approached an estimated FCI benchmark within a few mE_h, outperforming standard SQD (slower convergence) and even surpassing CISD in accuracy. These results indicate that randomized compiling can push quantum–chemistry calculations to 100+ qubits NISQ hardware, yielding error–mitigated energies competitive with coupled–cluster quality [599].

- **Entanglement Forging + SQD for Reactions (2025).** Smith et al. tackled a hydrogen-abstraction reaction (a radical chain reaction step) using a hybrid of Entanglement Forging (EF) and SQD. EF is a qubit-reduction technique that maps each qubit to a spatial orbital (rather than a spin-orbital), halving the qubit count required. They studied H-abstraction from 2,2-diphenyldipropane (a model radical in polymer degradation) and computed its activation energy and reaction energy on an IBM's 156-qubit Heron R2. By forging entanglement between two halves of the system (spin up and spin down), the largest active space (e.g. 40 spin-orbitals) could be mapped to 20 physical qubits, making the simulation feasible on hardware. Across various active-space sizes, the combined EF+SQD approach showed consistent accuracy, and the extracted reaction energetics closely matched those from classical quantum chemistry (within chemical accuracy) while using half the qubits otherwise needed. This work illustrates how mixing qubit reduction techniques with sampling-based algorithms can further extend the reach of quantum simulations – in this case

enabling a realistic chemical reaction energy profile to be calculated with a modest quantum resource count [603].

- **Closed-Loop SQD Optimization (2025).** In an extension of Robledo-Moreno *et al.* [42], Shirakawa *et al.* demonstrate a fully closed-loop SQD workflow in a quantum-centric supercomputing setting: the QPU supplies only Z-basis samples, while a large classical backend performs configuration recovery, projection, and Davidson diagonalization and feeds the resulting energies back to a population-based differential-evolution optimizer that updates the parameters of a LUCJ ansatz. The pipeline is asynchronous (sampling and diagonalization overlap in time) and massively parallel (many walkers and batches evaluated concurrently), so both quantum and classical resources are continuously utilized. By evaluating the objective from *samples*—instead of measuring thousands of Pauli expectations as in VQE—the per-iteration quantum cost becomes largely decoupled from the Hamiltonian term count; accuracy and wall time can be traded for classical nodes, while circuits remain shallow. On challenging Fe–S clusters, the closed-loop SQD runs improve upon prior (non-optimized) SQD baselines and, with energy–variance extrapolation, yield ground-state energies in close agreement with DMRG. Methodologically, inserting optimization inside the SQD loop is disruptive: it turns SQD from a one-shot sampler into a scalable variational driver that exploits HPC parallelism (many proposals, many subspace solves) while preserving SQD's hardware frugality. [606]

And finally the two works discussed in the next chapters in this thesis:

- **Surface Chemistry (Battery Reactions, 2025).** We combined Local Unitary Cluster Jastrow ansatz with an extended SQD (Ext-SQD) to simulate an electrochemical surface reaction (oxygen reduction on a Li–O2 battery cathode). We selected an active space of 32 orbitals for the Li–O2 surface and ran circuits on IBM's 156-qubit Heron R2 hardware. The largest circuit used 80 qubits (64 orbitals + 16 ancilla) with a two-qubit gate depth 2,714 and 5,455 CNOTs. Despite the circuit complexity, Ext-SQD improved ground-state energy estimates beyond coupled-cluster (CCSD), nearing CASCI (exact active-space CI) accuracy. For active spaces up to 80 qubits, the quantum-classical method achieved higher accuracy than CCSD (errors < chemical accuracy) for the Li–O2 reaction, showcasing the potential of sampling-based algorithms for complex surface chemistry problems on NISQ devices [6].
- **Band Gap Prediction in Materials (2025).** We applied SQD to periodic solid-state systems. Using 38- and 46-qubit circuits on IBM's 156-qubit Heron R2 processor, we simulated the ground-state of bulk silicon and metal oxides in a localized basis and successfully predicted their band gap in agreement with experiment. The workflow involved projecting DFT electron density into an Extended Hubbard Hamiltonian and preparing a LUCJ ansatz state. By sampling this state's bitstrings and diagonalizing the Hamiltonian in that subspace, they obtained the band gap with minimal error mitigation matching experimental band gaps using nearly 50 qubits on hardware [7].

These are summarized in Table 8

Note that this is not an exhaustive list. Currently, this is a moving target, and research on sampling based algorithms has had frequent publications over the past year.

Table 8 – **Sampling–based quantum chemistry and many–body case studies on hardware.**

Study (Year)	**Target system(s)**	**Qubits**	**Device**	**Method**
Robledo–Moreno (2025)	N_2; [2Fe–2S], [4Fe–4S]	58, 45, 77	Heron R1	SQD
Barison (2025)	N_2 (S_1/T_1); [2Fe–2S]	∼58, 45	Eagle	Ext–SQD
Shajan (2024)	H_{18} ring; cyclohexane (DMET frags)	27, 32	Eagle	DMET + SQD
Kaliakin (2024)	$(H_2O)_2$, $(CH_4)_2$ (PES)	27, 36, 54	Eagle	SQD
Liepuoniute (2024)	CH_2 (6e, 23o), S/T states	52	Eagle	SQD
Yoshioka (2025)	Interacting 2D lattice (Hubbard-like)	56	Eagle	KQD
Yu (2025)	SIAM (29/41 bath)	61, 85	Heron R1	SKQD
Piccinelli (2025)	Naphthalene; Coronene	48 (sim), ∼108 (hw)	Heron R2	SqDRIFT
Smith (2025)	H–abstraction (radical reaction)	∼10–20 (EF)	Heron R2	EF + SQD
Shirakawa (2025)	[2Fe–2S], [4Fe–4S]	77 (72 + 5 anc.)	Heron R3	SQD (optimizer)
Duriez (2025)	Bulk Si (periodic, band gap)	38, 46	Heron R2	Ext–SQD/SQD
Barroca (2025)	Li–O_2 surface (ORR model)	up to 80	Heron R2	Ext–SQD

4.5 Practical guidance and pitfalls

Sample–based diagonalization methods shift most of the algorithmic burden from long coherent evolutions and global variational searches to short, hardware–friendly state preparations, repeated Z–basis measurements, and classical diagonalization in compact, problem–adapted subspaces. This section distills implementation guidance and common failure modes from recent hardware and large–scale demonstrations in chemistry and many–body physics [7, 42–44, 47, 48, 599, 603, 604, 606].

Pre–processing: basis, active space, and mapping

- **Choose a sparse–friendly basis.** SQD/SKQD benefit when the target eigenstates have concentrated weight on a modest set of computational–basis configurations. In quantum chemistry, localized orbitals (Boys/PM/ER) often yield sparser determinant support than canonical orbitals; in lattice/impurity models, partially diagonalizing non–interacting blocks (e.g., transforming the bath to momentum space) can dramatically improve sparsity [48]. If the ground state is delocalized in the chosen basis, sampling must grow substantially to capture its support.

- **Right–size the active space.** Follow the chapter's basis/ECP workflow and freeze chemically inert cores; for embedded problems (DMET, Active Space selection, EF), ensure the fragment/bath captures the target correlations [7, 47]. Poor active–space choices manifest as persistent bias even as the selected subspace grows.

- **Exploit symmetries at the bitstring level.** Enforce (N, S_z), spin/parity tapering when the qubit mapping allows, and point–group irreps at the sampling and post–selection stages to reduce qubits and lower the effective sample complexity by discarding unphysical shots [42, 43].

Preparing the trial family

- **Trial state and circuit depth matters.** For SQD/QSCI, prepare a reference with non–negligible ground–state overlap (HF, LUCJ, short–depth UCC). Use shallow, hardware–tailored templates. For SKQD, the initial state can be simpler because short–time Krylov snapshots redistribute amplitude and expose support that is hidden in the initial state [48]. Prefer short segment time steps (e.g., $t = \Delta t$) and randomized compilation (qDRIFT) to average Trotter errors at fixed depth [599].

Sampling strategy and shot budgeting

- **Allocate shots adaptively.** Start with coarse sampling to estimate the tail of the distribution; allocate additional shots only until the selected set stabilizes (no new configurations above threshold appear). In SKQD, distribute shots unevenly across k to emphasize snapshots that empirically yield richer support.
- **Symmetry post–selection.** Discard bitstrings violating particle number/spin/parities; this boosts the effective sample efficiency and mitigates readout flips that change Hamming weight [42].
- **Configuration recovery.** After sampling, apply deterministic excitation generators to (re)introduce physically connected configurations that may have been missed due to noise or finite sampling (singles/doubles on selected determinants for molecules; analogous local moves for lattice models). Configuration recovery is essential both in SQD and SKQD to stabilize energies and improve subspace completeness [42, 43, 48].

Selecting and growing the subspace

- **Ranking and thresholding.** Rank by empirical frequency; pick top-K or use a probability floor. In noisy regimes, prefer a hybrid rule: keep all above a conservative floor and pad to K with the next most frequent valid bitstrings.
- **De–duplication across sources.** When merging subsets (multiple snapshots, multiple geometries, multiple fragments), de–duplicate and retain a provenance map (which state generated each configuration).
- **Monotone improvement.** Energies from subspace diagonalization are variational in the chosen basis; they should decrease (or stagnate) as the subspace grows. Lack of monotonicity signals numerical issues (ill–conditioning) or inconsistent Hamiltonian assembly.

Hamiltonian assembly and classical diagonalization

- **Matrix elements.** Use Slater–Condon rules (chemistry) or local operator tables (lattice) to assemble $H_{ij} = \langle D_i|H|D_j\rangle$ only for connected pairs (differing by ≤ 2 orbital substitutions for two–body Hamiltonians).

- **Conditioning.** Selected subspaces can be nearly linearly dependent (especially after configuration recovery and in embedded fragments). Use symmetric orthogonalization ($S^{-1/2}$) and stabilized eigensolvers (Davidson/Lanczos with preconditioning and DIIS–style mixing).
- **Scaling and memory.** Track K; classical cost is $O(K^2)$ storage and $O(K^3)$ diagonalization (less with iterative solvers). On HPC, distribute rows of H and use block Davidson to extract multiple low roots simultaneously; checkpoint matrices and provenance to enable reproducible restarts [42].

Excited states and Ext–SQD specifics

- **Orthogonality enforcement.** For multiple states, build either (i) a single augmented subspace that supports several low eigenpairs (preferred), or (ii) state–specific subspaces with explicit Gram–Schmidt orthogonalization constraints. Ext–SQD's deterministic excitation closure is designed to supply the configurations necessary for low–lying excited states [43].

SKQD hyperparameters: Δt and m

- **Short times, several snapshots.** Favor small Δt to keep circuits shallow and Trotter errors moderate; use several snapshots m to sweep amplitude across support [48]. Empirically, $m \in [3, 7]$ with hundreds–to–thousands of shots per snapshot worked well up to 85 qubits [48].
- **Randomized compilation.** qDRIFT (randomized Trotter) lowers coherent bias at fixed depth and performed well in 100+ qubit chemistry (SqDRIFT) [599]. Keep per–snapshot depth bounded; trade additional snapshots for depth when noise is dominant.

Error mitigation on samples

- **Readout/noise filtering.** Apply symmetry filters; optionally correct with calibration matrices for readout error. In some workflows, Bayesian denoising of histograms improves the tail estimate without over–selecting spurious states [42].
- **Observable–level mitigation.** Because energies are obtained via exact diagonalization of a *classically assembled* H, mitigation focuses on *selection bias*, not measurement of H on hardware. Configuration recovery and symmetry filtering are the main levers; zero–noise extrapolation is less central than in VQE.

Validation, stopping, and reporting

- **Internal validation.** (i) Monitor energy monotonicity under subspace growth; (ii) track overlap of successive eigenvectors; (iii) report subspace completeness proxies

(e.g., cumulative sampled weight $\sum_{x\in\Omega}\hat{p}_x$).

- **External checks.** Where possible, compare to CASCI/selected–CI or DMET fragment solvers on the same active space; in periodic/embedded contexts, reconcile with HF/DFT band edges or known gaps [7, 47].
- **Stopping criteria.** Stop when: (i) adding configurations changes E below a tolerance; (ii) re–sampling does not add new configurations above threshold; (iii) orthogonality–constrained excited–state gaps stabilize.

Common pitfalls (and fixes)

- **Non–sparse targets.** If states are delocalized in the computational basis, selection becomes expensive. *Fix:* change basis (localize orbitals; diagonalize non–interacting blocks), use embedding to reduce entanglement, or switch to SKQD to sweep support [48].
- **Selection bias from noise.** Spurious frequent bitstrings can enter the subspace. *Fix:* aggressive symmetry post–selection, configuration recovery from trusted seeds, cross–validation by re–sampling.
- **Ill–conditioned subspaces.** Near–linear dependence leads to unstable diagonalization. *Fix:* orthogonalize ($S^{-1/2}$), prune near–duplicate configurations, regularize low–norm vectors.
- **Under–resolved excited states.** Basic SQD subspaces may lack configurations for excited states. *Fix:* Ext–SQD excitation closure and spin/irrep filtering [43].
- **Overly deep Krylov circuits.** Trotter depth can dominate errors. *Fix:* smaller Δt, fewer Trotter steps, randomized compilation (SqDRIFT) [599].

Takeaways

For chemistry and materials on NISQ hardware, the most reliable pattern is: (i) embed or downfold to compact fragments/bands; (ii) prepare shallow, symmetry–respecting trials (LUCJ/short–UCC); (iii) sample aggressively with symmetry post–selection; (iv) apply configuration recovery; (v) diagonalize and iterate subspace growth until energies stabilize. For tougher, less–sparse targets, SKQD adds a principled Krylov sweep that (provably) uncovers missing support while retaining shallow circuits and Z–basis measurements [48]. The recent case studies discussed previously validate that these practices deliver robust, credible results on today's hardware [7, 42–44, 47, 48, 599, 603, 604].

4.6 Summary

This chapter surveyed sample-based diagonalization as a pragmatic alternative to deep-time-evolution and large-parameter variational approaches. The common template

is to (i) prepare one or more shallow states with non-negligible ground-state overlap, (ii) sample them in a convenient basis to discover a compact set of important configurations, and (iii) diagonalize the Hamiltonian in that data-defined subspace. Within this template we detailed three families: quantum selected CI (QSCI) and its chemistry incarnation, sampled quantum diagonalization (SQD); extended SQD (Ext-SQD), which augments a sampled subspace with systematically generated excitations to recover ground and excited states; and sample-based Krylov diagonalization (SKQD), which sweeps short-time evolutions to provably expose the support of low-energy eigenstates. Across these methods, we emphasized configuration-recovery, variance reduction, and error-mitigation as indispensable ingredients for chemical accuracy on pre–fault-tolerant hardware.

Scale remains the central challenge: realistic molecules, surfaces, and solids still generate large active spaces and dense Pauli decompositions. To keep quantum resources bounded, we tied sampling methods to embed and scale-out strategies: fragment the problem (DMET/active space selection/EF), and then apply SQD/Ext-SQD/SKQD as the high-accuracy solver inside each embedded subproblem. This "measure–then–diagonalize" workflow shifts depth and coherence demands from the QPU to classical HPC while retaining a variational or near-variational control knob via subspace enlargement. Hardware demonstrations in the literature already show this synergy at 50–100+ logical qubits with chemically meaningful observables, provided one enforces symmetries, prunes noise via configuration recovery, and exploits locality in both the ansatz and measurement.

The next two chapters instantiate this program on two demanding, but complementary, domains:

1. **Surface electrochemistry (battery reactions).** We apply Ext-SQD as a correlated solver for metal–oxygen surface models relevant to Li–O_2 batteries. Starting from LUCJ trial states, we sample, recover, and *extend* configuration spaces to capture reaction energetics. We show that, for active spaces up to $\sim$75 qubits, Ext-SQD systematically narrows the gap to CASCI and surpasses CCSD-level baselines for key steps of the oxygen-reduction reaction, all on contemporary superconducting hardware with modest mitigation [6].

2. **Periodic materials (band gaps).** We use SQD within a periodic embedding workflow to estimate band gaps in bulk silicon and transition metal oxides from Extended Hubbard Hamiltoninas with $\sim$36–48 qubit experiments. Short-time–prepared LUCJ states are sampled and diagonalized in a symmetry-adapted subspace, yielding band gaps in agreement with experiment with minimal mitigation, thereby illustrating that sampling–driven subspace solvers can operate directly on crystalline systems [7].

Together, these case studies test the methodological claims of this chapter under two distinct stressors: heterogeneous, strongly inhomogeneous chemistry at interfaces, and translationally symmetric, delocalized electronic structure in solids. In both settings we will make the design choices explicit (active-space construction, configuration recovery, excitation pools, measurement budgets) and quantify accuracy against trusted classical references. Beyond validating SQD/Ext-SQD as credible solvers inside embedded workflows, the results point toward a near-term pathway where quantum hardware contributes targeted many-body correlation to otherwise classical multiscale simulations.

5 Quantum Simulation of Battery Electrode Reactions

5.1 Adapted Article: Scaling active spaces in simulations of surface reactions through sample-based quantum diagonalization

Chapter 4 established the measure–then–diagonalize paradigm and the configuration-recovery + mitigation toolkit needed for chemical accuracy on pre–fault-tolerant devices. Here we instantiate that program for heterogeneous electrochemistry: oxygen-reduction steps at Li$-O_2$ battery surfaces. The workflow couples local embedding and active-space construction (density-difference analysis followed by CC-natural-orbital refinement) with sample-based subspace solvers built from shallow LUCJ trial states. The chapter quantifies accuracy vs. trusted references (CASCI/HCI/CCSD) while making design choices explicit (excitation pools, recovery, budgets), extending the narrative from Chapter 4.

Provenance and license. This chapter adapts material from: M. A. Barroca, T. Gujarati, V. Sharma, R. N. B. Ferreira, Y.-H. Na, M. Giammona, A. Mezzacapo, B. Wunsch, and M. Steiner, *Surface Reaction Simulations for Battery Materials through Sample-Based Quantum Diagonalization and Local Embedding*, arXiv:2503.10923 (2025), DOI: 10.48550/arXiv.2503.10923. The preprint is distributed under a Creative Commons Attribution (CC BY 4.0) license; this adaptation complies with those terms (link). Text is independently rewritten but figures, tables and numerical results are reproduced as in the original. The original work is currently undergoing peer-review.

Abstract. Accurate quantum chemistry calculations are essential for understanding electronic structures, yet exactly solving the many-body Schrödinger equation remains impractical due to electron correlation effects. While Density Functional Theory (DFT) is widely used, its inherent approximations limit its ability to capture strong correlation effects, particularly in surface reactions. Quantum computing offers a promising alternative, especially for localized electron interactions. In this work, we apply a quantum embedding approach to study the Oxygen Reduction Reaction at Lithium battery electrode surfaces. We employ an active space selection strategy, using Density Difference Analysis to identify key orbitals, which are subsequently refined with coupled-cluster natural orbitals. These are treated using Sample-Based Quantum Diagonalization (SQD) and its extended version (Ext-SQD), leveraging the Local Unitary Cluster Jastrow (LUCJ) ansatz for efficient state preparation on quantum hardware. Ext-SQD significantly improves upon standard SQD by incorporating excitation operators into the quantum-computed electron configurations used in classical Quantum Selected Configuration Interaction. Our approach demonstrates improved accuracy over Coupled Cluster Singles and Doubles (CCSD) calculations over ground state energies. The results, validated against Complete Active Space Configuration Interaction (CASCI), whenever computationally feasible, Heat-Bath Configuration Interaction (HCI) and its extended variant, provide new insights into Li$-O_2$ surface reactions,

underscoring the potential of quantum computing in materials discovery and reaction modeling.

5.1.1 Systems and reaction workflow

Accurate computational modeling of correlated electronic systems is essential for predicting chemical and materials properties with high precision. However, solving the electronic Schrödinger equation exactly is computationally prohibitive due to the exponential scaling of the many-body wavefunction. This has necessitated the development of approximate methods that balance computational feasibility with accuracy. As discussed in Section 3.1, while DTF is widely used for electronic structure calculations, its reliance on exchange–correlation approximations leads to limitations in capturing strong electron–electron correlation effects, particularly in complex molecular and surface systems [607, 608].

One technologically important setting where these limitations matter is surface electrochemistry in Li–O_2 battery environments. The reaction between lithium and oxygen in Li-metal anode batteries is critical for stabilizing the solid electrolyte interphase (SEI) and enhancing performance. Although these batteries offer high energy density, they face safety challenges from dendrite formation, which can cause short circuits and fires. Dissolved oxygen in the electrolyte can mitigate dendrite growth by promoting the formation of lithium oxides, thereby improving both stability and cycle life [11]. The primary reaction is

$$4\,\mathrm{Li} + \mathrm{O_2} \longrightarrow 2\,\mathrm{Li_2O},$$

resulting in lithium oxide (Li_2O) deposition on the SEI. Under different local conditions, additional lithium oxides—such as lithium peroxide (Li_2O_2) and lithium superoxide (LiO_2)—may form, influencing battery longevity and performance [12, 13]. Previous studies have proposed reaction mechanisms for various surface-mediated electrochemical processes [609–611]. Following a similar approach, we investigate the formation of Li^+ and $O_2{}^-$ species on the surface through quantum chemistry calculations. A minimal initial step is

$$\mathrm{Li} + \mathrm{O_2} \longrightarrow \mathrm{Li^+} + \mathrm{O_2{}^-(ads)},$$

where $O_2{}^-$ adsorption nucleates subsequent growth of peroxide/oxide. The adsorption and stabilization of $O_2{}^-$ play a critical role in the initial stages of lithium oxide formation; their surface binding dictates the pathway for further oxide growth and ultimately impacts electrochemical performance [12, 14, 15].

In Chapter 2 we saw that hybrid approaches that combine classical and quantum computing resources - "quantum-centric supercomputing" - are being actively explored to overcome mean-field limitations while keeping quantum depth modest [1]. Quantum algorithms such as QPE promise asymptotically favorable scaling for electronic structure [372, 612, 613], but current hardware constraints necessitate methods that integrate shallow quantum subroutines with classical embedding and post-processing [614]. Previously we positioned is active space selection as a critical enabler, which reduces computational burden while preserving key correlations local to the reactive site. We employ Density Difference Analysis (DDA), often refined with coupled-cluster natural orbitals [250], and note complementary routes such as AVAS [570] and DMET [53].

Within a reduced active space, the VQE [30, 615, 616] has been widely studied, yet practical challenges remain, such as ansatz design, barren plateaus, and large sampling overhead, which limits scalability on pre–fault-tolerant hardware [598, 614, 617, 618]. This motivates the sample-based subspace approaches from Chapter 4 that minimize circuit depth and measurement counts while retaining a variational or near-variational control knob.

As discussed previously, SQD constructs a compact, data-defined subspace by sampling bitstring configurations from a shallow trial state and then diagonalizing the Hamiltonian in that subspace [42]. A configuration-recovery step enforces particle number and spin sector and denoises hardware outputs. Extended SQD (Ext-SQD) improves accuracy by augmenting the sampled subspace with selected excitation operators (e.g., screened singles/doubles), enabling better ground and excited states without deep circuits [43]. To generate informative samples at low depth, we use a truncated LUCJ ansatz [42, 619], initialized by coupled-cluster amplitudes to bias sampling toward chemically relevant configurations. Recent demonstrations show that LUCJ-seeded sampling with SQD/Ext-SQD can achieve competitive accuracy under realistic noise/connectivity [42, 43, 604, 619].

We study lithium–oxygen surface chemistry using a workflow that couples periodic DFT, automated active-space selection, and sample-based quantum diagonalization. Geometry optimization and mean-field references are obtained for Li slabs with adsorbed oxygen; DDA and natural-orbital analysis identify and refine compact active spaces around the reactive region; shallow truncated LUCJ circuits are prepared within each active space and sampled on quantum hardware (or compressed simulators for larger sizes); finally, SQD/Ext-SQD deliver ground-state energies for reactant and product, and reaction energetics are extracted as

$$\Delta E \;=\; E_{\mathrm{prod}} - E_{\mathrm{reac}}.$$

This end-to-end process is summarized in Figure 19.

The geometry optimizations were performed using DFT as implemented in the VASP software package [620–622]. The structures were treated as periodic in the x–y plane, with Li slab thickness of 6.1 Å and a vacuum layer of 12 Å along z to minimize image interactions. The Brillouin zone was sampled using a $4 \times 4 \times 1$ Monkhorst–Pack grid; the O_2 molecule was allowed to relax on the Li surface to represent adsorption accurately. Core states were treated with PAW pseudopotentials; valence states used a plane-wave basis; positions were relaxed, and the convergence tolerance was 10^{-6} eV.

Following geometry optimization, we computed energies and electronic densities for both reactant and product with PySCF [587,588] using a Gaussian-type orbital basis (GTH-DZV) [623], PBE exchange–correlation [624], and GTH pseudopotentials [623]. Electronic structure was evaluated at the Γ-point of the supercell, as additional k-points did not significantly alter the densities. The resulting Γ-point reaction energy, $\Delta E = -3.897$ eV, underestimates literature values of ≈ -6 eV for lithium oxides [625–628], underscoring the need for correlated treatments of the reactive region.

In the remainder of this chapter we instantiate the workflow above: we (i) construct DDA→natural-orbital active spaces centered on HONO–LUNO [250], (ii) sample truncated two-layer LUCJ states on the `ibm_aachen` processor (and MPS emulation for larger sizes), and (iii) apply SQD/Ext-SQD [42, 43, 601] to compute ground states and reaction energies, benchmarking against CASCI (where feasible), HCI/Ext-HCI, and CCSD. We show that

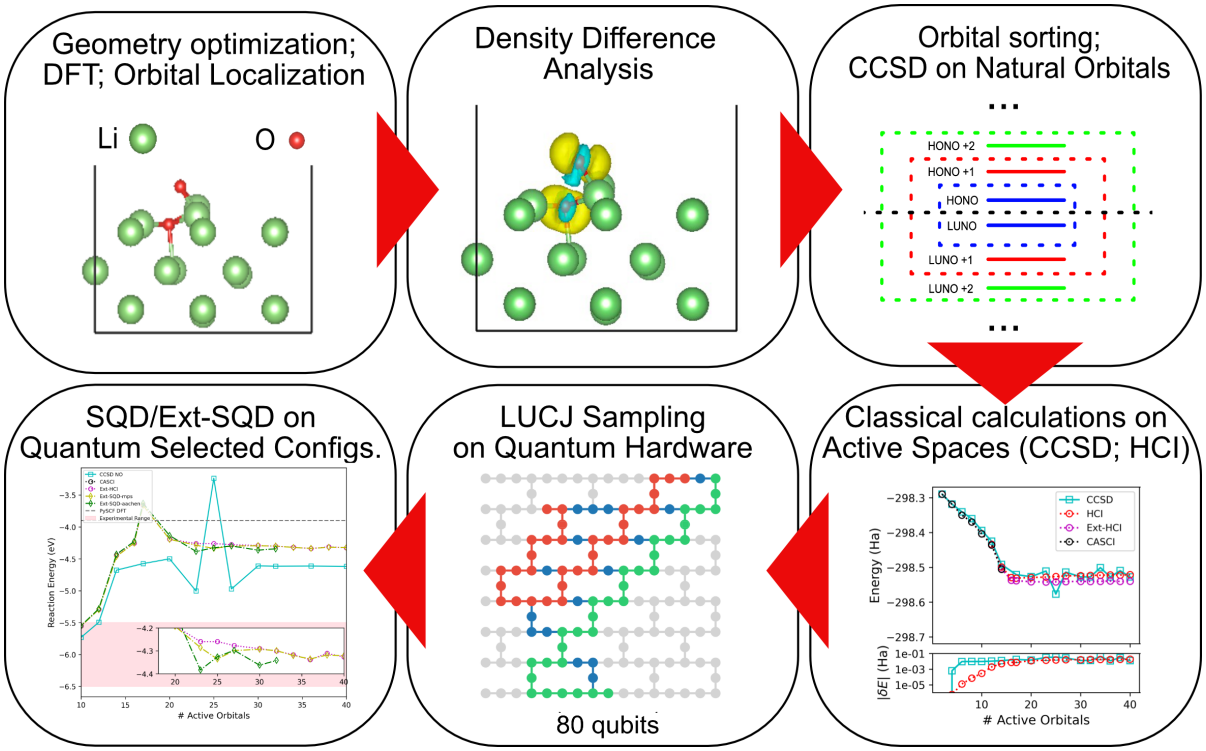


Figure 19 – **Hybrid quantum–classical framework for studying lithium–oxygen reactions.** Geometry optimization and DFT calculations lead into active-space selection via DDA and natural orbitals, identifying key orbitals near the reactive site. A truncated LUCJ ansatz is sampled on quantum hardware (or compressed simulators), and SQD/Ext-SQD compute reaction energies ($\Delta E = E_{\text{prod}} - E_{\text{reac}}$), providing insights into lithium oxide formation. *Reproduced from* M.A. Barroca, et al., [6] (DOI) CC BY 4.0 (link).

Ext-SQD systematically improves over SQD and compares favorably with variational classical references, while avoiding the sampling and depth burdens of large-parameter VQE. This supports sample-based diagonalization inside embedded, periodic workflows as a pragmatic route to chemically meaningful accuracy on pre–fault-tolerant hardware.

5.1.2 Methods and Results

Typically, the most effective active space selection methods involve valence electrons and orbitals, with further reductions permitted when chemically justified. It is crucial, however, that all orbitals contributing significantly to static correlation are included to maintain accuracy [570, 629, 630].

Studies have demonstrated that, in certain cases involving molecular adsorbates on surfaces, electron correlation effects are largely confined to a small number of orbitals and electrons, which suggests that constructing compact active spaces for surface reactions is feasible [631–634]. To construct such an active space we employ DDA and Natural orbitals as discussed in the previous section [250].

First, We localized the Kohn–Sham orbitals with the Pipek–Mezey procedure [635], using Mulliken populations formed from meta-Löwdin orthogonalized atomic orbitals [275,

636]. Localization was carried out independently for the occupied and virtual manifolds. To target charge rearrangements driven by the reaction, we evaluated the DFT-level density difference

$$\rho_{\mathrm{DD}}(\mathbf{r}) = \rho_{\mathrm{Li+O_2}}(\mathbf{r}) - \rho_{\mathrm{O_2}}(\mathbf{r}) - \rho_{\mathrm{Li}}(\mathbf{r}), \tag{5.1}$$

which highlights regions of electron accumulation and depletion (cf. Figure 20b).

Orbitals were ranked by their contributions to ρ_{DD}, using a threshold parameter η to suppress minor components and retain only physically meaningful features. A systematic scan identified a stable window near $\eta \approx 10^{-3}$ in which the ordering of occupied, localized orbitals is invariant; within this window we select the top-ranked occupied orbitals (and the corresponding virtual set) as the initial active space. This physics-aware protocol automates active-space definition from the mean-field description and enables controlled, systematic enlargement.

To further refine the space and ensure convergence with respect to size, we combined the density-difference ranking with natural orbitals from a CCSD calculation performed in the preliminary active space. The resulting natural occupations lie predominantly near 0 or 2, clearly separating high- and low-occupancy orbitals; when strong correlation is present, fractionally occupied orbitals are explicitly retained. The final active space is then built "inside out" around the Highest Occupied Natural Orbital (HONO) and Lowest Unoccupied Natural Orbital (LUNO).

In Figure 20 we can see the Active Space selection procedure, the electronic density for the chosen systems, and some orbitals that comprise the product's active space.

Classical Methods

Once the active-space orbitals were determined for the Γ-point, the rest were fixed and frozen at their HF levels, and the Born–Oppenheimer Hamiltonian was systematically reduced to the active space using established methods. We use the CCSD calculations of the active space Hamiltonian as an initial benchmark, as well as Complete Active Space Configuration Interaction (CASCI) and Heat-bath Configuration Interaction (HCI) calculations, to serve as comparisons for the quantum methods [285, 637].

In order to further refine the HCI results, we apply excitation operators to the final ground state according to a method we call "Extended Heat-bath Configuration Interaction" (Ext-HCI). The selection of which excitations to use can be adjusted to balance precision and efficiency, with higher-order excitations improving accuracy at an increased computational expense. Specifically we discard configurations in the wavefunction with amplitudes below a threshold of 10^{-2}. Furthermore, to keep the calculations manageable, we apply double excitation operators only to configurations which have amplitude larger than 10^{-1}, while only applying single excitations to the rest.

In Figure 21 we can see the classical methods results. Energy values start to converge at around 20 orbitals and continue to have small fluctuations up to 40 orbitals for HCI, Ext-HCI and CCSD. This suggests that an active space of 40 orbitals would be necessary to calculate the energy difference. Due to computational resource limitations, we cannot scale CASCI beyond 14 orbitals, but we can see reasonable agreement between the other three methods with errors within a few tens of mHa in this range. Ext-HCI, as a variational method, becomes the classical benchmark, and errors are calculated with

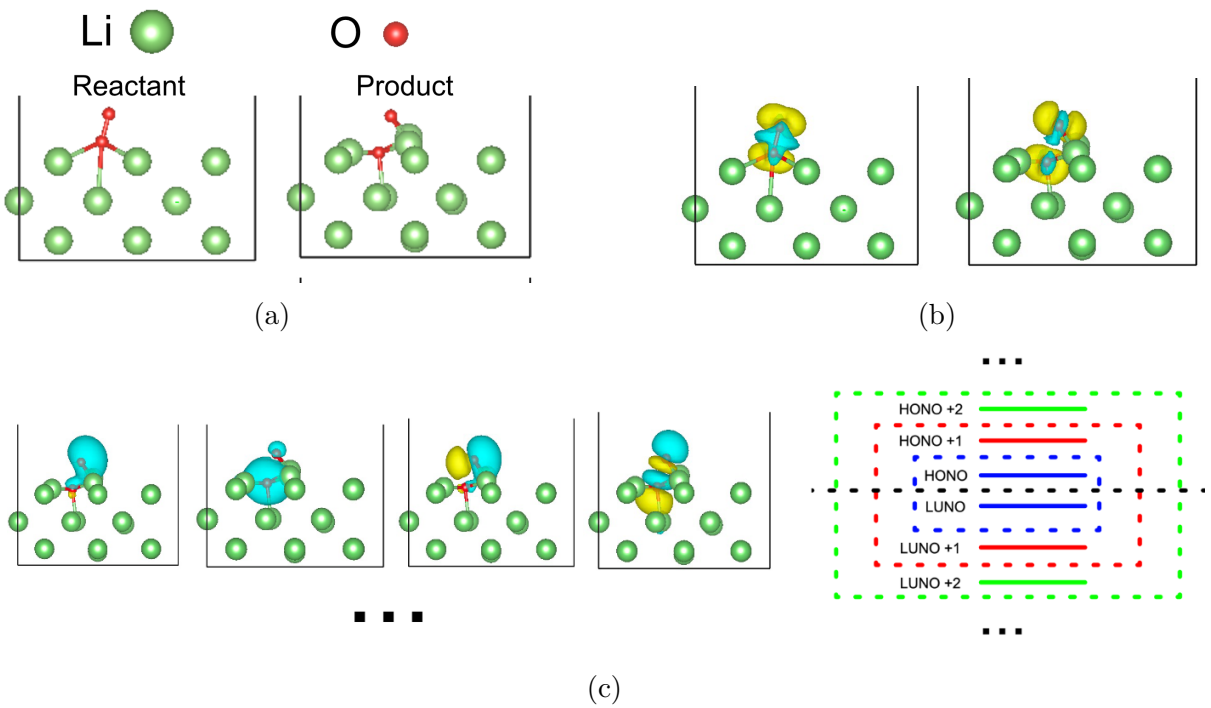


Figure 20 – **Active Space selection** We start with geometry optimization (a). Move to the Density difference analysis (b) and evaluate each orbital individual contribution to the density difference. Finally, we calculate the active space using the Natural Orbitals (c), where orbitals for the product are pictured. They were sorted based on their contribution to the Density difference. Natural Orbitals are calculated and the active space grows "inside out" from the HONO-LUNO level. *Reproduced from* M.A. Barroca, et al., [6] (DOI) CC BY 4.0 (link).

respect to it. It is relevant as well that the error increases as we add more orbitals to the active space, which is expected. As CCSD is not a variational method, it is possible that it is actually underestimating the actual ground state energies that we have from Ext-HCI. For this reason, errors are calculated with respect to Ext-HCI, the most accurate classical variational method utilized.

Quantum methods

As mentioned in Chapter 4 a critical differentiator between QSCI [601], and SQD [42] is the configuration recovery process, which ensures that sampled configurations maintain the correct electron count and spin sector. This is done iteratively, correcting errors by redistributing occupations in ways that align with the system's physical constraints. The ability to systematically expand the subspace with more configurations allows for controlled improvements in accuracy while maintaining manageable computational complexity. SQD calculations were conducted at around $\approx$ 5–10 iterations depending on the system size with 16 batches per iteration.

SQD offers a robust means of extracting ground-state properties while mitigating noise-related issues common in quantum simulations. The approach effectively clusters

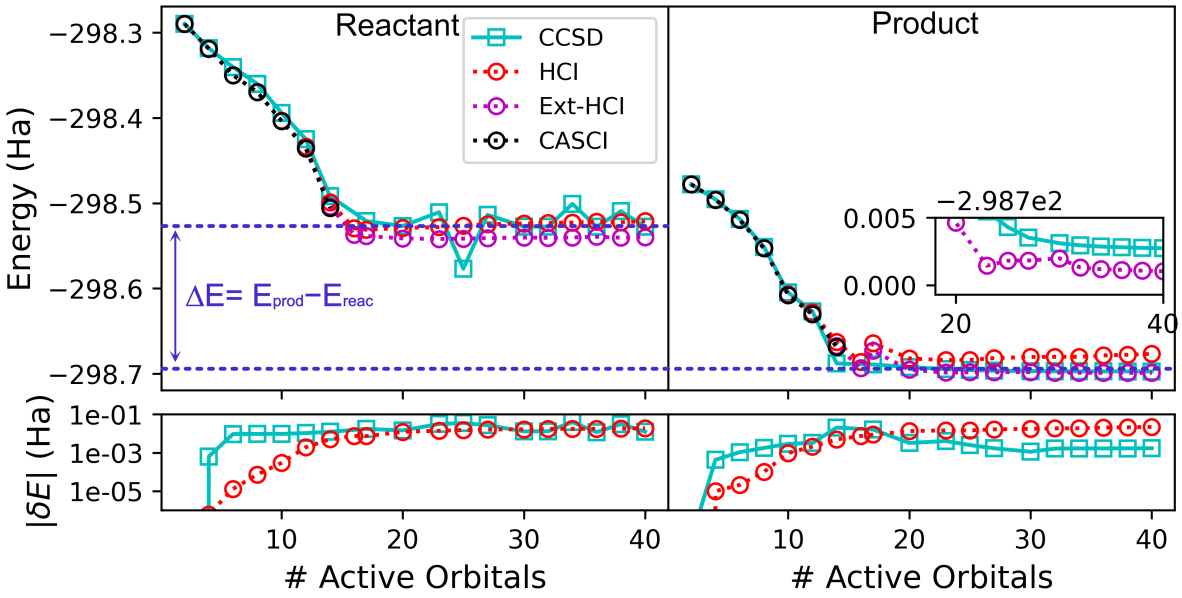


Figure 21 – **Comparison of ground state and reaction energy results for classical methods.** Classical simulations of the ground state energies as a function of active space size. The left curves are for the reactant and the right ones are for the product. *Reproduced from* M.A. Barroca, et al., [6] (DOI) CC BY 4.0 (link).

sampled configurations, refining noisy quantum data into physically meaningful representations, and provides a verifiable benchmark for comparing quantum simulations against standard Configuration Interaction (CI) methods.

To improve the quality of sampled configurations, we employ a truncated variation of the LUCJ ansatz [619]. Specifically, we adopt a two-layer circuit structure that omits the final orbital rotation and Jastrow interaction terms, reducing computational costs while maintaining accuracy and ensuring that sampled configurations do not completely cluster around the expected Hartree–Fock configuration. On top of that we initialize circuit parameters on the one and two body amplitudes form the CCSD benchmarks. This balance between efficiency and representational power ensures that the sampled configurations remain relevant for SQD, reducing errors in energy estimates [42]. Circuits were constructed using the `ffsim` software package [638].

In Figure 23, we show hardware-sampled histograms for the reactant in the (14e, 20o) active space, displaying only the ten most probable configurations (filled and unfilled circles denote occupied and unoccupied orbitals, respectively). In the non-truncated one-layer LUCJ circuit (Figure 23a), the Hartree–Fock determinant and its low-rank excitations dominate the distribution, consistent with prior observations [42]. By contrast, the truncated two-layer LUCJ circuit (Figure 23b), which omits the final orbital-rotation and Jastrow blocks, produces a broader, chemically relevant spread of configurations that is not HF-dominated [619]. Each histogram aggregates approximately 6×10^6 shots collected on `ibm_aachen`. Because the truncated two-layer variant yields higher-utility samples for SQD, we adopt it for all subsequent experiments.

For the quantum experiments, we also sample circuits on the `ibm_aachen` device, utilizing its heavy-hex topology in such a way that favors electronic-structure calculations

(22) [619]. When mapping circuits to the chip, we also take into account calibration data to avoid qubits with higher error rates using `mapomatic` [639]. Each circuit was sampled 6×10^6 times, and results were collected as counts dictionaries with bitstrings representing configurations, which are then passed to the SQD calculation.

High qubit count simulations include extra ancillae with respect to previous SQD work [42]. These were included after transpiling and mapping the circuit with `mapomatic`. This is especially necessary above 20 orbitals to improve sampling quality.

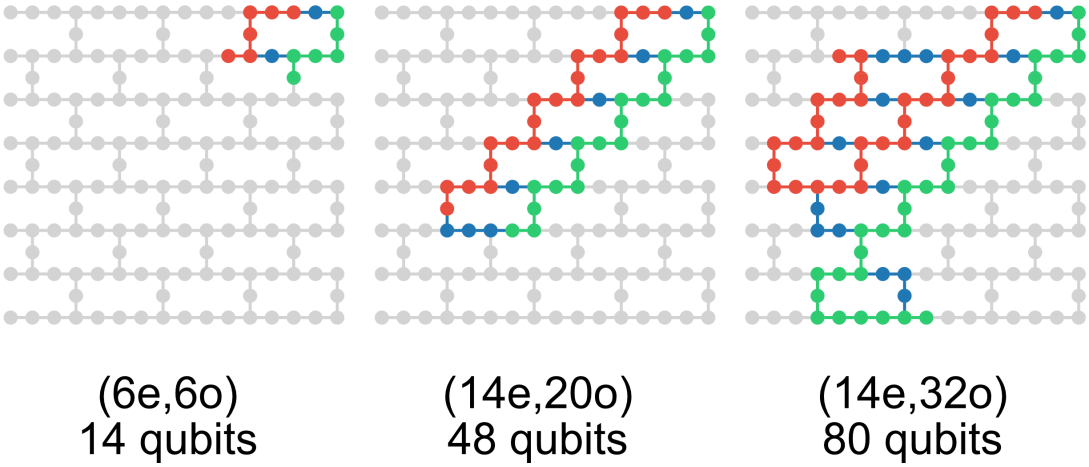
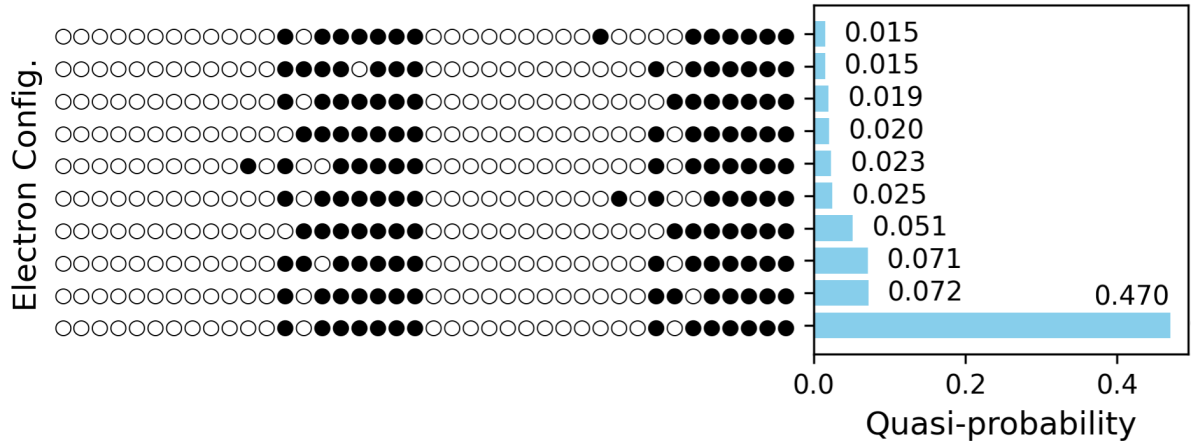


Figure 22 – **Circuit-to-chip mapping.** Green qubits represent α spin orbitals, red qubits represent β spin orbitals, while blue qubits act as ancillae. Numbers inside the qubits represent their index with respect to the device ordering. *Reproduced from* M.A. Barroca, et al., [6] (DOI) CC BY 4.0 (link).

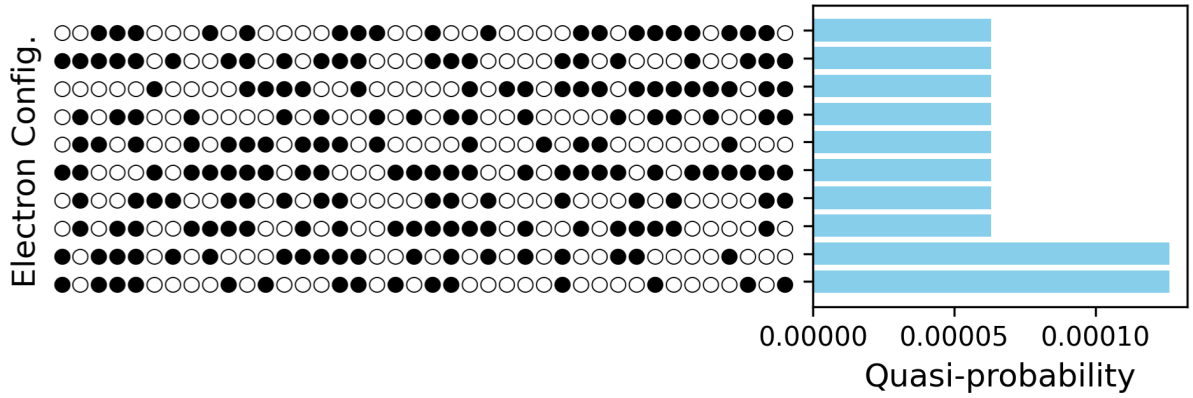
Table 9 – **Quantum metrics for select circuits.** q is the total number of qubits used, which is the sum of s qubits representing spin orbitals and a ancilla qubits. d is the 2-qubit circuit depth, CX is the number of CNOT gates and u the number of 1-qubit gates. Notice the small addition of two electrons in the product's active space makes circuits slightly larger after we include all occupied orbitals. *Reproduced from* M.A. Barroca, et al., [6] (DOI) CC BY 4.0 (link).

System	$q(s+a)$	d	CX	u
Prod(6e,6o)	14(12+2)	74	132	224
Prod(12e,12o)	30(24+6)	124	474	872
Prod(14e,20o)	48(40+8)	188	1228	2345
Prod(14e,32o)	80(64+16)	466	3026	5551
Reac(6e,6o)	14(12+2)	72	132	222
Reac(12e,12o)	30(24+6)	124	474	870
Reac(14e,20o)	48(40+8)	184	1194	2289
Reac(14e,32o)	80(64+16)	392	2874	5388

In Table 9 we see circuit sizes and qubit requirements for some of the circuits. Circuit size is described by depth and number of gates. When compared to previous work [42], we see our circuit requirements are similar for systems of comparable size. Bitstrings in the sampled distribution from `ibm_aachen` that have forbidden configurations with incorrect Hamming weights, which appear due to noise, are on average 93.57% of



(a)



(b)

Figure 23 – **Histograms of electron configurations sampled from the quantum hardware for the (14e,20o) reactant active space.** We display only the top-10 configurations. Filled (unfilled) circles represent occupied (unoccupied) orbitals. (a) Histogram for the non-truncated one-layer LUCJ ansatz. (b) Histogram for the truncated two-layer LUCJ ansatz. We observe that the truncated variant leads to a more relevant distribution of electronic configurations, that is not dominated by the reference Hartree–Fock state. *Reproduced from* M.A. Barroca, et al., [6] (DOI) CC BY 4.0 (link).

the samples. This is around 10 times better than a uniform configuration distribution which sits at an average 99.922% of forbidden configurations. At the maximum value of 32 orbitals, circuits get considerably more complicated, with 2935 CX gates we are very close to the limits of what can give usable results on hardware before noise starts to prominently affect the bitstrings. The percentage now sits at 96.54% for the reactant and 96.28% for the product, which worsens calculated energy values for SQD, ensuring results deteriorate for higher orbital counts. These results can seen on Figure 24. Fortunately, we can look beyond the standard SQD method with Ext-SQD, to enhance the accuracy of molecular ground-state energies.

Similarly to Ext-HCI, Ext-SQD refines accuracy by applying excitation operators

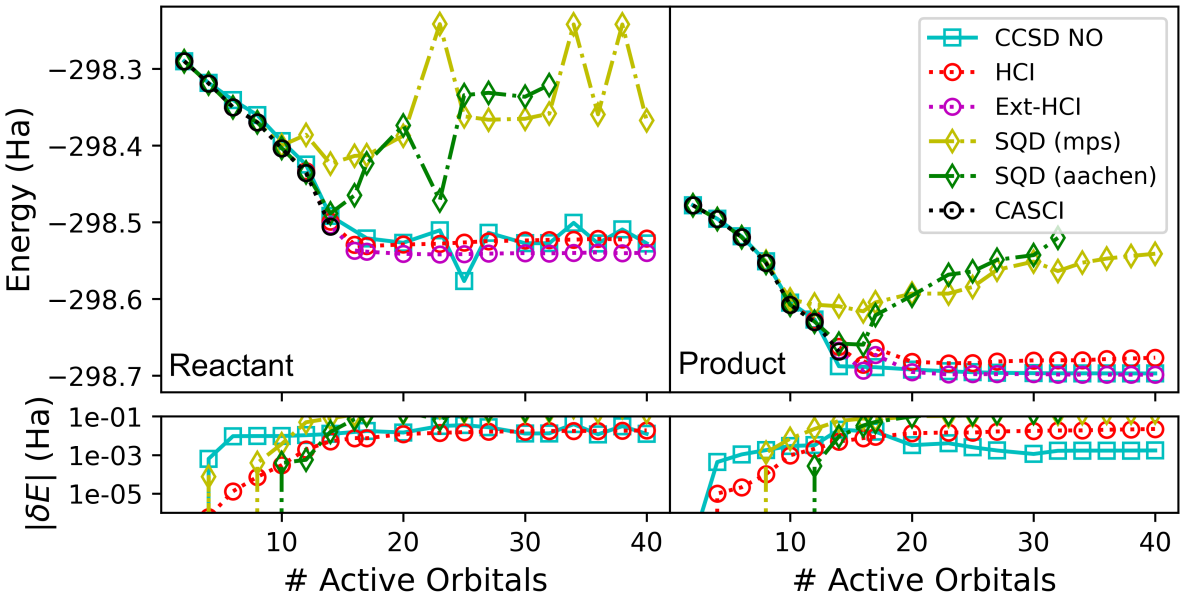


Figure 24 – **Comparison of ground state and reaction energy results for SQD methods.** SQD calculations of the ground state energies as a function of active space size. The left curves are for the reactant and the right ones are for the product. *Reproduced from* M.A. Barroca, et al., [6] (DOI) CC BY 4.0 (link).

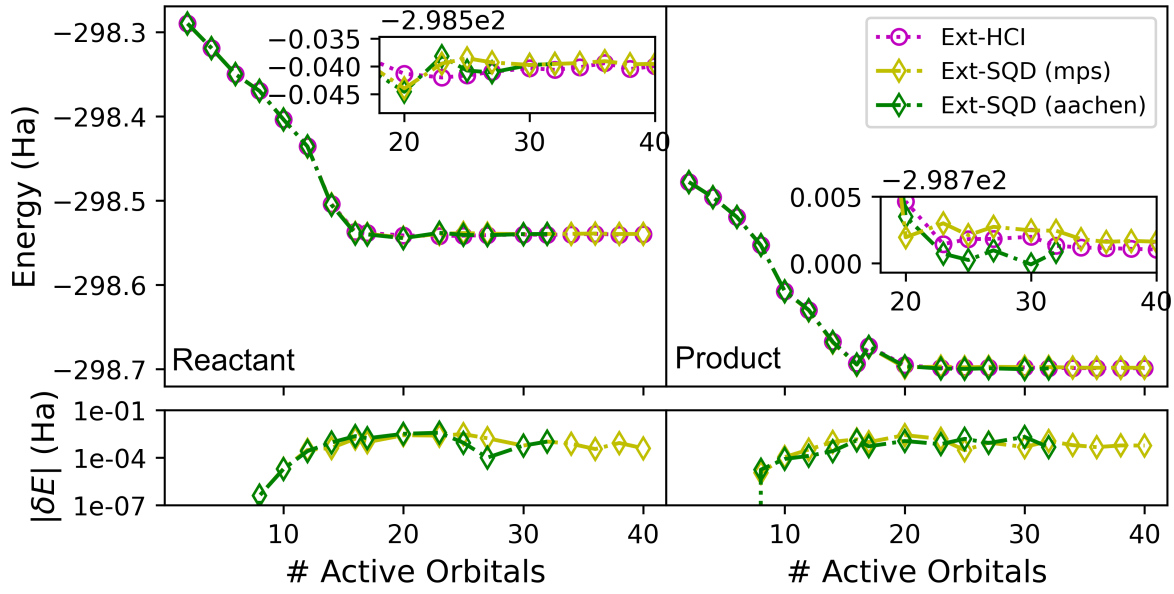
to sampled configurations [43]. Before applying these excitations, configurations in the wavefunction with amplitudes below a threshold of 10^{-2} are discarded. In order to have a fair comparison, we again discard configurations in the wavefunction with amplitudes below a threshold of 10^{-2}, and apply double excitation operators only to configurations that have amplitudes larger than 10^{-1}, while only including single excitations for the rest. Ext-SQD calculations do not rely on configuration recovery and simply start from the results of the previous SQD experiment; hence, only one iteration is necessary. We keep the number of batches at 16.

Despite not providing accurate results, SQD is an integral step before starting the Ext-SQD calculation. Those, in turn, provide much better results, staying consistently close to Ext-HCI (below the 1.6 mHa chemical accuracy threshold) when compared to standard HCI and CCSD, as seen in Figure 25a. Ext-SQD also reaches lower energy values for both the product and reactant beyond 14 orbitals. Unfortunately, scaling Ext-SQD on hardware beyond 32 orbitals proved challenging, as the quality of sampled configurations deteriorated. At 34 orbitals and beyond, we were unable to obtain any configuration with the correct Hamming weights among the 6×10^6 samples.

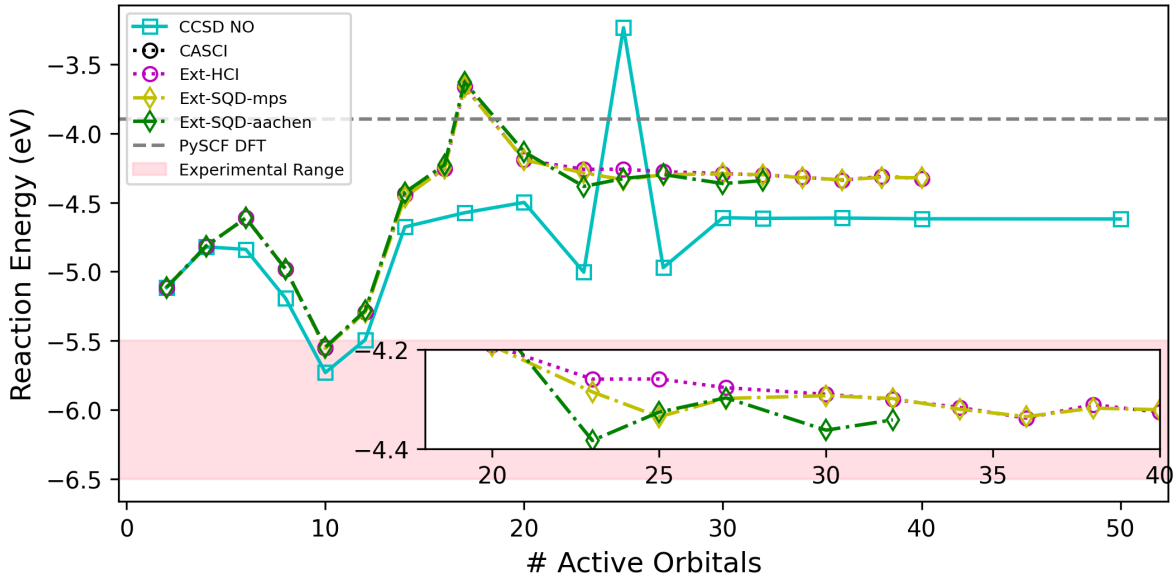
At this scale, there is not much one can do to improve results as the topology of the device limits the mapping of circuits to the diagonals displayed in Figure 22. With the increase in the number of orbitals (and, therefore, qubits) it becomes harder and harder to design circuits that avoids qubits and gates with high error rates, which ends up destroying the quality of the samples.

In order to still present SQD and Ext-SQD results up to 40 orbitals, we opted to approximate the LUCJ quantum circuits with tensor networks using a truncation value of 10^{-8} and sample utilizing Matrix Product State (MPS) simulators [640, 641]. These,

while not perfect, will at least provide us with a reasonable expectation of results that we would have had if noise were lower. These results can also be seen on Figure 25a labeled accordingly, where they scale up to the required 40 orbitals but give slightly higher energy values due to the truncation.



(a)



(b)

Figure 25 – **Comparison of ground state and reaction energy results for quantum experiments** (a) Simulated (yellow diamonds) and experimental (green diamonds) Ext-SQD results compared to Ext-HCI, the best classical method considered. Experimental Ext-SQD scales up to 32 orbitals, only. (b) Reaction energies for quantum and classical methods computed at the Γ-point. DFT results are shown as a reference but are known to be overestimated. *Reproduced from* M.A. Barroca, et al., [6] (DOI) CC BY 4.0 (link).

These calculations were conducted on a HPC nodes consiting of 192 x86 cores and 3 TB of available RAM utilizing Dice [642] as the diagonalizer. Note on Table 10 that as the number of orbitals increases, the subspace dimension for Ext-SQD grows faster than for standard SQD, and while it is consistently smaller than Ext-HCI, it is still in the same order of magnitude. That is because not only does the number of single excitations grow quadratically with the number of orbitals, but double excitations grow with the fourth power. This makes Ext-SQD challenging at higher orbital counts and severely limits the number of double excitations that can be included [43].

On Figure 25b we plot the reaction energy up to 50 orbitals in the active space and provide the DFT value for reference. Reaction energy values for select active spaces are made available in Table 11 We can see HCI, Ext-SQD, and CCSD in this active space provide lower values, which are more in line with those found in literature, usually rely on empirical corrections [626–628, 643]. We were unable to provide results up to 50 orbitals for all methods again due to a bottleneck on classical and quantum compute resources, but we can already see energy values starting to converge for both Ext-HCI and Ext-SQD much earlier than CCSD. While CCSD seems to provide a ΔE value closer to what is expected from experiments, it does so while overestimating reactant energy, so it cannot be taken at face value.

It is worth noting that while CCSD is not a variational method and can provide energy values below the actual ground state [541, 644], the same is not true for variational methods such as CASCI, HCI, Ext-HCI and both SQD and Ext-SQD which are based upon CI methods [601, 640]. These methods can only find upper bounds for the actual ground state energy for a given electronic Hamiltonian, which means lower energy values should always be more accurate [541].

Another relevant point is that atomic coordinates for both systems were derived from DFT relaxation, which means that all subsequent methods used for energy calculations will be affected by the mean-field nature of DFT in defining the reactant and product states. Despite that, we can see that all methods still improve the reaction energy value, getting closer to the experimental range.

5.1.3 Discussion and Summary

The results presented here demonstrate that a carefully engineered hybrid quantum–classical workflow (cf. Figure 19) can deliver competitive electronic-structure predictions for a prototypical surface reaction relevant to battery materials. Systematic active-space construction (Pipek–Mezey localization, DDA, and refinement with coupled-cluster natural orbitals) yields compact yet descriptive orbital subspaces for both reactant and product. Within those spaces, LUCJ circuits sampled on `ibm_aachen`, together with SQD and EXT-SQD, produce ground-state energies and reaction energetics that compare favorably to classical references (HCI, Ext-HCI, CCSD), with close agreement to CASCI at the smallest sizes (Figure 25, Figure 11). In particular, the truncated two-layer LUCJ ansatz generates a broader, chemically meaningful configuration distribution than the one-layer variant, improving SQD sample quality and downstream accuracy [42, 619].

Across the sizes studied, Ext-SQD remains consistently close or lower than Ext-HCI, often within chemical accuracy (< 1.6 mHa), while using subspaces of comparable

Table 10 – **Select Ext-SQD and Ext-HCI calculations for product and reactant.** $|\tilde{\chi}|$ is the number of samples taken from **ibm_aachen**, K is the number of configuration recovery iterations used in SQD, D is the dimension of the subspace being diagonalized by SQD, D_E is the dimension of the subspace being diagonalized by Ext-SQD, while D_{EH} is the dimension of the subspace being diagonalized by Ext-HCI. *Reproduced from* M.A. Barroca, et al., [6] (DOI) CC BY 4.0 (link).

System	$\|\tilde{\chi}\|$	K	D	D_E	D_{EH}
Prod(2e,2o)	$1.5 \cdot 10^6$	5	$2.7 \cdot 10^4$	$1.5 \cdot 10^4$	$1.5 \cdot 10^4$
Prod(12e,12o)	$1.5 \cdot 10^6$	7	$3.4 \cdot 10^5$	$2 \cdot 10^5$	$5.3 \cdot 10^5$
Prod(14e,20o)	$2 \cdot 10^6$	10	$1.4 \cdot 10^6$	$7.9 \cdot 10^6$	$9.6 \cdot 10^6$
Prod(14e,30o)	$2 \cdot 10^6$	10	$1.4 \cdot 10^6$	$33 \cdot 10^6$	$84 \cdot 10^6$
Reac(2e,2o)	$1.5 \cdot 10^6$	5	$2.9 \cdot 10^4$	$1.2 \cdot 10^4$	$1.2 \cdot 10^4$
Reac(12e,12o)	$1.5 \cdot 10^6$	7	$3.4 \cdot 10^5$	$1.8 \cdot 10^5$	$4.7 \cdot 10^5$
Reac(14e,20o)	$2 \cdot 10^6$	10	$1.3 \cdot 10^6$	$9 \cdot 10^6$	$11 \cdot 10^6$
Reac(14e,30o)	$2 \cdot 10^6$	10	$1.3 \cdot 10^6$	$28 \cdot 10^6$	$56 \cdot 10^6$

Table 11 – **Reaction Energies.** ΔE (in eV) compared across methods. *Reproduced from* M.A. Barroca, et al., [6] (DOI) CC BY 4.0 (link).

Act. Space	**CASCI**	**CCSD**	**Ext-HCI**	**Ext-SQD (MPS)**	**Ext-SQD (aachen)**
(2e,2o)	−5.116	−5.116	−5.116	−5.116	−5.116
(12e,12o)	−5.294	−5.494	−5.294	−5.294	−5.283
(14e,20o)	N/A	−4.500	−4.191	−4.193	−4.134
(14e,30o)	N/A	−4.610	−4.290	−4.293	−4.363
(14e,32o)	N/A	−4.615	−4.301	−4.299	−4.342

order (Tabela 10). The same trend is reflected at the level of reaction energies (Figura 25b, Tabela 11). Because HCI/Ext-HCI, CASCI, and SQD/Ext-SQD are variational CI-based methods, their energies form rigorous upper bounds to the true ground state, whereas CCSD can, in principle, undershoot [541, 644]. This explains cases where CCSD appears closer to empirical reaction energies while simultaneously over-stabilizing one state (typically the reactant), a behavior we observe in Table 11.

On current devices, performance is ultimately limited by circuit depth, entangling-gate counts, and noise accumulation. As the number of orbitals grows, the resources in Table 9 (depth, CX count) increase sharply, and the fraction of hardware samples with incorrect Hamming weights rises (into the $\sim 96\%$ range at 32 orbitals). SQD's configuration-recovery step helps sanitize noisy bitstrings into a physically valid subspace, clustering samples into meaningful determinants [42]. Nevertheless, beyond ∼32 orbitals on hardware, sample quality degrades and Ext-SQD falters; we therefore complemented hardware runs with tensor-network (MPS) sampling of the same LUCJ circuits to estimate the large-space trend (Figura 25a), acknowledging the small upward bias from MPS truncation [640, 641].

A key algorithmic finding is the trade-off between quantum sampling and classical

post-processing. Ext-SQD reliably improves on SQD by injecting targeted single and double excitations into the sampled subspace, but the extended diagonalizations scale rapidly with orbital count (singles $\sim O(n^2)$, doubles $\sim O(n^4)$), shifting the bottleneck to classical linear algebra (Tabela 10). This mirrors the classical HCI $\rightarrow$ Ext-HCI behavior and suggests that parallel eigensolvers, low-rank factorization, or Krylov/filtered-subspace strategies could be decisive for pushing beyond 30–40 orbitals.

Our comparison with traditional methods underscores the value of the hybrid route. Ext-SQD (experimental and MPS) reaches lower variational energies than HCI and converges the reaction energy earlier than CCSD at comparable active-space sizes (Figura 25). Importantly, while our DFT baseline at the Γ-point underestimates the exothermicity, [626–628, 643] hybrid results systematically correct toward literature ranges even with modest quantum resources. We note, however, that the underlying geometries originate from DFT relaxations and that all active-space energies inherit this structural context.

A natural next step is to recover dynamic correlation outside the active space via second-order perturbation theory. CASPT2 and NEVPT2 were designed precisely for this purpose, [278, 645] with CASPT2 typically stabilized by level shifts (real/imaginary) and the IPEA shift to mitigate intruders, [646, 647] while NEVPT2 is intruder-free and size-consistent in contracted form [278]. Algorithmically, Ext-SQD yields multi-reference wavefunctions that can serve as a PT2 reference over the external orbital space, in direct analogy with SCI+PT approaches [648, 649]. Although we do not tabulate PT2 numbers for our largest embedded systems, small-scale tests indicate that absolute energies drop while reaction-energy corrections partly cancel between states; the net benefit depends on differential external correlation. Implementing Ext-SQD+PT2 (and NEVPT2 atop Ext-SQD) is therefore a promising route to sub-chemical-accuracy thermochemistry without increasing quantum circuit depth.

In closing, this chapter establishes a practical blueprint for quantum-assisted surface chemistry: physics-informed active spaces; shallow, structure-preserving ansätze; sample-based diagonalization; and classical augmentation when needed. While hardware limitations remain (noise, depth, limited connectivity), the demonstrated agreement with strong classical baselines, together with clear paths to capture external correlation, points to a viable near-term role for quantum resources in materials discovery.

In closing, this chapter establishes a practical blueprint for quantum-assisted surface chemistry: physics-informed active spaces; shallow, structure-preserving ansätze; sample-based diagonalization; and classical augmentation when needed. While hardware limitations remain (noise, depth, limited connectivity), the demonstrated agreement with strong classical baselines, together with clear paths to capture external correlation, points to a viable near-term role for quantum resources in materials discovery.

The next chapter instantiates the same program in a translationally symmetric setting: We deploy SQD within a periodic embedding workflow to estimate band gaps in bulk semiconductors. Short-time LUCJ states are prepared, sampled, and diagonalized in symmetry-adapted subspaces that conserve crystal momentum and spin. With minimal mitigation, the resulting gaps align with classical references, illustrating that sampling–driven subspace solvers can operate directly on crystalline systems while retaining efficiency and accuracy.

Together, Chapters 5 and 6 test the sampling methodology under two complementary stressors—heterogeneous, strongly inhomogeneous interfacial chemistry versus delocalized, translationally symmetric bulk electronic structure—using the same design levers (configuration recovery, excitation pools, and measurement budgets) and quantifying accuracy against trusted classical references. This sets the stage for Chapter 6, where we detail the periodic setup, hamiltonian modelling, and benchmarking protocol.

6 Band Gap Estimation in Periodic Materials via Sample-Based Methods

6.1 Adapted Articles: Band–Gap Prediction via Sample-Based Quantum Diagonalization and lattice modeling of metal-organic frameworks for quantum computers

As we did in Chapter 5, we again carry forward the measure–then–diagonalize paradigm and the configuration-recovery/mitigation toolkit from Chapter 4. Here we merge two complementary strands to extend that roadmap from molecules and reactions to materials: (i) a worked recipe for computing crystalline band gaps with SQD, and (ii) a curated lattice-model resource for MOFs that supplies transferable tight-binding and extended-Hubbard Hamiltonians. For periodic band-gap prediction, we focus on compact binary oxides—ZrO_2 and HfO_2—rather than full MOF crystals, which remain beyond near-term qubit/shot budgets; shallow LUCJ trial states seed SQD subspaces and measurement-defined subspaces are classically diagonalized to extract gaps from $\sim$38–46-qubit Hamiltonians derived by projecting DFT densities onto localized bases and refining interactions. Accuracy and cost are tied to explicit design choices (basis/localization strategy, interaction construction, excitation pools, mitigation and configuration recovery, and sampling budgets) yielding a reproducible template for crystalline systems. In parallel, the TBHubbard collection furnishes tight-binding models for a large MOF corpus and extended-Hubbard instances (self-consistent U and V) for a transition-metal subset—including Zr- and Hf-containing frameworks—together with descriptors amenable to ML and quantum-algorithm benchmarking. In this chapter we use TBHubbard not as a direct SQD target, but as a standardized source of lattice Hamiltonians, descriptors, and periodic-embedding testbeds to stress-test SQD/LUCJ choices across chemically diverse motifs. This separation clarifies scope: the SQD case study treats representative crystalline oxides end-to-end today, while TBHubbard underpins benchmarking and paves a data-ready path to full-framework studies as hardware scales.

Provenance and license. This chapter adapts material from two CC-licensed publications on which the candidate is a co-author:

- A. Duriez, P. C. Carvalho, M. A. Barroca, F. Zipoli, B. Jaderberg, R. N. B. Ferreira, K. Sharma, A. Mezzacapo, B. Wunsch, and M. Steiner, *Computing band gaps of periodic materials via sample-based quantum diagonalization*, arXiv:2503.10901 (2025), DOI: 10.48550/arXiv.2503.10901. (Submitted to peer-review)
- P. C. Carvalho, F. Zipoli, A. C. Duriez, M. A. Barroca, R. N. B. Ferreira, B. Jones, B. Wunsch, and M. Steiner, *TBHubbard: tight-binding and extended Hubbard model database for metal-organic frameworks*, *Scientific Data* **12**, 1776 (2025), DOI: 10.1038/s41597-025-06054-w; dataset DOI: 10.7910/DVN/ZKLRLF.

The publications are released under a Creative Commons Attribution (CC BY 4.0) license; this adaptation complies with those terms (link). Text is independently rewritten; figures, tables, and numerical results are reproduced from the source with full attribution.

6.1.1 Motivation

Predictive simulation has become a central pillar of modern materials discovery and design [1], offering a scalable complement to costly or time-consuming experiments. For electronic-structure problems, the dominant workhorse is KS-DFT [22, 23], which replaces the many-body Hamiltonian with an auxiliary mean-field problem. As we discussed in Chapter~3, KS-DFT captures broad chemical trends, from metals to semiconductors to insulators, with remarkable efficiency [650]. Yet its mean-field nature also sets limits: families with localized open-shell d or f electrons often exhibit phenomena (e.g., Mott insulating behavior, unconventional superconductivity) that are poorly described at the semilocal level [651–655].

A canonical symptom is the band-gap underestimation. The fundamental gap $E_{gap} = E(N_e+1) + E(N_e-1) - 2E(N_e)$ [651, 656] differs from the KS eigenvalue gap because of the derivative discontinuity in the exchange–correlation potential [651, 656]. Various remedies exist—Wannier-function approaches [657], hybrid functionals [658, 659], and many-body perturbation theory (GW) [660]—each trading accuracy against cost. DFT+U+V, motivated by the Extended Hubbard model [661, 662], offers a pragmatic route by adding screened on-site (U) and inter-site (V) interactions. Importantly, recent linear-response/DFPT implementations compute U, V self-consistently rather than fitting them [662, 663]. Even so, outcomes can depend on projector choices and manifolds, and corrected gaps can remain smaller than experiment in challenging cases [664].

An alternative viewpoint models electrons directly with fermionic lattice Hamiltonians (e.g., tight-binding plus short-range interactions), a language that has illuminated systems from polyacetylene and organic conductors to twisted bilayer graphene and cuprates [665–668]. Exact solutions, however, are restricted to small instances [669]; approximate solvers face their own barriers: Quantum Monte Carlo suffers from sign problems for generic fermions [670–672], and tensor-network methods, while powerful in 1D and some 2D cases [264, 669, 673], confront steep scaling for 3D and wide entanglement [674].

These challenges motivate quantum algorithms tailored to materials simulation [1]. Early demonstrations with variational methods show promise for simple crystals and model systems [675, 676], and theory continues to chart conditions for quantum advantage [677]. Present-day devices, however, are noise-limited and resource-constrained [678], calling for hybrid strategies that minimize depth while retaining many-body fidelity. As we saw previsouly, SQD follows this philosophy: a shallow, physics-informed circuit generates a distribution over configurations; one then *classically* diagonalizes the Hamiltonian in the sampled subspace, augmented by configuration-recovery techniques to counteract noise [42, 43, 601].

We combine a material-informed lattice construction with data assets that enable transfer across chemistries. First, following a DFT→lattice workflow, we project KS states onto localized pseudo-atomic orbitals to obtain tight-binding (TB) hoppings and then append self-consistent U and V terms (DFT+U+V) to arrive at an Extended-

Hubbard (EH) Hamiltonian suited for many-body solvers. Second, we build and leverage *TBHubbard*—a CC-licensed dataset of MOF tight-binding models and a curated subset with self-consistent U, V—to motivate why the EH description is the minimal extension of TB needed to capture metal–ligand correlation trends relevant to band edges in porous frameworks. While full MOF unit cells remain out of reach for near-term quantum hardware, the same chemistry arises in compact binary oxides; accordingly, we validate the approach on representative ZrO_2 and HfO_2, where KS-DFT underestimates the gap but a lattice+many-body treatment can recover experimental/theoretical references.

(i) We formalize the DFT-projected Extended-Hubbard Hamiltonian used throughout and relate its fundamental-gap definition to many-body ground states. (ii) We explain, using TBHubbard statistics and examples, why including both U (metal d) and V (metal–ligand) is essential for MOF-like chemistries. (iii) We show how shallow LUCJ circuits feed SQD/Ext-SQD to estimate gaps from $\sim$38–46 spin-orbital Hamiltonians with minimal mitigation, positioning the workflow as a practical stepping stone from crystalline oxides to larger framework fragments. The remainder of the chapter introduces the lattice Hamiltonian (§6.1.2), justifies the EH model in the MOF context using TBHubbard (§6.1.3), and then deploys SQD to obtain band-gap predictions and resource/accuracy trade-offs in periodic materials(§6.1.5).

6.1.2 From DFT to a Material-Informed Lattice Hamiltonian for Crystals

Our starting point for periodic materials is an *ab initio*, material-informed lattice model that lifts the Kohn–Sham (KS) band structure into second quantization on a localized basis and then augments it with short-range interactions. Concretely, we (i) project the KS electronic structure onto pseudo-atomic orbitals to obtain a TB Hamiltonian and (ii) add screened on-site and near-neighbor density–density terms computed self-consistently via DFT+U+V linear response. The result is an EH Hamiltonian in which each site corresponds to a localized atomic orbital in the DFT pseudopotential basis (orbitals from the same atom become distinct sites). This construction preserves material specificity (through the TB hoppings) while capturing leading correlation effects (through U and V) [662, 663, 679–683].

At a fixed $\mathbf{k}$ (we use Γ in this work, but the construction is general), the lattice Hamiltonian reads

$$\hat{H} = \sum_{pq,\sigma} t_{pq}\, \hat{a}^{\dagger}_{p\sigma}\hat{a}_{q\sigma} + \sum_{p} U_p\, \hat{n}_{p\uparrow}\hat{n}_{p\downarrow} + \sum_{pq,\sigma\tau} V_{pq}\, \hat{n}_{p\sigma}\hat{n}_{q\tau}, \tag{6.1}$$

where t_{pq} come from the TB projection (matching the KS bands on the chosen grid), U_p penalizes double occupancy on localized d (or f) shells, and V_{pq} captures short-range metal–ligand density coupling. In practice, manifold choices such as U_d on the transition-metal and $V_{d\text{–}p}$ to nearest-neighbor ligand p orbitals are determined by the chemistry and by the DFPT projector setup [662, 663].

Given $\hat{H}$ in Equation 6.1, the direct band gap at the chosen $\mathbf{k}$ can be computed from ground-state energies in fixed-number sectors,

$$E_g \;=\; E(N_e+1) \;+\; E(N_e-1) \;-\; 2\,E(N_e), \tag{6.2}$$

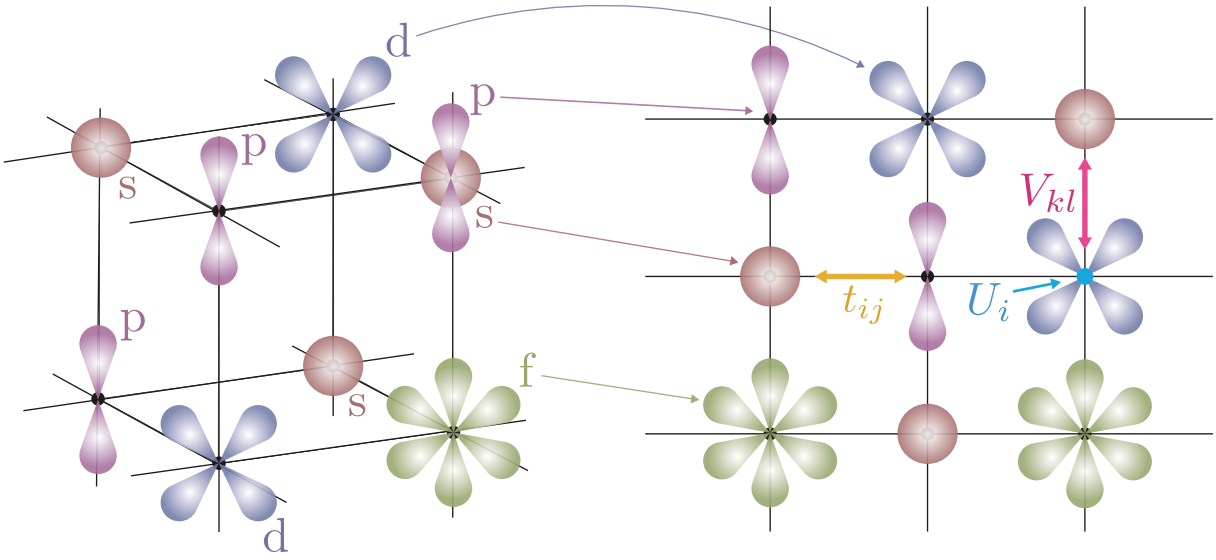


Figure 26 – **DFT→lattice workflow.** KS densities are projected to a localized AO basis to yield TB hoppings t_{pq}; intra-site U_p and inter-site V_{pq} are obtained self-consistently within DFT+U+V. Each AO becomes a lattice site; orbitals on the same atom map to different sites. The final model is the Extended-Hubbard Hamiltonian in Equation 6.1. *Reproduced from* Duriez et al. [7] (DOI), CC BY 4.0 (link).

with N_e the number of valence electrons defined by the pseudopotential basis [684]. This definition mirrors the fundamental gap and avoids known pitfalls of using KS eigenvalue differences in semilocal DFT.

The projection preserves the KS bandstructure by construction (within the chosen basis). Figure 27 demonstrates this for representative oxides: the PAOFLOW-projected TB bands (solid) accurately reproduce the DFT bands (dashed) near the Fermi level, while the corresponding TB matrices at Γ expose chemically interpretable sparsity/locality patterns.

Placing this lattice construction upfront clarifies scope and interfaces for what follows: TBHubbard (next section) provides datasets of t, U, V, and descriptors across MOFs; the periodic SQD study shows how Equation 6.1 plus many-body solvers yields band-gap predictions in chemically realistic oxides where full MOF unit cells are still beyond near-term qubit budgets. With the Hamiltonian and its provenance established here, §6.1.3 can focus on *why* EH is the minimal, effective extension of TB for MOFs and how TBHubbard feeds classical and quantum workflows.

6.1.3 The TBHubbard database

MOFs combine transition–metal clusters with conjugated organic linkers, yielding localized d–manifolds that hybridize with ligand p (and sometimes s) states. A TB Hamiltonian captures orbital topology and near–neighbor couplings, but it misses the leading on–site and short–range correlation effects that control gap formation and charge localization in many transition metal–containing frameworks. The EH model augments TB with orbital–resolved interactions: an intra–site term U that penalizes double occupation

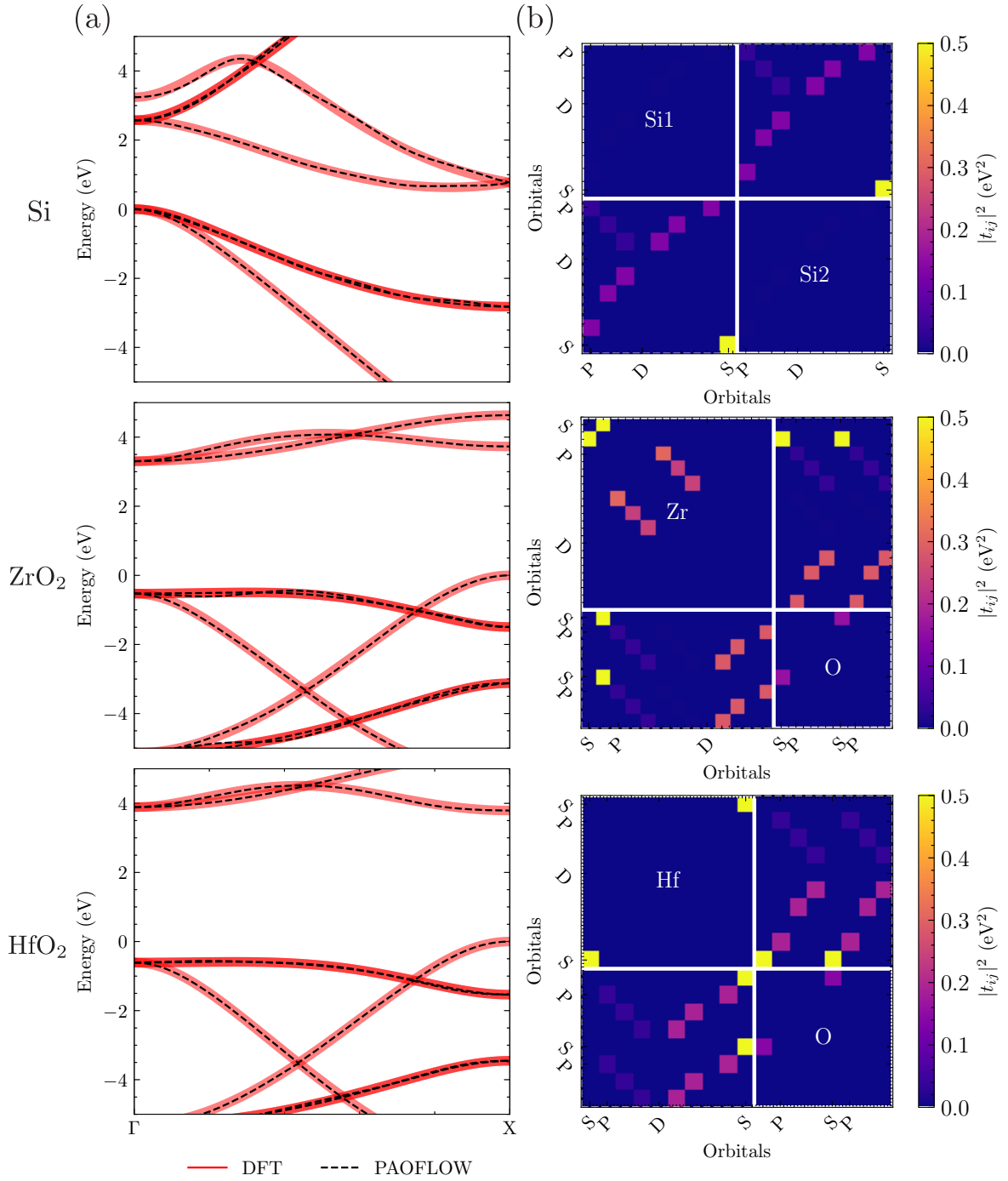


Figure 27 – Validation of the TB projection and resulting sparsity. **Left:** DFT (dashed) vs. PAOFLOW TB (solid) bandstructures for Si, ZrO_2, and HfO_2. **Right:** False-color $|t_{ij}|^2$ at Γ showing block structure (intra-site vs. inter-site) and dominant near-neighbor couplings. *Reproduced from* Duriez et al. [7] (DOI), CC BY 4.0 (link).

on the transition–metal d shell and an inter–site term V that encodes metal–ligand density–density coupling across bonds [662]. In the DFT context, these same physics appear as DFT+U and DFT+U+V, where U and V can be obtained from first principles via linear–response / DFPT rather than empirical fitting [663].

TBHubbard provides, for 10,435 MOFs, TB Hamiltonians (via PAOFLOW [679,

680]) and descriptors; for a 242–MOF subset it further supplies self–consistent U and V values computed in Quantum ESPRESSO [681–683]. The dataset thus pairs (structure, TB couplings) with (U, V) that are specific to a metal–ligand motif. Figure 28 (panels c–d of the source figure) illustrates how including U and V can shift band gaps relative to semilocal DFT, with the largest systematic corrections arising in frameworks hosting late–row d shells, while Zr/Hf–based nets (dominant in the EH subset by design) remain comparatively less affected—consistent with their more ionic bonding and less localized d charge. This pattern supports the modeling choice: TB captures connectivity/hybridization and EH supplies the minimal correlated correction that regularizes band–gap trends across diverse metals [651].

Beyond the familiar U correction to metal d states, the inter–site V term is essential for MOFs because the frontier density is often shared along M–X (e.g., M–O, M–N) bonds. The TBHubbard statistics show broad, metal–dependent spreads in V when the perturbed manifold involves ligand p states, and smaller (often slightly negative, by construction of the response kernel) values with s–manifold perturbations; see panels (a–b) in Figure 28. Physically, a nonzero V renormalizes effective M–X hopping and screens on–site U, improving band–edge placement in mixed–covalent motifs typical of conductive/catalytic MOFs.

In practice, TB+U+V (or EH) supplies a compact, transferable surrogate: start from TB matrix elements at Γ (or across a coarse k–mesh) and correct the metal–centered blocks by U, while renormalizing M–X couplings by V. TBHubbard provides precisely these ingredients at scale, enabling rapid screening or calibration of band–gap models before expensive hybrid–DFT/MBPT runs. The database's emphasis on Zr/Hf (UiO–like chemistry) ensures coverage of technologically relevant, hydrothermally stable families while still including late–row metals for which correlation effects are strongest.

For MOFs, an EH description is the smallest extension of TB that (i) honors metal–centered correlation and (ii) respects metal–ligand hybridization. TBHubbard empirically quantifies both, making it a principled foundation for band–gap modeling.

6.1.4 From Descriptors to Hamiltonians: How TBHubbard Feeds Band–Gap Prediction and Quantum Workflows

TBHubbard couples three assets: (i) TB Hamiltonians from PAOFLOW, (ii) self–consistent U, V for a curated EH subset, and (iii) aligned structural descriptors (Smooth Overlap of Atomic Positions (SOAP)) and TB–block embeddings that encode local chemistry and electronic couplings [679,680,685,685]. The dataset spans 10k+ MOFs (TB) with a focused 242–MOF EH core rich in Zr/Hf chemistry. Figure 29 visualizes the split and shows that the EH subset occupies a chemically meaningful but narrower manifold within the broader TB set—useful for model transfer and uncertainty estimation.

For fast screening, one can (i) assemble a TB Hamiltonian from the dataset, (ii) apply EH corrections with tabulated U, V (or ML–predicted surrogates learned from TB–block embeddings), and (iii) compute gaps on a coarse k grid. Because U, V are manifold–specific (e.g., d on the metal and p on the nearest ligand), the correction is local, scalable, and explainable. The empirical trends in Figure 28 guide when EH corrections

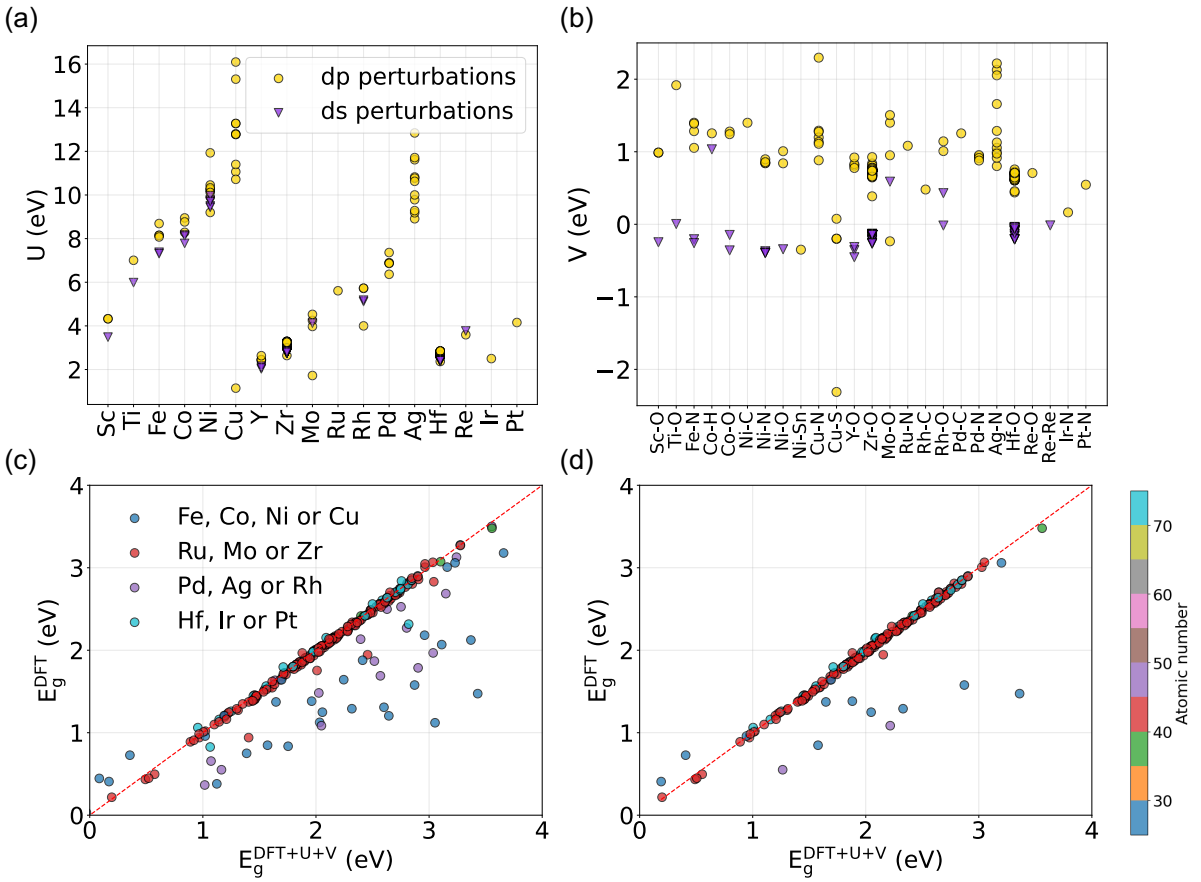


Figure 28 – **Hubbard parameters and band–gap shifts in the TBHubbard EH subset. (a–b)** Distributions of intra–site U and inter–site V across metals for $d-p$ and $d-s$ manifolds. **(c–d)** Scatter of E_g^{DFT} vs. $E_g^{\mathrm{DFT+U+V}}$ for $d-p$ and $d-s$ perturbations, respectively. Many Zr/Hf frameworks lie near the diagonal (small corrections), whereas late–row d systems show larger positive shifts, consistent with correlation–driven gap opening. Reproduced from Carvalho et al. [8] (DOI), CC BY 4.0 (link).

matter most (late–row d shells, stronger covalency).

Near–term quantum algorithms benefit from compact, structured models. EH Hamiltonians derived from TBHubbard satisfy three constraints that are crucial for quantum experiments: (i) *sparsity and locality* (dominant near–neighbor couplings in the TB matrix), (ii) *few–orbital, symmetry–aware active spaces* centered on band–edge states (which TB embeddings/SOAP help identify), and (iii) *data–driven parameterization* via U, V to avoid ad hoc fitting. Concretely:

- The TB matrix structure (see Figure 30) exposes block–localized orbital groups that map naturally to small active spaces and shallow ansätze.
- The U, V tables (and their ML predictors supplied in the original paper) define minimal correlated Hamiltonians for proof–of–principle quantum solvers (e.g., sample–based diagonalization with shallow LUCJ trial states), even when full MOF cells are still too large to fit on hardware.

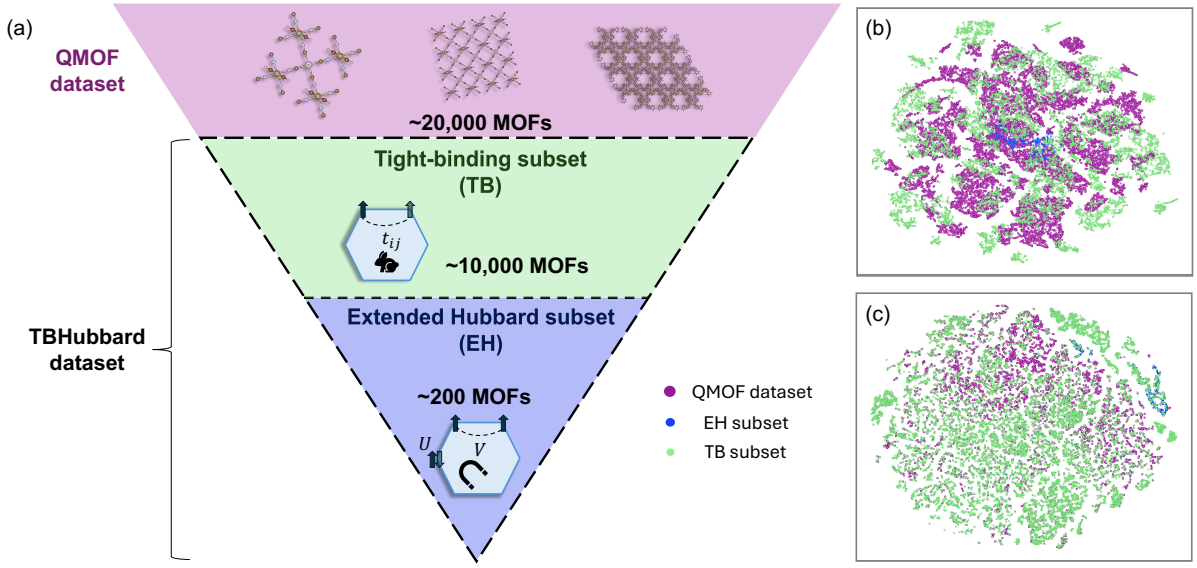


Figure 29 – **Dataset at a glance. (a)** Coverage of QMOF and its TB (green) and EH (blue) subsets. **(b)** t-SNE of TB matrices indicates distinct clustering among source databases and subsets. **(c)** t-SNE of SOAP–3 Å metal–center descriptors (after PCA) highlights chemically coherent neighborhoods—useful for model transfer and active learning. *Reproduced from* Carvalho et al. [8] (DOI), CC BY 4.0 (link).

- Zr/Hf–rich frameworks in the EH subset connect directly to oxide chemistries (e.g., ZrO_2, HfO_2) that are already tractable on current devices for band–edge benchmarks; TBHubbard supplies the parameter landscape for scaling toward full frameworks as qubit budgets grow.

Given a target MOF family (e.g., Zr–oxo clusters with carboxylate linkers): (1) extract TB from TBHubbard; (2) select band–edge orbitals (via TB embeddings/SOAP neighborhoods) to build a symmetry–aware active space; (3) attach EH corrections using tabulated or ML–predicted U, V(the dataset already provides some as seen in Figure 29); (4) reduce to a small k–point or real–space cluster and map to qubits; (5) run a shallow variational or sample–based diagonalization to estimate edge energies. Steps (1–3) are data lookups; only (4–5) touch quantum resources. This hybrid path turns TBHubbard into a *feeder dataset* for robust, low–depth quantum experiments tied directly to classical validation.

The pieces above set up a data-informed, low-depth route to band-edge physics: TBHubbard supplies standardized t (and, where available, U, V) plus descriptors to localize the relevant orbitals; the EH form in Eq. (6.1) delivers a compact, many-body Hamiltonian; and the fundamental-gap expression in Eq. (6.2) ties spectra to ground-state energies in fixed-N_e sectors. Full MOF cells remain outside near-term qubit and shot budgets, but the same metal–ligand physics appears in oxide building blocks (ZrO_2, HfO_2) that *are* tractable today. We therefore pivot from MOF-scale datasets to *compact binary oxides* as an executable benchmark for the sample-based approach.

We now instantiate this workflow on ZrO_2 and HfO_2:

1. Build material-informed EH Hamiltonians at Γ: project KS-DFT onto localized

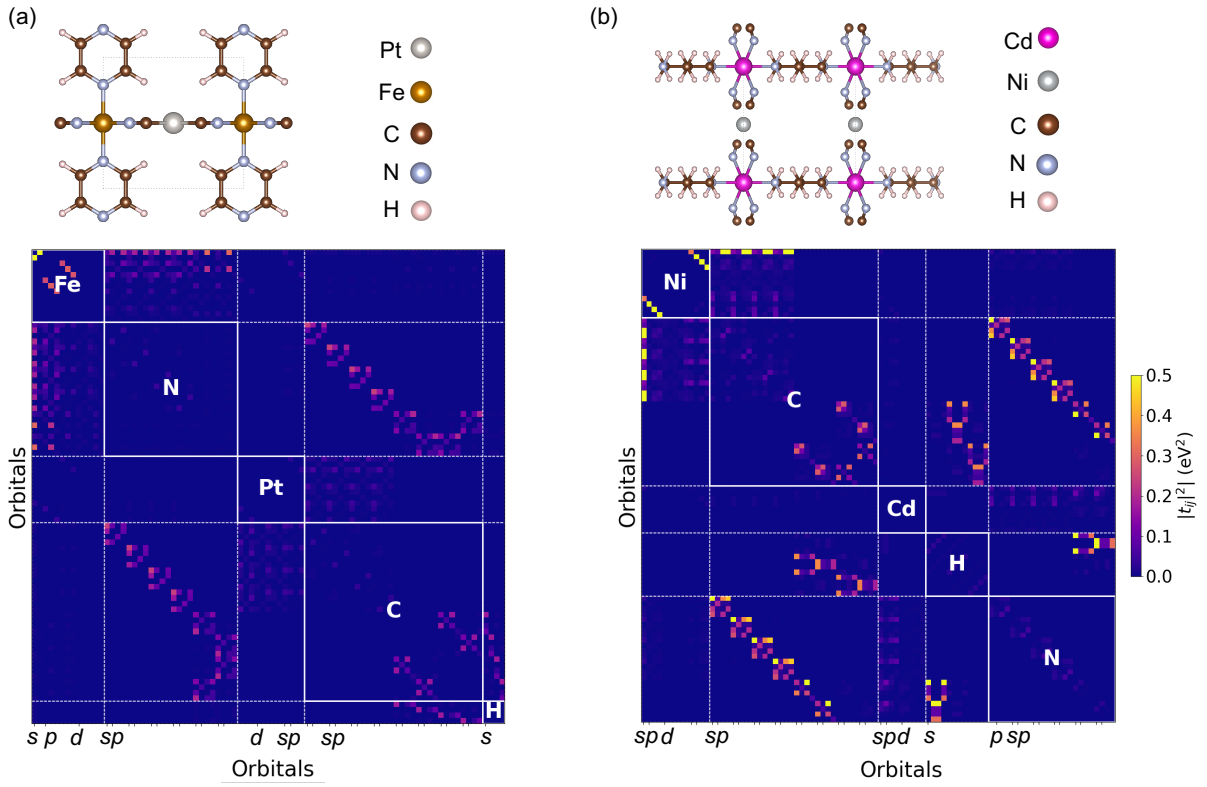


Figure 30 – **Representative tight–binding matrices at Γ for two MOFs.** False–color plots of $|t_{ij}|^2$ reveal strong intra–site blocks and chemically interpretable inter–site couplings (e.g., M–N, C–N, C–H). This sparsity/locality is exactly what near–term quantum solvers exploit when defining active spaces and excitation pools. *Reproduced from* Carvalho et al. [8] (DOI), CC BY 4.0 (link).

AOs (PAOFLOW) to obtain t_{pq}; add DFPT U_d and $V_{d\text{–}p}$ to capture metal–ligand correlation.

2. Map the EH model to a standard electronic Hamiltonian and define a symmetry-aware active space targeting band-edge orbitals (∼38–46 spin orbitals).

3. Prepare shallow LUCJ trial states from CCSD amplitudes; *sample* circuits to select configurations; apply configuration recovery; *classically* diagonalize in the sampled subspace (SQD / Ext-SQD).

4. Evaluate E_g via Eq. (6.2) in the N_e-1, N_e, N_e+1 sectors; compare against DFT, DFT+U+V, and trusted references.

5. Analyze accuracy vs. cost: effect of U vs. V, ansatz depth, sampling budget, and CI subspace size.

The next section (§6.1.5) details the SQD/Ext-SQD workflow and experimental settings; §6.1.5 reports the oxide band gaps (cf. Fig. 31, Fig. 32, and Table 12) and ablation studies that isolate the role of U and V in gap opening.

6.1.5 Results

As discussed previously, we validated that projecting KS-DFT onto localized pseudo-atomic orbitals reproduces the underlying bands near E_F while exposing a sparse, chemistry-aligned tight-binding (TB) structure. Figure 27 summarizes the validation process and interpretation (site = localized AO; orbitals on the same atom map to different sites). Independent band-structure checks show excellent agreement between DFT and the PAOFLOW TB projection around the gap region, and the TB matrices reveal block-locality.

Using the fundamental-gap definition $E_g = E(N_e + 1) + E(N_e - 1) - 2E(N_e)$ on the lattice Hamiltonian and the KS eigenvalue gap on DFT, we compare predictions for cubic HfO_2 and ZrO_2. Figure 31 collates four interaction settings on the lattice side—TB ($U = V = 0$), U-only, V-only, and $U + V$ (EH)—alongside DFT and literature references (GW and experiment) [9, 10]. Ground-state energies for the N_e-1, N_e, and N_e+1 sectors are obtained with HCI and treated as numerical ground truth for these instances.

Two robust trends emerge: (i) at the semilocal DFT level, adding U shifts gaps modestly; (ii) in the lattice picture, V is the dominant lever for gap opening in these oxides, with U providing a smaller, material-dependent correction. For both HfO_2 and ZrO_2, the EH gaps align more closely with GW benchmarks and fall within reported experimental ranges at Γ (noting phase/k-point variability in measurements [9, 10, 686]). A compact literature comparison for ZrO_2 is given in Table 12.

Method		Direct band gap (eV)
OLCAO [10]	Theory	4.9
LDA [687]	Theory	3.6
GW [687]	Theory	5.8
OLCAO [688]	Theory	3.8
DFT [689]	Theory	3.9
This work (EH/HCI)	**Theory**	**6.1**
Vacuum UV [10]	Experiment	6.1, 7.1

Table 12 – **Direct gap of cubic ZrO_2 at Γ:** literature vs. this work (EH value corresponds to the U+V point in Fig. 31). *Reproduced from* Duriez et al. [7] (DOI), CC BY 4.0 (link).

Hardware results: SQD/Ext-SQD ground states and gap accuracy

We next estimate E_g with sample-based diagonalization on hardware, using the same EH Hamiltonians at Γ and shallow LUCJ trial states seeded by CCSD amplitudes (see Chapters 4 and 5 for the workflow details already covered in the previously for the battery work). Figure 32 shows SQD and Ext-SQD ground-state energies for all three number sectors on `ibm_pittsburgh` with 38 qubits (HfO_2) and 46 qubits (ZrO_2). Errors are plotted versus the fraction of the configuration space included in the classical CI diagonalization.

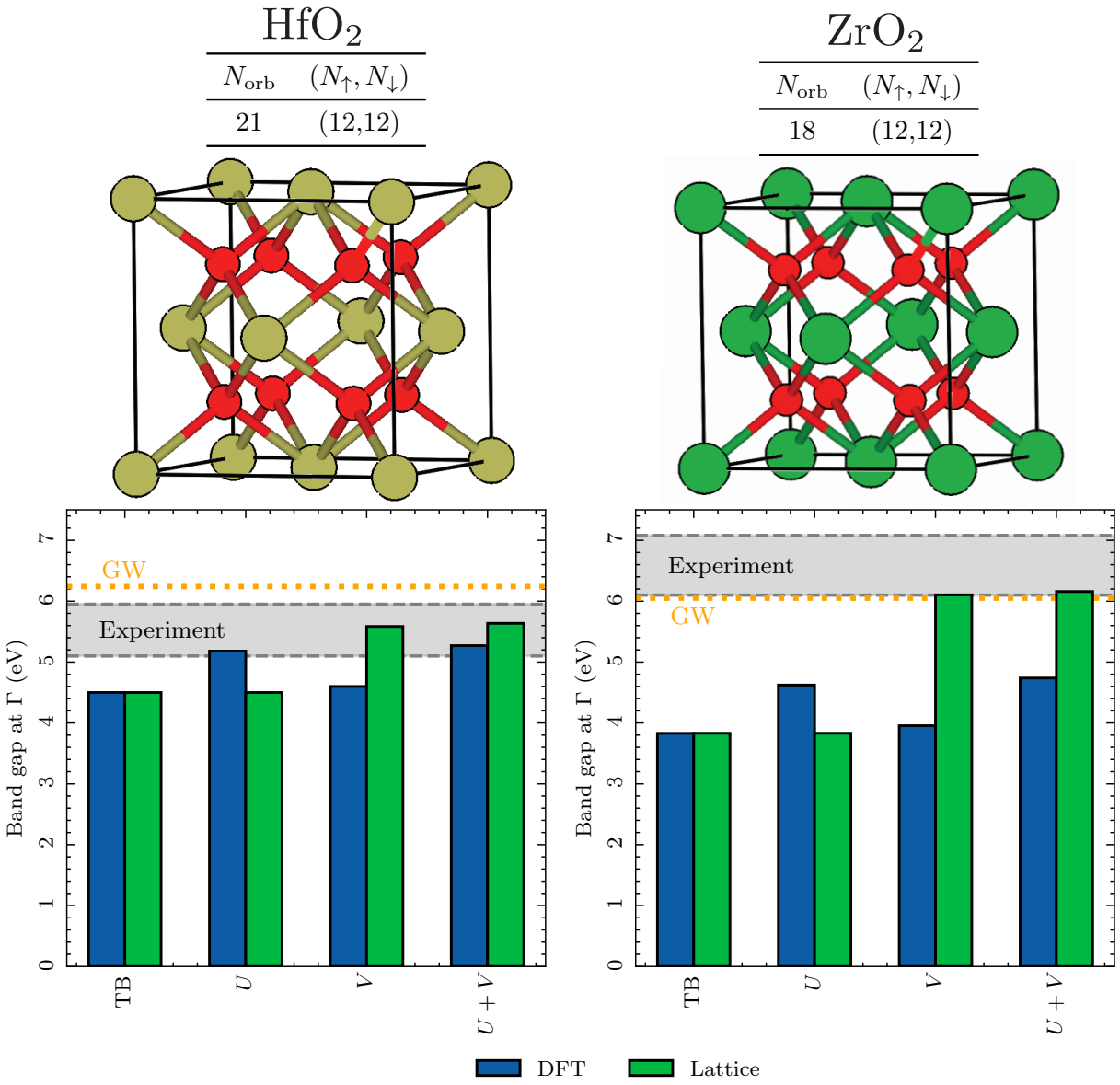


Figure 31 – **Direct band gaps at Γ for HfO_2 and ZrO_2 from DFT and from lattice Hamiltonians under different interaction settings (TB, U, V, $U+V$).** Lattice gaps are computed from ground-state energies in $N_e \pm 1$ and N_e sectors using HCI. GW and experimental references from [9, 10]. *Reproduced from* Duriez et al. [7] (DOI), CC BY 4.0 (link).

Across both materials, Ext-SQD (configuration recovery + excitation expansion) pushes energies below CCSD(T) error while requiring a smaller CI subspace than a comparable SQD run. Because band gaps are formed from differences of three high-accuracy ground-state energies, E_g from Ext-SQD matches the HCI reference to within $\mathcal{O}(10^{-7})$ eV (HfO_2: 1.6×10^{-8} eV; ZrO_2: 1.3×10^{-7} eV), effectively saturating numerical precision at the shot budgets used.

For the problem sizes here, HCI tightly approximates (and, for HfO_2 in N_e sectors, matches) exact diagonalization with vanishing variance (Fig. 33). Treating HCI as the classical baseline is therefore justified for these instances.

Wrapping up, Table 13 compiles the direct gaps at Γ for the systems treated here,

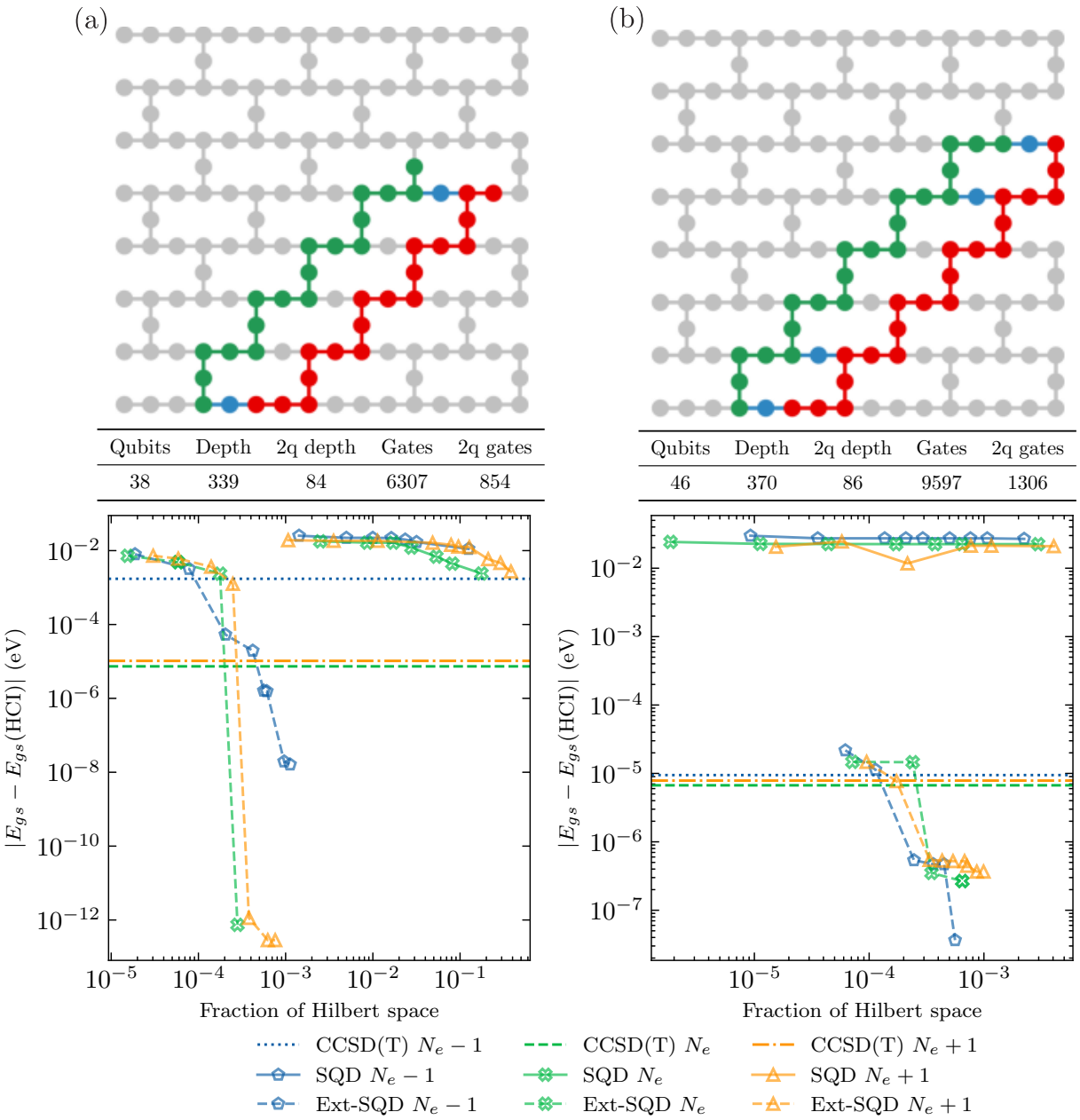


Figure 32 – **SQD/Ext-SQD on superconducting hardware for (a) HfO_2 (38 qubits) and (b) ZrO_2 (46 qubits). Errors vs. HCI (reference) as a function of CI-subspace fraction for N_e-1, N_e, and N_e+1 sectors.** Ext-SQD attains HCI-level accuracy with fewer configurations and surpasses CCSD(T). *Reproduced from* Duriez et al. [7] (DOI), CC BY 4.0 (link).

alongside representative theoretical and experimental references. Our entries report DFT (GGA), DFT+U+V, the tight-binding limit (U=V=0; TB), and EH model solved with HCI (as Ext-SQD essentialy match them) on the projected lattice Hamiltonians. These are the same values plotted in Figura 31.

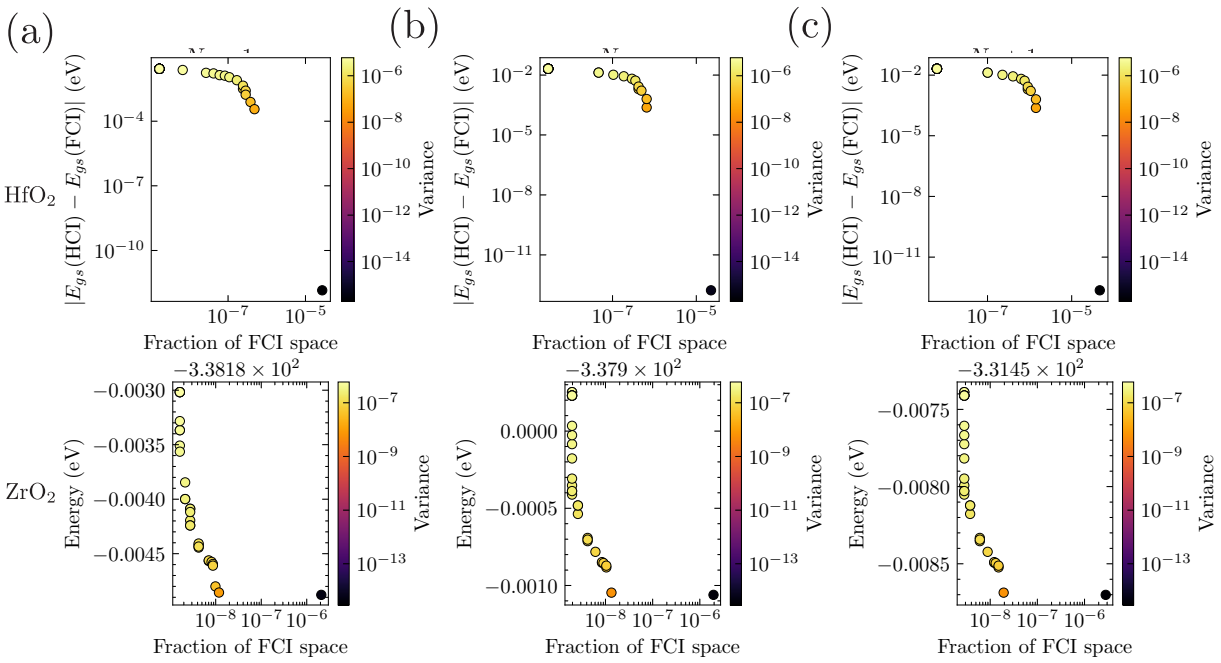


Figure 33 – **HCI convergence of ground-state energies for HfO_2 and ZrO_2 in the N_e (a), N_e-1 (b), and N_e+1 (c) electron-number sectors.** For HfO_2, we plot the deviation from FCI (exact) energies; for ZrO_2, we report absolute energies. The horizontal axis shows the fraction of the total configuration space included in the CI diagonalization (as in Fig. 32). Color bars encode the ground-state variance, $\mathrm{Var}(H) = (\langle H^2\rangle - \langle H\rangle^2)/\langle H\rangle^2$, with smaller values indicating higher-fidelity eigenstates. *Reproduced from* Duriez et al. [7] (DOI), CC BY 4.0 (link).

6.1.6 Discussion

The EH lattice route systematically improves DFT Γ-point gaps for HfO_2/ZrO_2 by incorporating short-range correlations absent at the semilocal level. In these oxides, inter-site V (metal–ligand) is the main driver of gap opening; the on-site U plays a secondary role. EH gaps align with GW and fall in the spread of experimental reports, where phase and k-point differences complicate one-to-one comparisons [9, 10, 686].

From a quantum-computing perspective, SQD/Ext-SQD extract chemically meaningful ground states and band gaps on current devices with shallow, physics-guided circuits and modest mitigation. Ext-SQD, in particular, leverages configuration recovery to reduce classical CI cost while improving accuracy beyond CCSD(T), yielding band-gap estimates numerically indistinguishable from HCI at the precisions considered.

With oxide benchmarks in hand, we conclude by outlining near-term extensions rather than presenting further results. This chapter is limited to direct gaps at Γ in cubic ZrO_2 and HfO_2. The same DFT→EH construction and SQD/Ext-SQD workflow can, in principle, be extended to additional **k**-points, larger active spaces, and other chemistries as resources permit; TBHubbard will then serve as a standardized source of EH test cases, not a target addressed here.

Looking ahead, the immediate technical avenues are clear. On the modeling side, the DFT→lattice projection combined with self-consistent U and V offers a portable route

Table 13 – **Direct band gaps at Γ (E_g^Γ, eV) for ZrO_2, HfO_2, and Si.** TB and Extended-Hubbard (EHM) values are from HCI on the projected lattice Hamiltonians; literature values are shown for context. *Reproduced from* Duriez et al. [7] (DOI), CC BY 4.0 (link).

Method	ZrO_2	HfO_2	Si bulk
This work			
DFT (GGA)	3.83	4.50	2.57
DFT+U+V	4.74	5.27	2.53
TB	3.83	4.50	2.57
EHM	6.16	5.64	2.53
Other works			
DFT	3.65[a] (LDA)	$\sim$ 4.2[b] (LDA)	2.57[c] (LDA)
GW	5.81[a]	$\sim$ 6[b] (GW_0)	3.27[c]
Experimental	6.1-7.08[e]	5.10 – 5.95[b]	3.40[c]

[a] Ref. [687] [b] Ref. [9] [c] Ref. [690] [e] Ref. [10]

to many-body Hamiltonians; the main bottleneck remains classical CI in the sampled subspace. Progress in selected-CI preselection, Davidson/Krylov solvers, and sparse/tensor accelerations can push to larger active spaces without altering the quantum front end. On the quantum side, shallow physics-informed trial states, configuration-recovery, and modest mitigation already deliver sub-millihartree accuracy on pre–fault-tolerant devices; incremental gains in device stability and shot budgets should enable multi-**k** sampling (indirect gaps, rudimentary dispersions) and modest basis growth. Beyond the present work, TBHubbard is best viewed as a standardized source of EH instances and descriptors to benchmark these choices across broader chemistries once hardware and classical resources allow. In the longer term, fault-tolerant primitives (e.g., QPE) provide a principled endpoint; until then, sample-based diagonalization remains a practical bridge between ab initio materials modeling and today's hardware.

This thesis has traced a pragmatic arc: from the limitations of variational approaches, to sample-based diagonalization on embedded fragments, and finally to periodic oxides where a material-informed EH model plus (Ext-)SQD yields band-gap predictions aligned with experiment and high-level theory. In the concluding chapter, we step back to synthesize these results, distill lessons about algorithm–model co-design in quantum-centric workflows, and outline concrete milestones for scaling from chemically realistic fragments to broader classes of materials and properties.

7 Conclusions and Future Directions

This thesis examined whether sample-based quantum algorithms—principally Sample-Based Quantum Diagonalization (SQD) and its extended form (Ext-SQD)—can deliver chemically meaningful predictions for realistic systems on pre–fault-tolerant hardware. The work positioned these methods within quantum-centric supercomputing workflows, in which quantum processors provide sampling over low-depth state families and classical HPC infrastructure performs subspace diagonalization and analysis. Two pillars grounded the investigation: (i) embedded surface chemistry relevant to lithium–oxygen battery interfaces; and (ii) periodic, material-informed EH Hamiltonians for band-edge physics in crystalline oxides. Across these settings, the central finding is that (Ext-)SQD yields reliable ground-state energies and derived observables (e.g., reaction energies and direct band gaps) at shallow circuit depths by decoupling state preparation from variational optimization and shifting the eigenvalue problem into a classically diagonalized, sampled subspace.

7.1 Consolidated Answers to Goals & Questions

Returning to the aims in Sec. 1.2, we summarize what the thesis establishes for each research question and point to the concrete artifacts (chapters, figures, and tables) where the claims are borne out.

- **RQ1 — Limits of variational algorithms and motivation to pivot.** Chapter 3 shows that qubit-ADAPT-VQE attains high accuracy on small active spaces but faces steep measurement cost and depth/optimizer fragility as systems grow, even with symmetry tapering and careful operator pools. These controlled baselines motivate replacing inner-loop optimization with sampling-plus-diagonalization in later chapters.
- **RQ2 — Adapting sample-based diagonalization to realistic fragments and periodic models.** Chapters 5 and 6 operationalize SQD/Ext-SQD in two settings: embedded surface fragments (active spaces defined by DDA $\rightarrow$ CC natural orbitals) and material-informed periodic Hamiltonians (DFT$\rightarrow$TB+U+V projection). The surface workflow and its quantum sampling are illustrated in Fig. 23 and the device mapping in Fig. 22; the periodic projection and TB validation are documented in Fig. 27.
- **RQ3 — Achieving chemically meaningful accuracy.** For Li–O_2 surface energetics, ground and reaction energies converge against CI baselines as active spaces grow (Fig. 21); Ext-SQD improves over SQD and tracks Ext-HCI (Fig. 25), with selected ΔE values compiled in Table 11. For oxides, direct gaps at Γ computed from $E(N_e \pm 1)$ sectors align with high-level references (Fig. 31; Table 12, Table 13), and SQD/Ext-SQD recover HCI-level sector energies at available shot budgets (Fig. 32), with classical convergence shown in Fig. 33.

- **RQ4 — Integration with embedding & effect of design choices.** Design levers are made explicit and tied to resources: active-space growth and accuracy (Fig. 21, Fig. 25); sampling quality vs. circuit design and mapping (Fig. 23, Fig. 22); and quantum resource profiles (Table 9). On the periodic side, the role of U vs. V in opening gaps is isolated (Fig. 31), while TBHubbard (Sec. 6.1.3) supplies standardized TB/EH instances and descriptors that inform active-space selection and parameter regimes.

- **RQ5 — Benchmarks on simplified systems to assess stability and cost.** Chapter 3 shows that—even with symmetry tapering and curated operator pools—qubit-ADAPT-VQE's depth and measurement cost grow steeply with active-space size, and optimizer stability degrades, making results beyond small/medium spaces impractical on today's hardware; this motivates the pivot rather than contradicting it. For surfaces (Chapter 5), hardware sampling quality collapses past ∼30–32 orbitals as invalid Hamming-weight rates spike (Fig. 23); tensor-network sampling extends diagnostics to ∼40 orbitals but then the classical CI step—not the quantum front end—becomes the bottleneck (Fig. 21, Fig. 25), and resource profiles indicate that deeper ansätze or larger spaces exceed feasible two-qubit budgets (Table 9). For periodic oxides (Chapter 6), (Ext-)SQD attains HCI-level sector energies in 38–46 spin-orbital windows (Fig. 32), but scaling to indirect gaps or dispersions would require multi-**k** sampling and larger active spaces, which are currently shot- and CI-limited; moreover, TB+U+V model fidelity is contingent on projector/manifold choices (Fig. 31). In short, full MOF unit cells, wide k-meshes, and deeply correlated fragments remain out of reach on present devices; stable, accurate results are achievable only within the above windows, with configuration recovery and careful active-space curation as non-negotiable ingredients.

7.2 Consolidated Contributions

- **Strategic perspective on materials simulation (Chapter 2).** Adapted from a perspectives work and tailored to this thesis, Chapter 2 formalizes the quantum-centric supercomputing (QCSC) stack for materials: it delineates pre–fault-tolerant vs. fault-tolerant regimes, emphasizes hybrid QC–HPC integration, and sets concrete reporting conventions (qubits, two-qubit depth/CX counts, sampling budgets) and a use-case rubric for realistic modeling goals.

- **Early-stage variational benchmarks (Chapter 3).** Established a control case for accuracy vs. depth/measurements and optimizer stability on simplified systems, setting the yardstick that later motivates sampling-based approaches (see chapter figures/tables). Scaling sensitivity becomes apparent beyond small active spaces, with wall-time and shot counts growing steeply. This motivates the pivot away from fragile variational optimization and frames the deployment/limits of (Ext-)SQD in later chapters.

- **Method integration and deployment (Chs. 5–6).** Integrated SQD/Ext-SQD with shallow LUCJ trials, configuration recovery, and reproducible diagnostics into

a portable QCSC workflow. Classical CI cost in the sampled subspace becomes the dominant bottleneck as the subspace grows.

- **Embedded Li–O_2 surface reactions (Chapter 5).** DDA → NO-refined active spaces plus (Ext-)SQD deliver reaction energies that track CI within chemical-accuracy tolerances across feasible sizes; see Fig. 21, Fig. 25, and Table 11, with resource profiles in Table 9. Hardware invalid–Hamming-weight rates rise sharply past ∼30–32 orbitals; tensor-network sampling to ∼40 orbitals (Fig. 23, Fig. 22) separates sampling effects from model truncation.
- **Periodic EH Hamiltonians for oxide band gaps (Chapter 6).** DFT → TB+U+V projections (Fig. 27) solved with (Ext-)SQD reproduce direct Γ-point gaps for ZrO_2/HfO_2; see Fig. 31, Table 12 and Table 13. Ext-SQD attains HCI-level sector energies with smaller CI subspaces (Fig. 32; HCI baseline in Fig. 33). Current demonstrations span 38–46 spin orbitals; V dominates gap opening, with U a secondary, material-dependent correction.
- **Data/descriptor bridge (TBHubbard; Chapter 6).** Curated TB/EH instances with U, V and aligned descriptors that inform active-space localization and provide standardized, CC-licensed benchmarks for future periodic SQD workflows. Full MOF unit cells remain beyond near-term qubit/shot budgets; dataset presently anchors oxide-scale proxies and framework fragments.

7.3 Review

We began by establishing the cost model and workflow scaffolding used throughout the thesis. We framed *quantum-centric supercomputing* (QCSC) as the practical integration point between quantum devices and HPC, then enumerated the algorithmic menu most relevant to materials—ground states (VQE, subspace/Krylov, QPE), short-time dynamics, open-system and finite-temperature primitives—clarifying which belong to pre-fault-tolerant versus fault-tolerant regimes (Sec. 2.2). We then turned to the classical information-processing levers that actually determine feasibility on current hardware: fermion/boson encodings and symmetry tapering (Sec. 2.3.1), measurement strategies and grouping/tomography trade-offs, problem-level preprocessing (active spaces, embedding) and variational ansatz design/initialisation (CAFQA/near-Clifford starts), as well as connectivity-aware compilation, SWAP networks, circuit cutting, and dynamic-circuit optimizations (Sec. 2.3, 2.3.5). Error-mitigation options (symmetry checks, ZNE/PEC families) and the continuing role of classical simulation backends (state-vector, density-matrix, MPS/TN) round out the execution stack (Sec. 2.4). Finally, a use-case rubric ties scientific impact to computational plausibility and motivates embedding/model-system targets over full ab initio workloads for near-term runs (Sec. 2.5).

Motivated by this, in Chapter 3 we developed a reproducible variational baseline built around qubit-ADAPT-VQE. Using the mapping and symmetry-tapering choices discussed in Sec. 2.3.1, we specify the operator pool, the selection rule, stopping criteria, and initialisation strategies —all tied to the resource model introduced in Sec. 2.2. To make results portable across simulators and hardware, we report qubit counts and wall-time

projections. Baseline runs are validated on classical backends (PySCF HF and DFT as well as direct diagonalization) using the same estimators, grouping, and optimizers intended for hardware when applicable.

Next, we articulated the design problems that drive the pivot to sampling-based algorithms in subsequent chapters: push coherent depth to a minimum; replace inner-loop variational optimization with classical linear algebra; and preserve compatibility with the active-space and embedding choices. In short, Chapter 3 is the underwhelming result that Chapter 2 made testable: it translates the execution-stack guidance into quantitative evidence for why sampling-based diagonalization offers the more plausible QCSC path forward on near-term hardware.

Following that, Chapter 4 operationalizes the pivot that Chapters 2 and 3 make compelling. This chapter formalizes a sampling-based diagonalization program. The central design choice is to trade coherent depth and global optimization for repeated Z-basis sampling plus classical solvers: prepare shallow trials, select a compact subspace from bitstrings (with symmetry post-selection and configuration recovery), and diagonalize H there. We present this under a unified umbrella that includes QSCI/SQD (variational, determinant-based), Ext-SQD (deterministic excitation closure for ground and excited states), and SKQD (short-time Krylov snapshots with provable support recovery).

Conceptually, Chapter 4 is system-agnostic: the algorithms are developed and benchmarked on multiple instances expressly to expose accuracy–cost trade-offs. We show how configuration recovery stabilizes sparse subspaces; how symmetry filters and lightweight mitigation convert raw histograms into reliable bases; and how SKQD's multi-snapshot union improves coverage when a single trial under-resolves support. Throughout, guarantees and diagnostics are tied to the QCSC perspective: variational upper bounds from subspace enlargement (SQD/Ext-SQD), polynomial sample schedules under sparsity assumptions (SKQD), and monotone energy behavior as a sanity check across growth steps.

Later on, we specify the interfaces to scale these methods in practice. Embedding and qubit-reduction tools motivated in Sec. 2.3 (active spaces, DMET-style fragments, entanglement forging) are treated here as first-class wrappers that shrink the quantum footprint and keep systems tractable, while classical HPC shoulders Hamiltonian assembly and eigensolves in the QCSC split. The outcome is a concrete playbook for constructing, conditioning, and solving sample-defined subspaces on pre-fault-tolerant hardware.

Chapter 5 then instantiates this playbook for heterogeneous electrochemistry at Li-O_2 battery surfaces. We build embedded active spaces via DDA refined with coupled-cluster natural orbitals, then prepare shallow LUCJ trial states to sample bitstring configurations. The SQD pipeline recovers physically valid configurations and diagonalizes the Hamiltonian in the sampled subspace; Ext-SQD augments that space with screened single/double excitations. Reaction energetics are reported as $\Delta E = E_{\mathrm{prod}} - E_{\mathrm{reac}}$, using matched active-space definitions for reactant and product. Quantitative convergence and accuracy are shown in Fig.21 and Fig.25, where Ext-SQD improves over SQD and tracks Ext-HCI across feasible active spaces; selected reaction-energy values are compiled in Table 11. Resource profiles for representative circuits appear in Table 9, while sampling distributions and chip mappings are illustrated in Fig.23 and Fig.22.

The chapter also quantifies practical limits. Hardware sampling exhibits rising

invalid Hamming-weight rates beyond ∼30–32 orbitals, at which point configuration recovery and symmetry filters are required to maintain quality. MPS-based sampling provides scale-out diagnostics to ∼40 orbitals, revealing convergence trends and separating sampling effects from model truncation. These results delineate both the promise and the boundaries of (Ext-)SQD for surface energetics under current device noise but, overall, physics-informed active-space selection plus shallow LUCJ sampling and sample-based diagonalization provides a practical near-term blueprint for quantum-assisted surface chemistry.

Chapter 6 repeats the same playbook to periodic materials. We construct material–informed lattice Hamiltonians by projecting KS–DFT onto localized pseudo–atomic orbitals to obtain tight–binding matrices, then add screened on–site U and inter–site V to form an Extended–Hubbard model (Sec. 6.1.2). Direct gaps at Γ are computed from many-body ground states via $E_g = E(N_e{+}1) + E(N_e{-}1) - 2E(N_e)$. Band-edge active spaces of ∼38–46 spin orbitals are mapped to shallow, truncated LUCJ trials; bitstrings are sampled, symmetry-filtered, and passed through configuration recovery, and the Hamiltonian is diagonalized in the measurement-defined subspace with SQD/Ext-SQD. Projection fidelity and lattice provenance are shown in Fig.27; Band-gap outcomes at Γ across TB, U, V, and $U{+}V$ are summarized in Fig.31, with a compact literature comparison in Table 12 (see also Table 13). SQD/Ext-SQD sector energies and gap accuracy relative to HCI are reported in Fig.32, with the classical HCI convergence baseline in Fig.33.

Two robust trends emerge on compact oxides (ZrO_2, HfO_2). First, V (metal–ligand) is the principal driver of gap opening, with U providing a smaller, material-dependent correction; the EH gaps align with GW and fall within experimental spreads at Γ. Second, Ext-SQD reaches HCI-level accuracy for the $N_e{-}1, N_e, N_e{+}1$ sectors using smaller CI subspaces than SQD, making the resulting E_g numerically indistinguishable from the classical reference at the shot budgets considered. Throughout, accuracy/cost is controlled by explicit design levers (basis/localization, U/V manifolds, excitation pools, recovery iterations, sampling budgets), matching the QCSC resource model.

Finally, TBHubbard (Sec.6.1.3) serves as a standardized feeder of TB/EH instances and descriptors: it motivates why U+V is the minimal correlated extension needed for MOF-like chemistries and provides transferable Hamiltonians for benchmarking quantum solvers as qubit budgets grow. In combination with the surface-reaction study of Chapter 5, the periodic results in Chapter 6 demonstrate that the same sampling-based diagonalization toolkit spans heterogeneous interfaces and crystalline band edges.

Taken together, the chapters trace a coherent arc: Chapter 2 motivates the need for near-term, accuracy-conscious workflows and formalizes modeling choices and baselines; Chapter 3 measures a variational control case (qubit-ADAPT-VQE) and exposes its scaling sensitivities; Chapter 4 replaces fragile optimization with sampling plus compact diagonalization; and Chapters 5 and 6 demonstrate domain relevance on embedded surfaces and periodic proxies with transparent resource accounts.

7.4 Implications and future directions

This thesis argued that sampling-plus-diagonalization is a practical near-term path for quantum electronic-structure: by trading deep, globally optimized circuits for shallow, physics-guided trials and classical diagonalization in a measurement-defined subspace, we recover CI-quality energies on problems that are otherwise out of reach for unstable variational training. The implication for computational chemistry is concrete: CI-grade benchmarks become accessible for problem sizes that matter to reactivity and band edges, using resources that match today's quantum and HPC hardware.

For quantum computing, the results supply credible milestones short of fault tolerance and a template for reporting them. Resource profiles (qubits, depth, CX counts) tied to chip topology and calibration-aware mapping make it straightforward to compare devices and compilers on the same Hamiltonians. Configuration recovery and lightweight symmetry checks convert raw histograms into reliable CI bases, and Ext-SQD provides a deterministic improvement knob without deeper circuits. Equally important are the honest limits: invalid-Hamming-weight rates rise sharply beyond ∼30–32 orbitals in current runs, at which point sample quality, not algorithmic design, becomes the bottleneck. This clarifies where hardware advances (gate fidelity, connectivity, mid-circuit features) most directly translate into scientific wins.

For materials science, the workflow delivers a screening-oriented capability that complements established DFT and many-body practice. On surfaces (Chapter 5), density-difference–guided active spaces refined by natural orbitals focus correlation at the reactive site, enabling reaction-energy trends with explicit error bars relative to CI references. In periodic settings (Chapter 6), a material-informed lattice model (TB+U+V) projected from DFT and solved by (Ext-)SQD recovers direct gaps at Γ in representative oxides, with transparent levers to balance accuracy and cost (choice of basis/localization, U/V manifolds, CI subspace size). Together these pieces support near-term studies that prioritize rank-ordering candidates and isolating mechanistic effects over full ab initio fidelity, while still remaining quantitatively anchored.

A second implication is methodological standardization through data assets. TBHubbard provides tight-binding and extended-Hubbard instances, descriptors, and metal–ligand interaction statistics that align with the Hamiltonians actually used in quantum runs. This turns dataset curation into an algorithm co-design tool: the same models support classical baselines (HCI, CC), machine-learning surrogates for U/V, and quantum experiments with shallow ansätze. Reproducibility improves because figures and tables point to portable inputs and unambiguous targets; comparability improves because resource reports and accuracy metrics are defined on a common set of instances. As classical CI becomes the bottleneck at larger subspaces, the QCSC split also highlights where HPC most effectively extends quantum reach. The overarching implication is that useful, verifiable materials predictions are already feasible on pre–fault-tolerant hardware, provided we co-design models, ansätze, and classical solvers within a disciplined QCSC framework.

Finally, the results suggest a staged roadmap. In the short term, Ext-SQD atop carefully built active spaces can push near or below chemical-accuracy thresholds without deeper circuits, while robust PT2 closures (CASPT2/NEVPT2) can recover external

correlation atop Ext-SQD references without altering the quantum front end. On the modeling side, automated active-space selection from learned descriptors will standardize what is currently a semi-manual step.

In the medium term, implications center on scaling and standardization. Periodic cases can expand beyond Γ via coarse k-meshes or compact supercells coupled to (Ext-)SQD, while tighter embedding integrations (DMET/LASSCF and LAS-UCC hybrids) increase chemical realism at modest depth. The field could benefit from open datasets and leaderboards that mirror our surface and EH-lattice benchmarks, turning SQD-style studies into comparable, regression-friendly exercises. On hardware, connectivity-aware LUCJ variants, SWAP-efficient routing, and symmetry-preserving compilation links device progress to immediate scientific payoff by improving sample quality at fixed qubits and depth.

Longer-term implications trace a continuity path from today's prototypes to asymptotically exact solvers. The same EH and embedded Hamiltonians can be revisited with fault-tolerant primitives (QPE-like) once available, providing a clean handoff from sample-defined subspaces to phase-estimation accuracy. End-to-end design loops—target property $\rightarrow$ dataset instance $\rightarrow$ SQD run $\rightarrow$ validation—can be coupled to lab automation for closed-loop discovery.

The central question posed at the outset was whether shallow, sampling-driven quantum workflows can already deliver chemically meaningful predictions. The evidence across Chapters 3–6 indicates that they can: by measuring judiciously and diagonalizing compact, physics-informed subspaces, we obtain CI-quality energetics and band edges with transparent, reproducible resource accounts. In this sense, the most effective near-term strategy is not to make circuits deeper, but to make models and subspaces sharper, anchoring quantum effort where it matters and outsourcing the rest to classical HPC. The same material-informed Hamiltonians and datasets that let us measure-then-diagonalize today become clean targets for phase-estimation tomorrow. In that sense, this work is less a destination than a calibrated on-ramp to certified, fault-tolerant electronic structure in the future.

Bibliography

1 ALEXEEV, Y. et al. “Quantum-centric supercomputing for materials science: A perspective on challenges and future directions”. *Future Gener. Comput. Syst.*, Elsevier BV, v. 160, p. 666–710, nov. 2024.

2 TAZHIGULOV, R. N. et al. “Simulating models of challenging correlated molecules and materials on the sycamore quantum processor”. *PRX Quantum*, APS, v. 3, n. 4, p. 040318, 2022.

3 MOSTAME, S. et al. “Emulation of complex open quantum systems using superconducting qubits”. *Quantum Information Processing*, v. 16, p. 44, 2016. Availabe at: <https://doi.org/10.1007/s11128-016-1489-3>.

4 KIM, S. et al. “High-performance transparent radiative cooler designed by quantum computing”. *ACS Energy Letters*, v. 7, p. 4134–4141, 2022.

5 BARROCA, M. A.; FERREIRA, R. N. B.; STEINER, M. “Exploring qubit-adapt-vqe for materials discovery in direct air capture”. *APL Quantum*, v. 1, n. 4, p. 046103, 10 2024. ISSN 2835-0103. Availabe at: <https://doi.org/10.1063/5.0219500>.

6 BARROCA, M. A. et al. “*Scaling active spaces in simulations of surface reactions through sample-based quantum diagonalization*”. 2025. Availabe at: <https://arxiv.org/abs/2503.10923>.

7 DURIEZ, A. et al. “*Computing band gaps of periodic materials via sample-based quantum diagonalization*”. 2025. Availabe at: <https://arxiv.org/abs/2503.10901>.

8 CARVALHO, P. C. et al. “TBHubbard: tight-binding and extended hubbard model dataset for metal-organic frameworks”. *Sci. Data*, Springer Science and Business Media LLC, v. 12, n. 1, nov. 2025.

9 JIANG, H. et al. “Electronic band structure of zirconia and hafnia polymorphs from the *gw* perspective”. *Phys. Rev. B*, American Physical Society, v. 81, p. 085119, Feb 2010. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevB.81.085119>.

10 FRENCH, R. H. et al. “Experimental and theoretical determination of the electronic structure and optical properties of three phases of zro_2”. *Phys. Rev. B*, American Physical Society, v. 49, p. 5133–5142, Feb 1994. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevB.49.5133>.

11 GIAMMONA, M. J. et al. “Oxygen assisted lithium-iodine batteries: Towards practical iodine cathodes and viable lithium metal protection strategies”. *Adv. Mater. Interfaces*, Wiley, v. 10, n. 17, jun. 2023.

12 GIORDANI, V. et al. “Synergistic effect of oxygen and LiNO3on the interfacial stability of lithium metal in a Li/O2Battery”. *J. Electrochem. Soc.*, The Electrochemical Society, v. 160, n. 9, p. A1544–A1550, 2013.

13 ZHENG, S. et al. "Air-exposed lithium metal as a highly stable anode for low-temperature energy storage applications". *Energy Materials*, v. 2, n. 6, 2022. ISSN 2770-5900. Availabe at: <https://www.oaepublish.com/articles/energymater.2022.66>.

14 SHEN, C. et al. "Li2O-reinforced solid electrolyte interphase on three-dimensional sponges for dendrite-free lithium deposition". *Front. Chem.*, Frontiers Media SA, v. 6, p. 517, nov. 2018.

15 WANG, E. et al. "Effects of atmospheric gases on li metal cyclability and solid-electrolyte interphase formation". *ACS Energy Lett.*, American Chemical Society (ACS), v. 5, n. 4, p. 1088–1094, abr. 2020.

16 BIZ, C.; FIANCHINI, M.; GRACIA, J. "Strongly correlated electrons in catalysis: Focus on quantum exchange". *ACS Catalysis*, v. 11, n. 22, p. 14249–14261, 2021. Availabe at: <https://doi.org/10.1021/acscatal.1c03135>.

17 GAVINI, V. et al. "Roadmap on electronic structure codes in the exascale era". *Modelling and Simulation in Materials Science and Engineering*, IOP Publishing, v. 31, n. 6, p. 063301, aug 2023. Availabe at: <https://dx.doi.org/10.1088/1361-651X/acdf06>.

18 KNOWLES, P.; HANDY, N. "A new determinant-based full configuration interaction method". *Chemical Physics Letters*, v. 111, n. 4, p. 315–321, 1984. ISSN 0009-2614. Availabe at: <https://www.sciencedirect.com/science/article/pii/000926148485513X>.

19 HELGAKER, T.; JORGENSEN, P.; OLSEN, J. "*Molecular Electronic Structure Theory*". Nashville, TN: John Wiley & Sons, 2013.

20 JONES, R. O. "Density functional theory: Its origins, rise to prominence, and future". *Rev. Mod. Phys.*, American Physical Society, v. 87, n. 3, p. 897–923, Aug 2015. Availabe at: <https://link.aps.org/doi/10.1103/RevModPhys.87.897>.

21 BURKE, K. "Perspective on density functional theory". *The Journal of Chemical Physics*, v. 136, n. 15, p. 150901, 04 2012. ISSN 0021-9606. Availabe at: <https://doi.org/10.1063/1.4704546>.

22 HOHENBERG, P.; KOHN, W. "Inhomogeneous electron gas". *Phys. Rev.*, American Physical Society, v. 136, p. B864–B871, Nov 1964. Availabe at: <https://link.aps.org/doi/10.1103/PhysRev.136.B864>.

23 KOHN, W.; SHAM, L. J. "Self-consistent equations including exchange and correlation effects". *Phys. Rev.*, American Physical Society, v. 140, p. A1133–A1138, Nov 1965. Availabe at: <https://link.aps.org/doi/10.1103/PhysRev.140.A1133>.

24 CLEMENTI, E.; CHAKRAVORTY, S. J. "A comparative study of density functional models to estimate molecular atomization energies". *The Journal of Chemical Physics*, v. 93, n. 4, p. 2591–2602, 08 1990. ISSN 0021-9606. Availabe at: <https://doi.org/10.1063/1.458899>.

25 KLIMEŠ, J.; MICHAELIDES, A. "Perspective: Advances and challenges in treating van der waals dispersion forces in density functional theory". *J. Chem. Phys.*, AIP Publishing, v. 137, n. 12, p. 120901, set. 2012.

26 COHEN, A. J.; MORI-SáNCHEZ, P.; YANG, W. "Challenges for density functional theory". *Chemical Reviews*, v. 112, n. 1, p. 289–320, 2012. PMID: 22191548. Availabe at: <https://doi.org/10.1021/cr200107z>.

27 PERDEW, J. P.; ZUNGER, A. "Self-interaction correction to density-functional approximations for many-electron systems". *Phys. Rev. B*, American Physical Society, v. 23, n. 10, p. 5048–5079, May 1981. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevB.23.5048>.

28 PURVIS III, G. D.; BARTLETT, R. J. "A full coupled-cluster singles and doubles model: The inclusion of disconnected triples". *J. Chem. Phys.*, AIP Publishing, v. 76, n. 4, p. 1910–1918, fev. 1982.

29 BARTLETT, R. J.; MUSIAŁ, M. "Coupled-Cluster Theory in Quantum Chemistry". *RMP*, APS, v. 79, n. 1, p. 291–352, 2007. Availabe at: <https://link.aps.org/doi/10.1103/RevModPhys.79.291>.

30 PERUZZO, A. et al. "A variational eigenvalue solver on a photonic quantum processor". *Nature Communications*, v. 5, n. 1, p. 4213, 2014. Availabe at: <https://doi.org/10.1038/ncomms5213>.

31 O'MALLEY, P. J. J. et al. "Scalable quantum simulation of molecular energies". *Phys. Rev. X*, American Physical Society, v. 6, p. 031007, Jul 2016. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevX.6.031007>.

32 KANDALA, A. et al. "Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets". *Nature*, Macmillan Publishers Limited, part of Springer Nature. All rights reserved. SN -, v. 549, p. 242, 09 2017. Availabe at: <https://doi.org/10.1038/nature23879>.

33 COLLESS, J. I. et al. "Computation of molecular spectra on a quantum processor with an error-resilient algorithm". *Phys. Rev. X*, American Physical Society, v. 8, p. 011021, Feb 2018. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevX.8.011021>.

34 DIRAC, P. A. M.; POLKINGHORNE, J. C. "*the principles of quantum mechanics*". *Phys. Today*, AIP Publishing, v. 11, n. 6, p. 32–33, jun. 1958.

35 WIERICHS, D.; GOGOLIN, C.; KASTORYANO, M. "Avoiding local minima in variational quantum eigensolvers with the natural gradient optimizer". *Phys. Rev. Res.*, American Physical Society, v. 2, p. 043246, Nov 2020. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevResearch.2.043246>.

36 LAROCCA, M. et al. "Barren plateaus in variational quantum computing". *Nature Reviews Physics*, Springer Science and Business Media LLC, v. 7, n. 4, p. 174–189, mar. 2025. ISSN 2522-5820. Availabe at: <http://dx.doi.org/10.1038/s42254-025-00813-9>.

37 MCCLEAN, J. R. et al. "Barren plateaus in quantum neural network training landscapes". *Nat. Commun.*, Springer Science and Business Media LLC, v. 9, n. 1, p. 4812, nov. 2018.

38 TEMME, K.; BRAVYI, S.; GAMBETTA, J. M. "Error mitigation for short-depth quantum circuits". *Physical Review Letters*, American Physical Society (APS), v. 119, n. 18, nov. 2017. ISSN 1079-7114. Availabe at: <http://dx.doi.org/10.1103/PhysRevLett.119.180509>.

39 HUGGINS, W. J. et al. "Efficient and noise resilient measurements for quantum chemistry on near-term quantum computers". *Npj Quantum Inf.*, Springer Science and Business Media LLC, v. 7, n. 1, p. 1–9, fev. 2021. Availabe at: <https://doi.org/10.1038/s41534-020-00341-7>.

40 GONTHIER, J. F. et al. "Measurements as a roadblock to near-term practical quantum advantage in chemistry: Resource analysis". *Phys. Rev. Res.*, American Physical Society (APS), v. 4, n. 3, ago. 2022.

41 ENDO, S. et al. "Hybrid quantum-classical algorithms and quantum error mitigation". *J. Phys. Soc. Jpn.*, Physical Society of Japan, v. 90, n. 3, p. 032001, mar. 2021.

42 ROBLEDO-MORENO, J. et al. "Chemistry beyond the scale of exact diagonalization on a quantum-centric supercomputer". *Science Advances*, American Association for the Advancement of Science (AAAS), v. 11, n. 25, jun. 2025. ISSN 2375-2548. Availabe at: <http://dx.doi.org/10.1126/sciadv.adu9991>.

43 BARISON, S.; MORENO, J. R.; MOTTA, M. "Quantum-centric computation of molecular excited states with extended sample-based quantum diagonalization". *Quantum Science and Technology*, IOP Publishing, v. 10, n. 2, p. 025034, fev. 2025. ISSN 2058-9565. Availabe at: <http://dx.doi.org/10.1088/2058-9565/adb781>.

44 KALIAKIN, D. et al. "*Accurate quantum-centric simulations of supramolecular interactions*". 2024. Availabe at: <https://arxiv.org/abs/2410.09209>.

45 KALIAKIN, D. et al. "*Implicit solvent sample-based quantum diagonalization*". 2025. Availabe at: <https://arxiv.org/abs/2502.10189>.

46 YOSHIOKA, N. et al. "Krylov diagonalization of large many-body hamiltonians on a quantum processor". *Nature Communications*, Springer Science and Business Media LLC, v. 16, n. 1, jun. 2025. ISSN 2041-1723. Availabe at: <http://dx.doi.org/10.1038/s41467-025-59716-z>.

47 SHAJAN, A. et al. "*Towards quantum-centric simulations of extended molecules: sample-based quantum diagonalization enhanced with density matrix embedding theory*". 2024. Availabe at: <https://arxiv.org/abs/2411.09861>.

48 YU, J. et al. "*Quantum-Centric Algorithm for Sample-Based Krylov Diagonalization*". 2025. Availabe at: <https://arxiv.org/abs/2501.09702>.

49 ABRAMS, D. S.; LLOYD, S. "Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors". *Phys. Rev. Lett.*, American Physical Society, v. 83, p. 5162–5165, Dec 1999. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevLett.83.5162>.

50 PRESKILL, J. "Quantum Computing in the NISQ era and beyond". *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 2, p. 79, ago. 2018. ISSN 2521-327X. Availabe at: <https://doi.org/10.22331/q-2018-08-06-79>.

51 NELSON, J. S.; BACZEWSKI, A. D. "Assessment of quantum phase estimation protocols for early fault-tolerant quantum computers". *Phys. Rev. A*, American Physical Society, v. 110, p. 042420, Oct 2024. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevA.110.042420>.

52 SUN, Q.; CHAN, G. K.-L. "Quantum embedding theories". *Accounts of Chemical Research*, American Chemical Society (ACS), v. 49, n. 12, p. 2705–2712, nov. 2016. ISSN 1520-4898. Availabe at: <http://dx.doi.org/10.1021/acs.accounts.6b00356>.

53 KNIZIA, G.; CHAN, G. K.-L. "Density matrix embedding: A simple alternative to dynamical mean-field theory". *Physical Review Letters*, v. 109, p. 186404, 2012.

54 KEMPE, J.; KITAEV, A.; REGEV, O. "The complexity of the local hamiltonian problem". *SIAM Journal on Computing*, v. 35, n. 5, p. 1070–1097, 2006. Availabe at: <https://doi.org/10.1137/S0097539704445226>.

55 KITAEV, A.; SHEN, A. H.; VYALYI, M. N. "*Classical and Quantum Computation*". [S.l.]: American Mathematical Society, Providence, RI, 2002. v. 47. (Graduate Studies in Mathematics, v. 47).

56 BOOKATZ, A. D. "*QMA-complete problems*". 2013. Availabe at: <https://arxiv.org/abs/1212.6312>.

57 LUCAS, A. "Ising formulations of many NP problems". *Front. Phys.*, Frontiers Media SA, v. 2, 2014.

58 ICHIKAWA, K. et al. "*Optimal elemental configuration search in crystal using quantum approximate optimization algorithm*". 2025. Availabe at: <https://arxiv.org/abs/2503.09356>.

59 XU, Z. et al. "Quantum annealing-assisted lattice optimization". *Npj Comput. Mater.*, Springer Science and Business Media LLC, v. 11, n. 1, jan. 2025.

60 FARHI, E.; GOLDSTONE, J.; GUTMANN, S. "*A Quantum Approximate Optimization Algorithm*". 2014. Availabe at: <https://arxiv.org/abs/1411.4028>.

61 SPEZIALI, S. et al. "*Solving Sensor Placement Problems In Real Water Distribution Networks Using Adiabatic Quantum Computation*". 2021. Availabe at: <https://arxiv.org/abs/2108.04075>.

62 RIEFFEL, E. G. et al. "A case study in programming a quantum annealer for hard operational planning problems". *Quantum Information Processing*, Springer Science and Business Media LLC, v. 14, n. 1, p. 1–36, dez. 2014. ISSN 1573-1332. Availabe at: <http://dx.doi.org/10.1007/s11128-014-0892-x>.

63 FEYNMAN, R. P. "Simulating physics with computers". *International Journal of Theoretical Physics*, v. 21, n. 6/7, 1982.

64 Argonne Leadership Computing Facility. "*Materials Science*". 2023. <https://www.alcf.anl.gov/domains/materials-science>.

65 AUSTIN, B. et al. "*Nersc-10 workload analysis*". 2020. <https://portal.nersc.gov/project/m888/nersc10/workload/N10_Workload_Analysis.latest.pdf>.

66 CONNOR, C. "*Institutional Computing for Open Science and Engineering*". Los Alamos, New Mexico, USA, 2023.

67 Donostia International Physics Center Supercomputing Center. "*Supercomputing Center*". 2023. <https://dipc.ehu.eus/en/supercomputing-center>. [Online; accessed 8-December-2023].

68 SZABO, A.; OSTLUND, N. S. "*Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*". Mineola, NY: Dover Publications, 2012.

69 TUBMAN, N. M. et al. "A deterministic alternative to the full configuration interaction quantum monte carlo method". *The Journal of Chemical Physics*, v. 145, n. 4, p. 044112, 2016. Availabe at: <https://doi.org/10.1063/1.4955109>.

70 SCHRIBER, J. B.; EVANGELISTA, F. A. "Adaptive configuration interaction for computing challenging electronic excited states with tunable accuracy". *Journal of Chemical Theory and Computation*, v. 13, n. 11, p. 5354–5366, 2017. PMID: 28892621. Availabe at: <https://doi.org/10.1021/acs.jctc.7b00725>.

71 TUBMAN, N. M. et al. "Modern approaches to exact diagonalization and selected configuration interaction with the adaptive sampling ci method". *Journal of Chemical Theory and Computation*, v. 16, n. 4, p. 2139–2159, 2020. Availabe at: <https://doi.org/10.1021/acs.jctc.8b00536>.

72 MEJUTO-ZAERA, C.; TUBMAN, N. M.; WHALEY, K. B. "Dynamical mean field theory simulations with the adaptive sampling configuration interaction method". *Phys. Rev. B*, American Physical Society, v. 100, p. 125165, Sep 2019. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevB.100.125165>.

73 OLSEN, J.; JøRGENSEN, P.; SIMONS, J. "Passing the one-billion limit in full configuration-interaction (fci) calculations". *Chemical Physics Letters*, v. 169, n. 6, p. 463–472, 1990. ISSN 0009-2614. Availabe at: <https://www.sciencedirect.com/science/article/pii/000926149085633N>.

74 SOLOMONIK, E. et al. "A massively parallel tensor contraction framework for coupled-cluster computations". *Journal of Parallel and Distributed Computing*, Elsevier, v. 74, n. 12, p. 3176–3190, 2014.

75 KIM, J. et al. "Qmcpack: an open source ab initio quantum monte carlo package for the electronic structure of atoms, molecules and solids". *Journal of Physics: Condensed Matter*, IOP Publishing, v. 30, n. 19, p. 195901, apr 2018. Availabe at: <https://dx.doi.org/10.1088/1361-648X/aab9c3>.

76 WILLIAMS-YOUNG, D. B. et al. "A parallel, distributed memory implementation of the adaptive sampling configuration interaction method". *The Journal of Chemical Physics*, v. 158, n. 21, p. 214109, 06 2023. ISSN 0021-9606. Availabe at: <https://doi.org/10.1063/5.0148650>.

77 VALIEV, M. et al. "Nwchem: A comprehensive and scalable open-source solution for large scale molecular simulations". *Computer Physics Communications*, v. 181, n. 9, p. 1477–1489, 2010. ISSN 0010-4655. Availabe at: <https://www.sciencedirect.com/science/article/pii/S0010465510001438>.

78 FOYEVTSOVA, K. et al. "Ab initio quantum monte carlo calculations of spin superexchange in cuprates: The benchmarking case of ca_2cuo_3". *Phys. Rev. X*, American Physical Society, v. 4, p. 031003, Jul 2014. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevX.4.031003>.

79 HOLZMANN, M. et al. “Theory of finite size effects for electronic quantum monte carlo calculations”. *Phys. Rev. B*, American Physical Society, v. 94, p. 035126, Jul 2016. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevB.94.035126>.

80 BOOTH, G. H. et al. “Towards an exact description of electronic wavefunctions in real solids”. *Nature*, v. 493, n. 7432, p. 365–370, 2013.

81 ERIKSEN, J. J. et al. “The ground state electronic energy of benzene”. *The Journal of Physical Chemistry Letters*, v. 11, n. 20, p. 8922–8929, 2020. PMID: 33022176. Availabe at: <https://doi.org/10.1021/acs.jpclett.0c02621>.

82 CHANG, C.-C.; RUBENSTEIN, B. M.; MORALES, M. A. “Auxiliary-field-based trial wave functions in quantum monte carlo calculations”. *Phys. Rev. B*, American Physical Society, v. 94, p. 235144, Dec 2016. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevB.94.235144>.

83 VLČEK, V. et al. “Swift *gw* beyond 10,000 electrons using sparse stochastic compression”. *Phys. Rev. B*, American Physical Society, v. 98, n. 7, p. 075107, ago. 2018. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevB.98.075107>.

84 JONG, W. A. D. et al. “Utilizing high performance computing for chemistry: parallel computational chemistry”. *Physical Chemistry Chemical Physics*, Royal Society of Chemistry, v. 12, n. 26, p. 6896–6920, 2010.

85 TUBMAN, N. M. et al. “Prospects for release-node quantum Monte Carlo”. *The Journal of Chemical Physics*, v. 135, n. 18, p. 184109, 11 2011. ISSN 0021-9606.

86 LLOYD, S. “Universal quantum simulators”. *Science*, JSTOR, p. 1073–1078, 1996.

87 LOW, G. H.; CHUANG, I. L. “Hamiltonian simulation by qubitization”. *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 3, p. 163, 2019.

88 CHILDS, A. M. et al. “Theory of trotter error with commutator scaling”. *Phys. Rev. X*, American Physical Society, v. 11, p. 011020, Feb 2021. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevX.11.011020>.

89 HAAH, J. et al. “Quantum algorithm for simulating real time evolution of lattice hamiltonians”. In: *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*. [S.l.: s.n.], 2018. p. 350–360.

90 TRAN, M. C. et al. “Locality and digital quantum simulation of power-law interactions”. *Phys. Rev. X*, American Physical Society, v. 9, p. 031006, Jul 2019. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevX.9.031006>.

91 DU, J. et al. “Nmr implementation of a molecular hydrogen quantum simulation with adiabatic state preparation”. *Phys. Rev. Lett.*, American Physical Society, v. 104, p. 030502, 2010. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevLett.104.030502>.

92 LANYON, B. P. et al. “Towards quantum chemistry on a quantum computer”. *Nature Chemistry*, Nature Research, v. 2, n. 2, p. 106–111, 2010.

93 WANG, Y. et al. “Quantum simulation of helium hydride cation in a solid-state spin register”. *ACS Nano*, ACS Publications, v. 9, n. 8, p. 7769–7774, 2015.

94 SHEN, Y. et al. "Quantum implementation of the unitary coupled cluster for simulating molecular electronic structure". *Physical Review A*, APS, v. 95, n. 2, p. 020501, 2017.

95 PAESANI, S. et al. "Experimental bayesian quantum phase estimation on a silicon photonic chip". *Phys. Rev. Lett.*, American Physical Society, v. 118, p. 100503, 2017. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevLett.118.100503>.

96 HEMPEL, C. et al. "Quantum chemistry calculations on a trapped-ion quantum simulator". *Phys. Rev. X*, American Physical Society, v. 8, p. 031022, Jul 2018. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevX.8.031022>.

97 SANTAGATI, R. et al. "Witnessing eigenstates for quantum simulation of hamiltonian spectra". *Science Advances*, American Association for the Advancement of Science, v. 4, n. 1, 2018. Availabe at: <https://advances.sciencemag.org/content/4/1/eaap9646>.

98 DUMITRESCU, E. F. et al. "Cloud quantum computing of an atomic nucleus". *Phys. Rev. Lett.*, v. 120, p. 210501, 2018.

99 KOKAIL, C. et al. "Self-verifying variational quantum simulation of lattice models". *Nature*, v. 569, n. 7756, p. 355–360, 2019. Availabe at: <https://doi.org/10.1038/s41586-019-1177-4>.

100 KANDALA, A. et al. "Error mitigation extends the computational reach of a noisy quantum processor". *Nature*, v. 567, n. 7749, p. 491–495, 2019. Availabe at: <https://doi.org/10.1038/s41586-019-1040-7>.

101 GANZHORN, M. et al. "Gate-efficient simulation of molecular eigenstates on a quantum computer". *Physical Review Applied*, APS, v. 11, n. 4, p. 044092, 2019.

102 SAGASTIZABAL, R. et al. "Experimental error mitigation via symmetry verification in a variational quantum eigensolver". *Physical Review A*, APS, v. 100, n. 1, p. 010302, 2019.

103 MCCASKEY, A. J. et al. "Quantum chemistry as a benchmark for near-term quantum computers". *npj Quantum Information*, Nature Publishing Group, v. 5, n. 1, p. 1–8, 2019.

104 SMART, S. E.; MAZZIOTTI, D. A. "Quantum-classical hybrid algorithm using an error-mitigating n-representability condition to compute the mott metal-insulator transition". *Physical Review A*, APS, v. 100, n. 2, p. 022517, 2019.

105 NAM, Y. et al. "Ground-state energy estimation of the water molecule on a trapped-ion quantum computer". *npj Quantum Information*, v. 6, n. 1, p. 33, 2020. Availabe at: <https://doi.org/10.1038/s41534-020-0259-3>.

106 LIU, J. et al. "Representation learning via quantum neural tangent kernels". *PRX Quantum*, APS, v. 3, n. 3, p. 030323, 2022.

107 WANG, Q. et al. "Resource-optimized fermionic local-hamiltonian simulation on a quantum computer for quantum chemistry". *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 5, p. 509, 2021.

108 LIU, J. et al. “Analytic theory for the dynamics of wide quantum neural networks”. *Physical Review Letters*, APS, v. 130, n. 15, p. 150601, 2023.

109 ZHENG, H. et al. “Speeding up learning quantum states through group equivariant convolutional quantum ansätze”. *PRX Quantum*, APS, v. 4, n. 2, p. 020327, 2023.

110 WANG, Q. et al. “Ever more optimized simulations of fermionic systems on a quantum computer”. In: *2023 60th ACM/IEEE Design Automation Conference (DAC)*. [S.l.: s.n.], 2023. p. 1–6.

111 PAGANO, G. et al. “Quantum approximate optimization of the long-range ising model with a trapped-ion quantum simulator”. *Proceedings of the National Academy of Sciences*, v. 117, n. 41, p. 25396–25401, 2020. Availabe at: <https://www.pnas.org/doi/abs/10.1073/pnas.2006373117>.

112 LOTSHAW, P. C. et al. “Simulations of frustrated ising hamiltonians using quantum approximate optimization”. *Philosophical Transactions of the Royal Society A*, The Royal Society, v. 381, n. 2241, p. 20210414, 2023.

113 KITAEV, A. Y. “Quantum measurements and the abelian stabilizer problem”. *arxiv preprint, arXiv:quant-ph/9511026*, 1995. Availabe at: <https://doi.org/10.48550/arXiv.quant-ph/9511026>.

114 LIN, L.; TONG, Y. “Near-optimal ground state preparation”. *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 4, p. 372, dez. 2020. ISSN 2521-327X. Availabe at: <https://doi.org/10.22331/q-2020-12-14-372>.

115 LIN, L.; TONG, Y. “Heisenberg-limited ground-state energy estimation for early fault-tolerant quantum computers”. *PRX Quantum*, American Physical Society, v. 3, p. 010318, Feb 2022. Availabe at: <https://link.aps.org/doi/10.1103/PRXQuantum.3.010318>.

116 DONG, Y.; LIN, L.; TONG, Y. “Ground-state preparation and energy estimation on early fault-tolerant quantum computers via quantum eigenvalue transformation of unitary matrices”. *PRX Quantum*, American Physical Society, v. 3, p. 040305, Oct 2022. Availabe at: <https://link.aps.org/doi/10.1103/PRXQuantum.3.040305>.

117 DING, Z.; LIN, L. et al. “Single-ancilla ground state preparation via Lindbladians”. *arXiv preprint arXiv:2308.15676*, 2023.

118 MCCLEAN, J. R. et al. “Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states”. *Phys. Rev. A*, American Physical Society, v. 95, p. 042308, Apr 2017. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevA.95.042308>.

119 PARRISH, R. M.; MCMAHON, P. L. “Quantum filter diagonalization: Quantum eigendecomposition without full quantum phase estimation”. *arXiv*, 2019.

120 MOTTA, M. et al. “Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution”. *Nature Physics*, v. 16, n. 2, p. 205–210, 2020. Availabe at: <https://doi.org/10.1038/s41567-019-0704-4>.

121 TAKESHITA, T. et al. “Increasing the representation accuracy of quantum simulations of chemistry without extra quantum resources”. *Phys. Rev. X*, American Physical Society, v. 10, p. 011004, Jan 2020. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevX.10.011004>.

122 HUGGINS, W. J. et al. “A non-orthogonal variational quantum eigensolver”. *New Journal of Physics*, IOP Publishing, v. 22, n. 7, p. 073009, jul 2020. Availabe at: <https://doi.org/10.1088/1367-2630/ab867b>.

123 STAIR, N. H.; HUANG, R.; EVANGELISTA, F. A. “A multireference quantum Krylov algorithm for strongly correlated electrons”. *Journal of Chemical Theory and Computation*, American Chemical Society, v. 16, n. 4, p. 2236–2245, 04 2020. Availabe at: <https://doi.org/10.1021/acs.jctc.9b01125>.

124 URBANEK, M. et al. “Chemistry on quantum computers with virtual quantum subspace expansion”. *Journal of Chemical Theory and Computation*, American Chemical Society, v. 16, n. 9, p. 5425–5431, 09 2020. Availabe at: <https://doi.org/10.1021/acs.jctc.0c00447>.

125 COHN, J.; MOTTA, M.; PARRISH, R. M. “Quantum filter diagonalization with compressed double-factorized hamiltonians”. *PRX Quantum*, APS, v. 2, n. 4, p. 040352, Dec 2021. Availabe at: <https://link.aps.org/doi/10.1103/PRXQuantum.2.040352>.

126 YOSHIOKA, N. et al. “Generalized quantum subspace expansion”. *Phys. Rev. Lett.*, American Physical Society, v. 129, p. 020502, Jul 2022. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevLett.129.020502>.

127 SEKI, K.; YUNOKI, S. “Quantum power method by a superposition of time-evolved states”. *PRX Quantum*, American Physical Society, v. 2, p. 010333, Feb 2021.

128 CORTES, C. L.; GRAY, S. K. “Quantum Krylov subspace algorithms for ground- and excited-state energy estimation”. *Phys. Rev. A*, American Physical Society, v. 105, p. 022417, Feb 2022. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevA.105.022417>.

129 KLYMKO, K. et al. “Real-time evolution for ultracompact hamiltonian eigenstates on quantum hardware”. *PRX Quantum*, American Physical Society, v. 3, p. 020323, May 2022. Availabe at: <https://link.aps.org/doi/10.1103/PRXQuantum.3.020323>.

130 Baek, U. et al. “Say NO to Optimization: A Nonorthogonal Quantum Eigensolver”. *PRX Quantum*, APS, v. 4, n. 3, p. 030307, jul. 2023.

131 TKACHENKO, N. V. et al. “Quantum Davidson algorithm for excited states”. *Quantum Science and Technology*, IOP Publishing, v. 9, n. 3, p. 035012, apr 2024. Availabe at: <https://dx.doi.org/10.1088/2058-9565/ad3a97>.

132 LEE, G.; LEE, D.; HUH, J. “Sampling error analysis in quantum krylov subspace diagonalization”. *Quantum*, Verein zur Forderung des Open Access Publizierens in den Quantenwissenschaften, v. 8, p. 1477, set. 2024. ISSN 2521-327X. Availabe at: <http://dx.doi.org/10.22331/q-2024-09-19-1477>.

133 ZHANG, Z. et al. “Measurement-efficient quantum krylov subspace diagonalisation”. *Quantum*, Verein zur Forderung des Open Access Publizierens in den Quantenwissenschaften, v. 8, p. 1438, ago. 2024. ISSN 2521-327X. Availabe at: <http://dx.doi.org/10.22331/q-2024-08-13-1438>.

134 KIRBY, W.; MOTTA, M.; MEZZACAPO, A. "Exact and efficient Lanczos method on a quantum computer". *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 7, p. 1018, maio 2023. ISSN 2521-327X. Availabe at: <https://doi.org/10.22331/q-2023-05-23-1018>.

135 SHEN, Y. et al. "Real-Time Krylov Theory for Quantum Computing Algorithms". *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 7, p. 1066, jul. 2023. ISSN 2521-327X. Availabe at: <https://doi.org/10.22331/q-2023-07-25-1066>.

136 MOTTA, M. et al. "Subspace methods for electronic structure simulations on quantum computers". *Electronic Structure*, IOP Publishing, v. 6, n. 1, p. 013001, mar 2024. Availabe at: <https://dx.doi.org/10.1088/2516-1075/ad3592>.

137 SHEN, Y. et al. "Estimating eigenenergies from quantum dynamics: A unified noise-resilient measurement-driven approach". In: *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*. [S.l.: s.n.], 2023. v. 02, p. 302–303.

138 EPPERLY, E. N.; LIN, L.; NAKATSUKASA, Y. "A theory of quantum subspace diagonalization". *SIAM Journal on Matrix Analysis and Applications*, v. 43, n. 3, p. 1263–1290, 2022.

139 KIRBY, W. "Analysis of quantum krylov algorithms with errors". *Quantum*, Verein zur Forderung des Open Access Publizierens in den Quantenwissenschaften, v. 8, p. 1457, ago. 2024. ISSN 2521-327X. Availabe at: <http://dx.doi.org/10.22331/q-2024-08-29-1457>.

140 LINDBLAD, G. "On the generators of quantum dynamical semigroups". *Communications in Mathematical Physics*, Springer, v. 48, p. 119–130, 1976.

141 WANG, H.; ASHHAB, S.; NORI, F. "Quantum algorithm for simulating the dynamics of an open quantum system". *Physical Review A*, APS, v. 83, n. 6, p. 062317, 2011.

142 KLIESCH, M. et al. "Dissipative quantum church-turing theorem". *Physical review letters*, APS, v. 107, n. 12, p. 120501, 2011.

143 DING, Z.; LI, X.; LIN, L. "Simulating open quantum systems using Hamiltonian simulations". *arXiv preprint arXiv:2311.15533*, 2023.

144 SWEKE, R. et al. "Digital quantum simulation of many-body non-markovian dynamics". *Physical Review A*, APS, v. 94, n. 2, p. 022317, 2016.

145 CHILDS, A. M.; LI, T. "Efficient simulation of sparse Markovian quantum dynamics". *Quantum Info. Comput.*, Rinton Press, Incorporated, Paramus, NJ, v. 17, n. 11–12, p. 901–947, sep 2017. ISSN 1533-7146.

146 CLEVE, R.; WANG, C. "Efficient Quantum Algorithms for Simulating Lindblad Evolution". In: CHATZIGIANNAKIS, I. et al. (Ed.). *44th International Colloquium on Automata, Languages, and Programming (ICALP 2017)*. Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2017. (Leibniz International Proceedings in Informatics (LIPIcs), v. 80), p. 17:1–17:14. ISBN 978-3-95977-041-5. ISSN 1868-8969. Availabe at: <https://drops-dev.dagstuhl.de/entities/document/10.4230/LIPIcs.ICALP.2017.17>.

147 KAMAKARI, H. et al. “Digital quantum simulation of open quantum systems using quantum imaginary–time evolution”. *PRX Quantum*, APS, v. 3, n. 1, p. 010320, 2022.

148 POCRNIC, M.; SEGAL, D.; WIEBE, N. “Quantum simulation of Lindbladian dynamics via repeated interactions”. *arXiv preprint arXiv:2312.05371*, 2023.

149 PATEL, D.; WILDE, M. M. “Wave matrix lindbladization i: Quantum programs for simulating markovian dynamics”. *Open Systems & Information Dynamics*, World Scientific, v. 30, n. 02, p. 2350010, 2023.

150 PATEL, D.; WILDE, M. M. “Wave matrix lindbladization ii: General lindbladians, linear combinations, and polynomials”. *Open Systems & Information Dynamics*, World Scientific, v. 30, n. 03, p. 2350014, 2023.

151 GUIMARÃES, J. D. et al. “Noise-assisted digital quantum simulation of open systems using partial probabilistic error cancellation”. *PRX Quantum*, APS, v. 4, n. 4, p. 040329, 2023.

152 OLMOS, B.; LESANOVSKY, I.; GARRAHAN, J. P. “Facilitated spin models of dissipative quantum glasses”. *Physical Review Letters*, APS, v. 109, n. 2, p. 020403, 2012.

153 MAY, V.; KÜHN, O. “*Charge and energy transfer dynamics in molecular systems*”. [S.l.]: John Wiley & Sons, 2023.

154 NITZAN, A. “*Chemical dynamics in condensed phases: relaxation, transfer and reactions in condensed molecular systems*”. [S.l.]: Oxford university press, 2006.

155 CHEN, C.-F. et al. “*Quantum Thermal State Preparation*”. 2023.

156 CUBITT, T. S. “Dissipative ground state preparation and the dissipative quantum eigensolver”. *arXiv preprint arXiv:2303.11962*, 2023.

157 HUBISZ, J.; SAMBASIVAM, B.; UNMUTH-YOCKEY, J. “Quantum algorithms for open lattice field theory”. *Physical Review A*, APS, v. 104, n. 5, p. 052420, 2021.

158 SCHLIMGEN, A. W. et al. “Quantum simulation of the Lindblad equation using a unitary decomposition of operators”. *Physical Review Research*, APS, v. 4, n. 2, p. 023216, 2022.

159 GALLINA, F.; BRUSCHI, M.; FRESCH, B. “Strategies to simulate dephasing-assisted quantum transport on digital quantum computers”. *New Journal of Physics*, IOP Publishing, v. 24, n. 2, p. 023039, fev. 2022. ISSN 1367-2630. Availabe at: <http://dx.doi.org/10.1088/1367-2630/ac512f>.

160 AVDIC, I.; SAGER-SMITH, L. M.; MAZZIOTTI, D. A. “Open quantum system violates generalized pauli constraints on quantum device”. *Communications Physics*, Springer Science and Business Media LLC, v. 6, n. 1, jul. 2023. ISSN 2399-3650. Availabe at: <http://dx.doi.org/10.1038/s42005-023-01295-w>.

161 AVDIC, I. et al. “Quantum sensing using multiqubit quantum systems and the pauli polytope”. *Physical Review Research*, American Physical Society (APS), v. 5, n. 4, out. 2023. ISSN 2643-1564. Availabe at: <http://dx.doi.org/10.1103/PhysRevResearch.5.043097>.

162 HU, Z.; XIA, R.; KAIS, S. "A quantum algorithm for evolving open quantum dynamics on quantum computing devices". *Scientific Reports*, Springer Science and Business Media LLC, v. 10, n. 1, fev. 2020. ISSN 2045-2322. Availabe at: <http://dx.doi.org/10.1038/s41598-020-60321-x>.

163 PLENIO, M. B.; HUELGA, S. F. "Dephasing-assisted transport: quantum networks and biomolecules". *New Journal of Physics*, IOP Publishing, v. 10, n. 11, p. 113019, nov. 2008. ISSN 1367-2630. Availabe at: <http://dx.doi.org/10.1088/1367-2630/10/11/113019>.

164 POULIN, D.; WOCJAN, P. "Sampling from the thermal quantum gibbs state and evaluating partition functions with a quantum computer". *Physical Review Letters*, American Physical Society (APS), v. 103, n. 22, nov. 2009. ISSN 1079-7114. Availabe at: <http://dx.doi.org/10.1103/PhysRevLett.103.220502>.

165 CHOWDHURY, A. N.; SOMMA, R. D. "Quantum algorithms for Gibbs sampling and hitting-time estimation". *Quantum Info. Comput.*, Rinton Press, Incorporated, Paramus, NJ, v. 17, n. 1–2, p. 41–64, Feb 2017. ISSN 1533-7146.

166 COTLER, J. et al. "Quantum virtual cooling". *Physical Review X*, American Physical Society (APS), v. 9, n. 3, jul. 2019. ISSN 2160-3308. Availabe at: <http://dx.doi.org/10.1103/PhysRevX.9.031013>.

167 TEMME, K. et al. "Quantum metropolis sampling". *Nature*, Springer Science and Business Media LLC, v. 471, n. 7336, p. 87–90, mar. 2011. ISSN 1476-4687. Availabe at: <http://dx.doi.org/10.1038/nature09770>.

168 YUNG, M.-H.; ASPURU-GUZIK, A. "A quantum–quantum metropolis algorithm". *Proceedings of the National Academy of Sciences*, Proceedings of the National Academy of Sciences, v. 109, n. 3, p. 754–759, jan. 2012. ISSN 1091-6490. Availabe at: <http://dx.doi.org/10.1073/pnas.1111758109>.

169 WOCJAN, P.; TEMME, K. "Szegedy walk unitaries for quantum maps". *Communications in Mathematical Physics*, Springer Science and Business Media LLC, v. 402, n. 3, p. 3201–3231, jul. 2023. ISSN 1432-0916. Availabe at: <http://dx.doi.org/10.1007/s00220-023-04797-4>.

170 DAVIES, E. "*Quantum Theory of Open Systems*". Academic Press, 1976. ISBN 9780122061509. Availabe at: <https://books.google.com/books?id=I5kuAAAAIAAJ>.

171 RALL, P.; WANG, C.; WOCJAN, P. "Thermal state preparation via rounding promises". *Quantum*, Verein zur Forderung des Open Access Publizierens in den Quantenwissenschaften, v. 7, p. 1132, out. 2023. ISSN 2521-327X. Availabe at: <http://dx.doi.org/10.22331/q-2023-10-10-1132>.

172 CHEN, C.-F.; KASTORYANO, M. J.; GILYéN, A. "*An efficient and exact noncommutative quantum Gibbs sampler*". 2023.

173 CHEN, C.-F. et al. "*Local minima in quantum systems*". 2023.

174 HARROW, A. W.; HASSIDIM, A.; LLOYD, S. "Quantum algorithm for linear systems of equations". *Physical review letters*, APS, v. 103, n. 15, p. 150502, 2009.

175 MONTANARO, A.; PALLISTER, S. "Quantum algorithms and the finite element method". *Phys. Rev. A*, American Physical Society, v. 93, p. 032324, Mar 2016. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevA.93.032324>.

176 LINDEN, N.; MONTANARO, A.; SHAO, C. "Quantum vs. classical algorithms for solving the heat equation". *Communications in Mathematical Physics*, Springer Berlin Heidelberg, v. 395, n. 2, p. 601–641, oct 2022. ISSN 0010-3616.

177 ZHANG, J.; FENG, F.; ZHANG, Q. J. "Quantum method for finite element simulation of electromagnetic problems". In: *2021 IEEE MTT-S International Microwave Symposium (IMS)*. [S.l.: s.n.], 2021. p. 120–123.

178 CHILDS, A. M.; KOTHARI, R.; SOMMA, R. D. "Quantum algorithm for systems of linear equations with exponentially improved dependence on precision". *SIAM Journal on Computing*, v. 46, n. 6, p. 1920–1950, 2017. Availabe at: <https://doi.org/10.1137/16M1087072>.

179 FANG, D.; LIN, L.; TONG, Y. "Time-marching based quantum solvers for time-dependent linear differential equations". *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 7, p. 955, mar. 2023. ISSN 2521-327X. Availabe at: <https://doi.org/10.22331/q-2023-03-20-955>.

180 KROVI, H. "Improved quantum algorithms for linear and nonlinear differential equations". *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 7, p. 913, fev. 2023. ISSN 2521-327X. Availabe at: <https://doi.org/10.22331/q-2023-02-02-913>.

181 AN, D. et al. "*A theory of quantum differential equation solvers: limitations and fast-forwarding*". 2023.

182 LEWIS, D. et al. "Limitations for quantum algorithms to solve turbulent and chaotic systems". *Quantum*, Verein zur Forderung des Open Access Publizierens in den Quantenwissenschaften, v. 8, p. 1509, out. 2024. ISSN 2521-327X. Availabe at: <http://dx.doi.org/10.22331/q-2024-10-24-1509>.

183 JIN, S.; LIU, N.; YU, Y. "Quantum simulation of partial differential equations: Applications and detailed analysis". *Phys. Rev. A*, American Physical Society, v. 108, p. 032603, Sep 2023. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevA.108.032603>.

184 WU, L.-A.; LIDAR, D. "Qubits as parafermions". *Journal of Mathematical Physics*, American Institute of Physics, v. 43, n. 9, p. 4506–4525, 2002.

185 SOMMA, R. D. "Quantum computation, complexity, and many-body physics". *arXiv preprint arXiv:quant-ph/0512209*, 2005.

186 MACRIDIN, A. et al. "Electron-phonon systems on a universal quantum computer". *Phys. Rev. Lett.*, American Physical Society, v. 121, p. 110504, Sep 2018. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevLett.121.110504>.

187 MCARDLE, S. et al. "Digital quantum simulation of molecular vibrations". *Chem. Sci.*, The Royal Society of Chemistry, v. 10, p. 5725–5735, 2019. Availabe at: <http://dx.doi.org/10.1039/C9SC01313J>.

188 SAWAYA, N. P. D.; HUH, J. "Quantum algorithm for calculating molecular vibronic spectra". *The Journal of Physical Chemistry Letters*, American Chemical Society, v. 10, n. 13, p. 3586–3591, 07 2019. Availabe at: <https://doi.org/10.1021/acs.jpclett.9b01117>.

189 KAN, A.; NAM, Y. "Lattice quantum chromodynamics and electrodynamics on a universal quantum computer". *arXiv preprint arXiv:2107.12769*, 2021.

190 KRESHCHUK, M. et al. "Quantum simulation of quantum field theory in the light-front formulation". *Phys. Rev. A*, American Physical Society, v. 105, p. 032418, Mar 2022. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevA.105.032418>.

191 KAN, A.; NAM, Y. "Simulating lattice quantum electrodynamics on a quantum computer". *Quantum Science and Technology*, IOP Publishing, v. 8, n. 1, p. 015008, 2022.

192 JORDAN, P.; WIGNER, E. "Über das paulische äquivalenzverbot". *Zeitschrift für Physik*, v. 47, n. 9, p. 631–651, 1928. Availabe at: <https://doi.org/10.1007/BF01331938>.

193 BRAVYI, S. B.; KITAEV, A. Y. "Fermionic quantum computation". *Annals of Physics*, Elsevier, v. 298, n. 1, p. 210–226, 2002.

194 VERSTRAETE, F.; CIRAC, J. I. "Mapping local hamiltonians of fermions to local hamiltonians of spins". *Journal of Statistical Mechanics: Theory and Experiment*, IOP Publishing, v. 2005, n. 09, p. P09012–P09012, sep 2005. Availabe at: <https://doi.org/10.1088/1742-5468/2005/09/p09012>.

195 SEELEY, J. T.; RICHARD, M. J.; LOVE, P. J. "The bravyi-kitaev transformation for quantum computation of electronic structure". *The Journal of Chemical Physics*, v. 137, n. 22, p. 224109, 2012. Availabe at: <https://doi.org/10.1063/1.4768229>.

196 BRAVYI, S. et al. "Tapering off qubits to simulate fermionic Hamiltonians". *arXiv preprint, arXiv:1701.08213*, 2017. Availabe at: <https://arxiv.org/abs/1701.08213>.

197 SETIA, K.; WHITFIELD, J. D. "Bravyi-Kitaev Superfast simulation of electronic structure on a quantum computer". *The Journal of Chemical Physics*, v. 148, n. 16, p. 164104, 04 2018. ISSN 0021-9606. Availabe at: <https://doi.org/10.1063/1.5019371>.

198 STEUDTNER, M.; WEHNER, S. "Fermion-to-qubit mappings with varying resource requirements for quantum simulation". *New Journal of Physics*, IOP Publishing, v. 20, n. 6, p. 063010, June 2018. Availabe at: <https://doi.org/10.1088/1367-2630/aac54f>.

199 STEUDTNER, M. "*Methods to simulate fermions on quantum computers with hardware limitations*". Tese (Doutorado) — Leiden University, 2019. Availabe at: <https://hdl.handle.net/1887/80413>.

200 DERBY, C. et al. "Compact fermion to qubit mappings". *Phys. Rev. B*, American Physical Society, v. 104, p. 035118, Jul 2021. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevB.104.035118>.

201 DERBY, C.; KLASSEN, J. "A compact fermion to qubit mapping part 2: Alternative lattice geometries". *arXiv preprint, arXiv:2101.10735*, 2021. Availabe at: <https://arxiv.org/abs/2101.10735>.

202 KIRBY, W. et al. "Second-quantized fermionic operators with polylogarithmic qubit and gate complexity". *PRX Quantum*, American Physical Society, v. 3, p. 020351, Jun 2022. Availabe at: <https://link.aps.org/doi/10.1103/PRXQuantum.3.020351>.

203 CHIEN, R. W.; KLASSEN, J. "Optimizing fermionic encodings for both hamiltonian and hardware". *arXiv preprint, arXiv:2210.05652*, 2021. Availabe at: <https://arxiv.org/abs/2210.05652>.

204 NAM, Y. et al. "Automated optimization of large quantum circuits with continuous parameters". *npj Quantum Information*, Nature Publishing Group UK London, v. 4, n. 1, p. 23, 2018.

205 MILLER, A. et al. "Bonsai algorithm: Grow your own fermion-to-qubit mappings". *PRX Quantum*, American Physical Society, v. 4, p. 030314, Aug 2023. Availabe at: <https://link.aps.org/doi/10.1103/PRXQuantum.4.030314>.

206 CHEN, Y.-A.; XU, Y. "Equivalence between fermion-to-qubit mappings in two spatial dimensions". *PRX Quantum*, American Physical Society, v. 4, p. 010326, Mar 2023. Availabe at: <https://link.aps.org/doi/10.1103/PRXQuantum.4.010326>.

207 MARTÍNEZ-MARTÍNEZ, L. A.; YEN, T.-C.; IZMAYLOV, A. F. "Assessment of various hamiltonian partitionings for the electronic structure problem on a quantum computer using the trotter approximation". *Quantum*, v. 7, p. 1086, 2023.

208 SHANGNAN, Z. "Quantum semi-supervised learning with quantum supremacy". *arXiv preprint arXiv:2110.02343*, 2021.

209 CHILDS, A. M. et al. "Toward the first quantum simulation with quantum speedup". *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 115, n. 38, p. 9456–9461, 2018.

210 NAM, Y.; MASLOV, D. "Low-cost quantum circuits for classically intractable instances of the hamiltonian dynamics simulation problem". *npj Quantum Information*, Nature Publishing Group UK London, v. 5, n. 1, p. 44, 2019.

211 BERRY, D. W. et al. "Efficient quantum algorithms for simulating sparse hamiltonians". *Communications in Mathematical Physics*, v. 270, n. 2, p. 359–371, 2007. Availabe at: <https://doi.org/10.1007/s00220-006-0150-x>.

212 TRANTER, A. et al. "Ordering of trotterization: Impact on errors in quantum simulation of electronic structure". *Entropy*, v. 21, n. 12, 2019. ISSN 1099-4300. Availabe at: <https://www.mdpi.com/1099-4300/21/12/1218>.

213 BERRY, D. W.; CHILDS, A. M. "Black-box hamiltonian simulation and unitary implementation". *Quantum Info. Comput.*, Rinton Press, Incorporated, Paramus, NJ, v. 12, n. 1-2, p. 29–62, jan. 2012. ISSN 1533-7146. Availabe at: <http://dl.acm.org/citation.cfm?id=2231036.2231040>.

214 BERRY, D. W.; CHILDS, A. M.; KOTHARI, R. "Hamiltonian simulation with nearly optimal dependence on all parameters". In: IEEE. *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on*. [S.l.], 2015. p. 792–809.

215 BERRY, D. W. et al. "Simulating hamiltonian dynamics with a truncated taylor series". *Physical review letters*, APS, v. 114, n. 9, p. 090502, 2015.

216 BERRY, D. W. et al. “Qubitization of arbitrary basis quantum chemistry leveraging sparsity and low rank factorization”. *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 3, p. 208, 2019.

217 CHILDS, A. M.; WIEBE, N. “Hamiltonian simulation using linear combinations of unitary operations”. *Quantum Info. Comput.*, Rinton Press, Incorporated, Paramus, NJ, v. 12, n. 11–12, p. 901–924, nov 2012. ISSN 1533-7146. Availabe at: <http://dl.acm.org/citation.cfm?id=2481569.2481570>.

218 LEE, J. et al. “Even more efficient quantum computations of chemistry through tensor hypercontraction”. *PRX Quantum*, American Physical Society, v. 2, p. 030305, Jul 2021. Availabe at: <https://link.aps.org/doi/10.1103/PRXQuantum.2.030305>.

219 LOAIZA, I. et al. “Reducing molecular electronic hamiltonian simulation cost for linear combination of unitaries approaches”. *Quantum Sci. Technol.*, v. 8, p. 035019, 2023.

220 LOAIZA, I.; IZMAYLOV, A. F. “Block-invariant symmetry shift: Preprocessing technique for second-quantized hamiltonians to improve their decompositions to linear combination of unitaries”. *Journal of Chemical Theory and Computation*, American Chemical Society, v. 19, n. 22, p. 8201–8209, 11 2023. Availabe at: <https://doi.org/10.1021/acs.jctc.3c00912>.

221 CHOI, S.; IZMAYLOV, A. F. “Measurement optimization techniques for excited electronic states in near-term quantum computing algorithms”. *J. Chem. Theory Comput. (arXiv:2302.11421)*, v. 19, n. 11, p. 3184–3193, 2023.

222 WANG, G. et al. “Minimizing estimation runtime on noisy quantum computers”. *PRX Quantum*, American Physical Society, v. 2, p. 010346, Mar 2021. Availabe at: <https://link.aps.org/doi/10.1103/PRXQuantum.2.010346>.

223 HUANG, H.-Y.; KUENG, R.; PRESKILL, J. “Predicting many properties of a quantum system from very few measurements”. *Nat. Phys.*, Nature Publishing Group, v. 16, n. 10, p. 1050–1057, 2020. Availabe at: <https://doi.org/10.1038/s41567-020-0932-7>.

224 ZHAO, A.; RUBIN, N. C.; MIYAKE, A. “Fermionic partial tomography via classical shadows”. *Phys. Rev. Lett.*, American Physical Society, v. 127, p. 110504, Sep 2021. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevLett.127.110504>.

225 O’GORMAN, B. “*Fermionic tomography and learning*”. 2022.

226 HUANG, H.-Y.; KUENG, R.; PRESKILL, J. “Efficient estimation of Pauli observables by derandomization”. *Phys. Rev. Lett.*, American Physical Society, v. 127, p. 030503, Jul 2021. Availabe at: <https://doi.org/10.1103/PhysRevLett.127.030503>.

227 HADFIELD, C. et al. “Measurements of quantum hamiltonians with locally-biased classical shadows”. *Communications in Mathematical Physics*, Springer, v. 391, n. 3, p. 951–967, 2022. Availabe at: <https://doi.org/10.1007/s00220-022-04343-8>.

228 HADFIELD, C. “Adaptive Pauli shadows for energy estimation”. *arXiv:2105.12207*, 2021. Availabe at: <https://arxiv.org/abs/2105.12207>.

229 LUKENS, J. M.; LAW, K. J.; BENNINK, R. S. “A Bayesian analysis of classical shadows”. *npj Quantum Information*, Nature Publishing Group, v. 7, n. 1, p. 1–10, 2021.

230 SHLOSBERG, A. et al. “Adaptive estimation of quantum observables”. *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 7, p. 906, 1 2023. ISSN 2521-327X. Availabe at: <https://doi.org/10.22331/q-2023-01-26-906>.

231 KOH, D. E.; GREWAL, S. “Classical shadows with noise”. *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 6, p. 776, 2022.

232 ELBEN, A. et al. “The randomized measurement toolbox”. *Nature Reviews Physics*, v. 5, n. 1, p. 9–24, 2023. Availabe at: <https://doi.org/10.1038/s42254-022-00535-2>.

233 HILLMICH, S. et al. “Decision diagrams for quantum measurements with shallow circuits”. In: *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)*. Los Alamitos, CA, USA: IEEE Computer Society, 2021. p. 24–34. Availabe at: <https://doi.ieeecomputersociety.org/10.1109/QCE52317.2021.00018>.

234 DUTT, A. et al. “Practical benchmarking of randomized measurement methods for quantum chemistry hamiltonians”. *arXiv*, 2023.

235 GARCÍA-PÉREZ, G. et al. “Learning to measure: Adaptive informationally complete generalized measurements for quantum algorithms”. *PRX Quantum*, American Physical Society, v. 2, p. 040342, Nov 2021. Availabe at: <https://link.aps.org/doi/10.1103/PRXQuantum.2.040342>.

236 ACHARYA, A.; SAHA, S.; SENGUPTA, A. M. “Shadow tomography based on informationally complete positive operator-valued measure”. *Phys. Rev. A*, American Physical Society, v. 104, p. 052418, 11 2021. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevA.104.052418>.

237 GLOS, A. et al. “*Adaptive POVM implementations and measurement error mitigation strategies for near-term quantum devices*”. 2022.

238 FILIPPOV, S. et al. “*Scalable tensor-network error mitigation for near-term quantum computing*”. 2023.

239 MOTTA, M. et al. “Low rank representations for quantum simulation of electronic structure”. *npj Quantum Information*, Nature Publishing Group UK London, v. 7, n. 1, p. 83, 2021.

240 YEN, T.-C.; IZMAYLOV, A. F. “Cartan subalgebra approach to efficient measurements of quantum observables”. *PRX Quantum*, v. 2, p. 040320, 2021.

241 CHOI, S.; LOAIZA, I.; IZMAYLOV, A. F. “Fluid fermionic fragments for optimizing quantum measurements of electronic hamiltonians in the variational quantum eigensolver”. *Quantum*, v. 7, p. 889, 2023.

242 CHOI, S.; YEN, T.-C.; IZMAYLOV, A. F. “Improving quantum measurements by introducing “ghost” pauli products”. *J. Chem. Theory Comput. (arXiv:2208.06563)*, v. 18, p. 7394, 2022.

243 YEN, T.-C.; GANESHRAM, A.; IZMAYLOV, A. F. “Deterministic improvements of quantum measurements with grouping of compatible operators, non-local transformations, and covariance estimates”. *npj Quantum Inf*, v. 9, n. 14, 2023.

244 CAI, Z. "Quantum error mitigation using symmetry expansion". *Quantum*, Verein zur Forderung des Open Access Publizierens in den Quantenwissenschaften, v. 5, p. 548, set. 2021. ISSN 2521-327X. Availabe at: <http://dx.doi.org/10.22331/q-2021-09-21-548>.

245 CAI, Z. et al. "Quantum error mitigation". *Rev. Mod. Phys.*, American Physical Society, v. 95, p. 045005, Dec 2023. Availabe at: <https://link.aps.org/doi/10.1103/RevModPhys.95.045005>.

246 CADE, C. et al. "Strategies for solving the fermi-hubbard model on near-term quantum computers". *Phys. Rev. B*, American Physical Society, v. 102, p. 235122, Dec 2020. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevB.102.235122>.

247 CAI, Z. "Resource estimation for quantum variational simulations of the hubbard model". *Phys. Rev. Appl.*, American Physical Society, v. 14, p. 014059, Jul 2020. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevApplied.14.014059>.

248 CLINTON, L. et al. "Towards near-term quantum simulation of materials". *Nature Communications*, v. 15, n. 1, p. 211, 2024. Availabe at: <https://doi.org/10.1038/s41467-023-43479-6>.

249 WOLFOWICZ, G. et al. "Quantum guidelines for solid-state spin defects". *Nature Reviews Materials*, Nature Publishing Group UK London, v. 6, n. 10, p. 906–925, 2021.

250 GUJARATI, T. P. et al. "Quantum computation of reactions on surfaces using local embedding". *npj Quantum Information*, Nature Publishing Group UK London, v. 9, n. 1, p. 88, 2023.

251 SARKAR, A. et al. "Understanding Antiferromagnetic and Ligand Field Effects on Spin Crossover in a Triple-Decker Dimeric Cr(II) Complex". *Journal of the American Chemical Society*, American Chemical Society, v. 145, n. 41, p. 22394–22402, oct 2023. ISSN 0002-7863. Availabe at: <https://doi.org/10.1021/jacs.3c05277>.

252 MAZE, J. R. et al. "Properties of nitrogen-vacancy centers in diamond: the group theoretic approach". *New Journal of Physics*, IOP Publishing, v. 13, n. 2, p. 025025, 2011.

253 MITRA, A. et al. "Excited states of crystalline point defects with multireference density matrix embedding theory". *J. Phys. Chem. Lett.*, ACS Publications, v. 12, n. 48, p. 11688–11694, 2021.

254 HALDAR, S. et al. "Local excitations of a charged nitrogen vacancy in diamond with multireference density matrix embedding theory". *J. Phys. Chem. Lett.*, ACS Publications, v. 14, p. 4273–4280, 2023.

255 VERMA, S. et al. "Optical Properties of Neutral F Centers in Bulk MgO with Density Matrix Embedding". *Journal of Physical Chemistry Letters*, American Chemical Society, p. 7703–7710, aug 2023. ISSN 19487185. Availabe at: <https://doi.org/10.1021/acs.jpclett.3c01875>.

256 DOHERTY, M. W. et al. "The nitrogen-vacancy colour centre in diamond". *Physics Reports*, Elsevier, v. 528, n. 1, p. 1–45, 2013.

257 Sauza-de la Vega, A. et al. "Multiconfiguration Pair-Density Functional Theory for Chromium(IV) Molecular Qubits". *JACS Au*, American Chemical Society, v. 2, n. 9, p. 2029–2037, sep 2022. Availabe at: <https://doi.org/10.1021/jacsau.2c00306>.

258 BAYLISS, S. L. et al. "Optically addressable molecular spins for quantum information processing". *Science*, American Association for the Advancement of Science, v. 370, n. 6522, p. 1309–1312, dec 2020. Availabe at: <https://doi.org/10.1126/science.abb9352>.

259 MA, H.; GOVONI, M.; GALLI, G. "Quantum simulations of materials on near-term quantum computers". *npj Computational Materials*, Nature Publishing Group UK London, v. 6, n. 1, p. 85, 2020.

260 SHENG, N. et al. "Green's function formulation of quantum defect embedding theory". *Journal of Chemical Theory and Computation*, ACS Publications, v. 18, n. 6, p. 3512–3522, 2022.

261 ROOS, B. O.; TAYLOR, P. R.; SIGBAHN, P. E. "A complete active space scf method (casscf) using a density matrix formulated super-ci approach". *Chem. Phys.*, v. 48, n. 2, p. 157 – 173, 1980. ISSN 0301-0104. Availabe at: <http://www.sciencedirect.com/science/article/pii/0301010480800450>.

262 SIEGBAHN, P. E. M. et al. "The complete active space scf (casscf) method in a newton–raphson formulation with application to the hno molecule". *J. Chem. Phys.*, v. 74, n. 4, p. 2384–2396, 1981. Availabe at: <https://doi.org/10.1063/1.441359>.

263 SIEGBAHN, P. et al. "A comparison of the super-CI and the newton-raphson scheme in the complete active space SCF method". *Phys. Scr.*, IOP Publishing, v. 21, n. 3-4, p. 323–327, jan 1980. Availabe at: <https://doi.org/10.1088/0031-8949/21/3-4/014>.

264 WHITE, S. R. "Density matrix formulation for quantum renormalization groups". *Physical Review Letters*, American Physical Society, v. 69, n. 19, p. 2863–2866, nov 1992. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevLett.69.2863>.

265 KNIZIA, G.; CHAN, G. K.-L. "Density matrix embedding: A strong-coupling quantum embedding theory". *Journal of Chemical Theory and Computation*, American Chemical Society, v. 9, n. 3, p. 1428–1432, 03 2013. Availabe at: <https://doi.org/10.1021/ct301044e>.

266 WOUTERS, S. et al. "A practical guide to density matrix embedding theory in quantum chemistry". *Journal of Chemical Theory and Computation*, American Chemical Society, v. 12, n. 6, p. 2706–2719, 06 2016. Availabe at: <https://doi.org/10.1021/acs.jctc.6b00316>.

267 PHAM, H. Q.; BERNALES, V.; GAGLIARDI, L. "Can density matrix embedding theory with the complete activate space self-consistent field solver describe single and double bond breaking in molecular systems?" *Journal of Chemical Theory and Computation*, v. 14, p. 1960–1968, 2018.

268 HERMES, M. R.; GAGLIARDI, L. "Multiconfigurational Self-Consistent Field Theory with Density Matrix Embedding: The Localized Active Space Self-Consistent Field Method". *Journal of Chemical Theory and Computation*, American Chemical Society, v. 15, n. 2, p. 972–986, feb 2019. ISSN 1549-9618. Availabe at: <https://doi.org/10.1021/acs.jctc.8b01009>.

269 PHAM, H. Q.; HERMES, M. R.; GAGLIARDI, L. "Periodic electronic structure calculations with the density matrix embedding theory". *Journal of Chemical Theory and Computation*, ACS Publications, v. 16, n. 1, p. 130–140, 2019.

270 WOUTERS, S.; JIMéNEZ-HOYOS, C. A.; CHAN, G. K. "Five years of density matrix embedding theory". In: _____. *Fragmentation*. John Wiley & Sons, Ltd, 2017. cap. 8, p. 227–243. ISBN 9781119129271. Availabe at: <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119129271.ch8>.

271 CUI, Z.-H.; ZHU, T.; CHAN, G. K.-L. "Efficient implementation of ab initio quantum embedding in periodic systems: Density matrix embedding theory". *Journal of Chemical Theory and Computation*, ACS Publications, v. 16, n. 1, p. 119–129, 2020.

272 LAU, B. T. G.; KNIZIA, G.; BERKELBACH, T. C. "Regional embedding enables high-level quantum chemistry for surface science". *The Journal of Physical Chemistry Letters*, v. 12, p. 1104–1109, 2021.

273 CUI, Z.-H. et al. "Systematic electronic structure in the cuprate parent state from quantum many-body simulations". *Science*, v. 377, p. 1192–1198, 2022.

274 MITRA, A.; HERMES, M. R.; GAGLIARDI, L. "Density Matrix Embedding Using Multiconfiguration Pair-Density Functional Theory". *Journal of Chemical Theory and Computation*, American Chemical Society, v. 19, p. 3498–3508, jun 2023. ISSN 15499626. Availabe at: <https://doi.org/10.1021/acs.jctc.3c00247>.

275 SUN, Q.; CHAN, G. K.-L. "Exact and optimal quantum mechanics/molecular mechanics boundaries". *Journal of Chemical Theory and Computation*, American Chemical Society, v. 10, n. 9, p. 3784–3790, 09 2014. Availabe at: <https://doi.org/10.1021/ct500512f>.

276 ANGELI, C. et al. "Introduction of n-electron valence states for multireference perturbation theory". *J. Chem. Phys.*, v. 114, n. 23, p. 10252–10264, 2001. Availabe at: <https://doi.org/10.1063/1.1361246>.

277 ANGELI, C. et al. "A quasidegenerate formulation of the second order n-electron valence state perturbation theory approach". *J. Chem. Phys.*, v. 121, n. 9, p. 4043–4049, 2004. Availabe at: <https://doi.org/10.1063/1.1778711>.

278 ANGELI, C.; CIMIRAGLIA, R.; MALRIEU, J.-P. "N-electron valence state perturbation theory: A fast implementation of the strongly contracted variant". *J. Chem. Phys.*, v. 114, p. 10663–10669, 2001.

279 Li Manni, G. et al. "Multiconfiguration Pair-Density Functional Theory". *J. Chem. Theory Comput.*, American Chemical Society, v. 10, n. 9, p. 3669–3680, sep 2014. ISSN 1549-9618. Availabe at: <https://doi.org/10.1021/ct500483t>.

280 GAGLIARDI, L. et al. "Multiconfiguration pair-density functional theory: A new way to treat strongly correlated systems". *Acc. Chem. Res.*, v. 50, n. 1, p. 66–73, 2017. Availabe at: <https://doi.org/10.1021/acs.accounts.6b00471>.

281 ZHOU, C. et al. "Electronic structure of strongly correlated systems: recent developments in multiconfiguration pair-density functional theory and multiconfiguration nonclassical-energy functional theory". *Chemical Science*, The Royal Society of Chemistry, v. 13, n. 26, p. 7685–7706, 2022. ISSN 2041-6520. Availabe at: <http://dx.doi.org/10.1039/D2SC01022D>.

282 HERMES, M. R.; PANDHARKAR, R.; GAGLIARDI, L. "Variational Localized Active Space Self-Consistent Field Method". *J. Chem. Theory Comput.*, American Chemical Society, v. 16, n. 8, p. 4923–4937, aug 2020. ISSN 1549-9618. Availabe at: <https://doi.org/10.1021/acs.jctc.0c00222>.

283 PANDHARKAR, R. et al. "Localized Active Space-State Interaction: a Multireference Method for Chemical Insight". *Journal of Chemical Theory and Computation*, American Chemical Society, v. 18, n. 11, p. 6557–6566, nov 2022. ISSN 1549-9618. Availabe at: <https://doi.org/10.1021/acs.jctc.2c00536>.

284 OTTEN, M. et al. "Localized Quantum Chemistry on Quantum Computers". *Journal of Chemical Theory and Computation*, American Chemical Society, v. 18, n. 12, p. 7205–7217, dec 2022. ISSN 1549-9618. Availabe at: <https://doi.org/10.1021/acs.jctc.2c00388>.

285 HOLMES, A. A.; TUBMAN, N. M.; UMRIGAR, C. J. "Heat-bath configuration interaction: An efficient selected configuration interaction algorithm inspired by heat-bath sampling". *J. Chem. Theory Comput.*, American Chemical Society (ACS), v. 12, n. 8, p. 3674–3680, ago. 2016.

286 STEIN, C. J.; REIHER, M. "autoCAS: A Program for Fully Automated Multiconfigurational Calculations". *Journal of Computational Chemistry*, John Wiley & Sons, Ltd, v. 40, n. 25, p. 2216–2226, sep 2019. ISSN 0192-8651. Availabe at: <https://doi.org/10.1002/jcc.25869>.

287 BOOTH, G. H.; THOM, A. J. W.; ALAVI, A. "Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space". *The Journal of Chemical Physics*, v. 131, n. 5, p. 54106, aug 2009. ISSN 0021-9606. Availabe at: <https://doi.org/10.1063/1.3193710>.

288 LI, J.; JONES, B. A.; KAIS, S. "Toward perturbation theory methods on a quantum computer". *Science Advances*, American Association for the Advancement of Science, v. 9, n. 19, p. eadg4576, nov 2023. Availabe at: <https://doi.org/10.1126/sciadv.adg4576>.

289 ANAND, A. et al. "A quantum computing view on unitary coupled cluster theory". *Chemical Society Reviews*, Royal Society of Chemistry, v. 51, n. 5, p. 1659–1684, 2022.

290 D'CUNHA, R. et al. "Challenges in the use of quantum computing hardware-efficient ansätze in electronic structure theory". *The Journal of Physical Chemistry A*, ACS Publications, v. 127, n. 15, p. 3437–3448, 2023.

291 TANG, H. L. et al. "Qubit-adapt-vqe: An adaptive algorithm for constructing hardware-efficient ansätze on a quantum processor". *PRX Quantum*, American Physical Society (APS), v. 2, n. 2, abr. 2021. ISSN 2691-3399. Availabe at: <http://dx.doi.org/10.1103/PRXQuantum.2.020310>.

292 RYABINKIN, I. G. et al. "Qubit coupled-cluster method: A systematic approach to quantum chemistry on a quantum computer". *J. Chem Theory Comput.*, v. 14, p. 6317–6326, 2018.

293 RYABINKIN, I. et al. "Iterative qubit coupled cluster approach with efficient screening of generators". *J. Chem. Theory Comput.*, v. 16, n. 2, p. 1055–1063, 2020.

294 HUANG, B. et al. "Quantum simulations of fermionic Hamiltonians with efficient encoding and ansatz schemes". *Journal of Chemical Theory and Computation*, v. 19, p. 1487–1498, 2023.

295 NYKäNEN, A. et al. "*Mitigating the measurement overhead of ADAPT-VQE with optimised informationally complete generalised measurements*". 2023.

296 RAVI, G. S. et al. "Cafqa: A classical simulation bootstrap for variational quantum algorithms". In: *Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 1*. New York, NY, USA: Association for Computing Machinery, 2022. (ASPLOS 2023), p. 15–29. ISBN 9781450399159. Availabe at: <https://doi.org/10.1145/3567955.3567958>.

297 Khan, A.; Clark, B. K.; Tubman, N. M. "Pre-optimizing variational quantum eigensolvers with tensor networks". *arXiv e-prints*, p. arXiv:2310.12965, out. 2023.

298 Mullinax, J. W.; Tubman, N. M. "Large-scale sparse wavefunction circuit simulator for applications with the variational quantum eigensolver". *arXiv e-prints*, p. arXiv:2301.05726, jan. 2023.

299 KIVLICHAN, I. D. et al. "Quantum simulation of electronic structure with linear depth and connectivity". *Physical review letters*, APS, v. 120, n. 11, p. 110501, 2018.

300 TOMESH, T. et al. "Coreset clustering on small quantum computers". *Electronics*, MDPI, v. 10, n. 14, p. 1690, 2021.

301 HASHIM, A. et al. "Optimized swap networks with equivalent circuit averaging for qaoa". *Physical Review Research*, APS, v. 4, n. 3, p. 033028, 2022.

302 GRZESIAK, N. et al. "Efficient arbitrary simultaneously entangling gates on a trapped-ion quantum computer". *Nature communications*, Nature Publishing Group UK London, v. 11, n. 1, p. 2963, 2020.

303 BRAVYI, S.; MASLOV, D.; NAM, Y. "Constant-cost implementations of clifford operations and multiply-controlled gates using global interactions". *Physical Review Letters*, APS, v. 129, n. 23, p. 230501, 2022.

304 ANG, J. et al. "Architectures for multinode superconducting quantum computers". *arXiv preprint arXiv:2212.06167*, 2022.

305 WU, A.; DING, Y.; LI, A. "Collcomm: Enabling efficient collective quantum communication based on epr buffering". *arXiv preprint arXiv:2208.06724*, 2022.

306 ZHANG, Y. et al. "Variational quantum eigensolver with reduced circuit complexity". *npj Quantum Information*, Springer Science and Business Media LLC, v. 8, n. 1, ago. 2022. ISSN 2056-6387. Availabe at: <http://dx.doi.org/10.1038/s41534-022-00599-z>.

307 PENG, T. et al. "Simulating large quantum circuits on a small quantum computer". *Physical review letters*, APS, v. 125, n. 15, p. 150504, 2020.

308 PIVETEAU, C.; SUTTER, D. "Circuit knitting with classical communication". *IEEE Transactions on Information Theory*, IEEE, 2023.

309 TANG, W.; MARTONOSI, M. "Scaleqc: A scalable framework for hybrid computation on quantum and classical processors". *arXiv preprint arXiv:2207.00933*, 2022.

310 IBM. "*Introducing Quantum Serverless, a new programming model for leveraging quantum and classical resources*". 2021. <https://research.ibm.com/blog/quantum-serverless-programming>. Accessed: Nov 6, 2023.

311 COWTAN, A.; SIMMONS, W.; DUNCAN, R. "A generic compilation strategy for the unitary coupled cluster ansatz". *arXiv preprint arXiv:2007.10515*, 2020.

312 MARTIEL, S.; BRUGIÈRE, T. G. de. "Architecture aware compilation of quantum circuits via lazy synthesis". *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 6, p. 729, 2022.

313 AMY, M.; AZIMZADEH, P.; MOSCA, M. "On the controlled-not complexity of controlled-not–phase circuits". *Quantum Science and Technology*, IOP Publishing, v. 4, n. 1, p. 015002, 2018.

314 VANDAELE, V.; MARTIEL, S.; BRUGIÈRE, T. G. de. "Phase polynomials synthesis algorithms for nisq architectures and beyond". *Quantum Science and Technology*, IOP Publishing, v. 7, n. 4, p. 045027, 2022.

315 NAM, Y.; SU, Y.; MASLOV, D. "Approximate quantum fourier transform with o (n log (n)) t gates". *NPJ Quantum Information*, Nature Publishing Group UK London, v. 6, n. 1, p. 26, 2020.

316 BäUMER, E. et al. "Efficient long-range entanglement using dynamic circuits". *PRX Quantum*, American Physical Society (APS), v. 5, n. 3, ago. 2024. ISSN 2691-3399. Availabe at: <http://dx.doi.org/10.1103/PRXQuantum.5.030339>.

317 COPPERSMITH, D. "An approximate fourier transform useful in quantum factoring". *arXiv preprint quant-ph/0201067*, 2002.

318 CAMPBELL, C. et al. "Superstaq: Deep optimization of quantum programs". In: *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*. [S.l.: s.n.], 2023. v. 01, p. 1020–1032.

319 GIDNEY, C.; EKERÅ, M. "How to factor 2048 bit rsa integers in 8 hours using 20 million noisy qubits". *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 5, p. 433, 2021.

320 PABST, S.; NAM, Y. "A quantum algorithm for network reliability". *arXiv preprint arXiv:2203.10201*, 2022.

321 DALZELL, A. M. et al. "Quantum algorithms: A survey of applications and end-to-end complexities". *arXiv preprint arXiv:2310.03011*, 2023.

322 MAKSYMOV, A. et al. "Enhancing quantum computer performance via symmetrization". *arXiv preprint arXiv:2301.07233*, 2023.

323 STEIN, S. et al. "Q-beep: Quantum bayesian error mitigation employing poisson modeling over the hamming spectrum". In: *Proceedings of the 50th Annual International Symposium on Computer Architecture*. [S.l.: s.n.], 2023. p. 1–13.

324 BERG, E. V. D. et al. “Probabilistic error cancellation with sparse pauli–lindblad models on noisy quantum processors”. *Nature Physics*, Nature Publishing Group UK London, 2023.

325 KIM, Y. et al. “Scalable error mitigation for noisy quantum circuits produces competitive expectation values”. *Nature Physics*, Springer Science and Business Media LLC, v. 19, n. 5, p. 752–759, fev. 2023. ISSN 1745-2481. Availabe at: <http://dx.doi.org/10.1038/s41567-022-01914-3>.

326 LANDSMAN, K.; KEESAN, M.; MONROE, C. “Toward convergence of effective-field-theory simulations on digital quantum computers”. *Physical Review A*, APS, v. 100, n. 6, p. 062319, 2019.

327 GIURGICA-TIRON, T. et al. “Digital zero noise extrapolation for quantum error mitigation”. In: *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*. [S.l.: s.n.], 2020. p. 306–316.

328 MAJUMDAR, R. et al. “Best practices for quantum error mitigation with digital zero-noise extrapolation”. In: *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*. [S.l.: s.n.], 2023. v. 01, p. 881–887.

329 KIM, Y. et al. “Evidence for the utility of quantum computing before fault tolerance”. *Nature*, v. 618, n. 7965, p. 500–505, jun. 2023. ISSN 1476-4687. Availabe at: <https://doi.org/10.1038/s41586-023-06096-3>.

330 WAHL, M. A. et al. “Zero noise extrapolation on logical qubits by scaling the error correction code distance”. In: *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*. [S.l.: s.n.], 2023. v. 01, p. 888–897.

331 VIOLA, L.; KNILL, E.; LLOYD, S. “Dynamical decoupling of open quantum systems”. *Physical Review Letters*, v. 82, n. 12, p. 2417–2421, 1999.

332 UHRIG, G. S. “Keeping a quantum bit alive by optimized π-pulse sequences”. *Physical Review Letters*, v. 98, n. 10, p. 100504, 2007.

333 PAZ-SILVA, G. A.; VIOLA, L. “General transfer-function approach to noise filtering in open-loop quantum control”. *Physical Review Letters*, v. 113, n. 25, p. 250501, 2014.

334 LANGE, G. de et al. “Universal dynamical decoupling of a single solid-state spin from a spin bath”. *Science*, v. 330, n. 6000, p. 60–63, 2010.

335 BYLANDER, J. et al. “Noise spectroscopy through dynamical decoupling with a superconducting flux qubit”. *Nature Physics*, v. 7, p. 565–570, 2011.

336 SOARE, A. et al. “Experimental noise filtering by quantum control”. *Nature Physics*, v. 10, n. 11, p. 825–829, 2014.

337 RAVI, G. S. et al. “Quancorde: Boosting fidelity with quantum canary ordered diverse ensembles”. In: *2022 IEEE International Conference on Rebooting Computing (ICRC)*. [S.l.: s.n.], 2022. p. 66–77.

338 JONES, T. et al. “Quest and high performance simulation of quantum computers”. *Scientific Reports*, v. 9, p. 10736, 2019. Availabe at: <https://doi.org/10.1038/s41598-019-47174-9>.

339 LI, A. et al. “Sv-sim: Scalable pgas-based state vector simulation of quantum circuits”. In: *SC21: International Conference for High Performance Computing, Networking, Storage and Analysis*. [S.l.: s.n.], 2021. p. 1–14.

340 LYKOV, Y. A. D. “Importance of diagonal gates in tensor network simulations”. *arXiv*, 2021. Availabe at: <https://arxiv.org/abs/2106.15740>.

341 FEDOROV, D. A. et al. “Unitary Selective Coupled-Cluster Method”. *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 6, p. 703, maio 2022. ISSN 2521-327X. Availabe at: <https://doi.org/10.22331/q-2022-05-02-703>.

342 “LAYING the Groundwork for Quantum Powered Use Cases”. 2023. <https://www.youtube.com/watch?v=Hd-IGvuARfE&ab_channel=IBMResearch>. [Video].

343 GOTTESMAN, D. “The heisenberg representation of quantum computers”. *arXiv preprint quant-ph/9807006*, 1998.

344 AARONSON, S.; GOTTESMAN, D. “Improved simulation of stabilizer circuits”. *Physical Review A*, APS, v. 70, n. 5, p. 052328, 2004.

345 GARCIA, H. J.; MARKOV, I. L. “Simulation of quantum circuits via stabilizer frames”. *IEEE Transactions on Computers*, IEEE, v. 64, n. 8, p. 2323–2336, 2014.

346 AHARONOV, D.; LANDAU, Z.; MAKOWSKY, J. “The quantum fft can be classically simulated”. *arXiv preprint quant-ph/0611156*, 2006.

347 YORAN, N.; SHORT, A. J. “Efficient classical simulation of the approximate quantum fourier transform”. *Physical Review A*, APS, v. 76, n. 4, p. 042321, 2007.

348 HILLMICH, S.; MARKOV, I. L.; WILLE, R. “Just like the real thing: Fast weak simulation of quantum computation”. In: IEEE. *2020 57th ACM/IEEE Design Automation Conference (DAC)*. [S.l.], 2020. p. 1–6.

349 HILLMICH, S. et al. “Approximating decision diagrams for quantum circuit simulation”. *ACM Transactions on Quantum Computing*, ACM New York, NY, v. 3, n. 4, p. 1–21, 2022.

350 FATIMA, A.; MARKOV, I. L. “Faster schrödinger-style simulation of quantum circuits”. In: IEEE. *2021 IEEE International Symposium on High-Performance Computer Architecture (HPCA)*. [S.l.], 2021. p. 194–207.

351 BARENCO, A. et al. “Elementary gates for quantum computation”. *Physical review A*, APS, v. 52, n. 5, p. 3457, 1995.

352 HÄNER, T.; STEIGER, D. S. “0.5 petabyte simulation of a 45-qubit quantum circuit”. In: ACM. *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. [S.l.], 2017. p. 33.

353 PEDNAULT, E. et al. “Breaking the 49-qubit barrier in the simulation of quantum circuits”. *arXiv preprint arXiv:1710.05867*, 2017.

354 PEDNAULT, E. et al. “Leveraging secondary storage to simulate deep 54-qubit sycamore circuits”. *arXiv preprint arXiv:1910.09534*, 2019.

355 LI, A. et al. “Sv-sim: scalable pgas-based state vector simulation of quantum circuits”. In: *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. [S.l.: s.n.], 2021. p. 1–14.

356 LI, A. et al. “Deep quantum circuit simulations of low-energy nuclear states”. *arXiv preprint arXiv:2310.17739*, 2023.

357 Hirsbrunner, M. R. et al. “Beyond MP2 initialization for unitary coupled cluster quantum circuits”. *arXiv e-prints*, p. arXiv:2301.05666, jan. 2023.

358 MARKOV, I. L. et al. “Massively parallel approximate simulation of hard quantum circuits”. In: IEEE. *2020 57th ACM/IEEE Design Automation Conference (DAC)*. [S.l.], 2020. p. 1–6.

359 LI, A. et al. “Density matrix quantum circuit simulation via the bsp machine on modern gpu clusters”. In: IEEE. *Sc20: international conference for high performance computing, networking, storage and analysis*. [S.l.], 2020. p. 1–15.

360 MARKOV, I. L.; SHI, Y. “Simulating quantum computation by contracting tensor networks”. *SIAM Journal on Computing*, SIAM, v. 38, n. 3, p. 963–981, 2008.

361 LYKOV, D.; ALEXEEV, Y. “Importance of diagonal gates in tensor network simulations”. In: *2021 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*. [S.l.: s.n.], 2021. p. 447–452.

362 LYKOV, D. et al. “Tensor network quantum simulator with step-dependent parallelization”. In: *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*. [S.l.: s.n.], 2022. p. 582–593.

363 LYKOV, D. et al. “Performance evaluation and acceleration of the qtensor quantum circuit simulator on gpus”. In: IEEE. *2021 IEEE/ACM Second International Workshop on Quantum Computing Software (QCS)*. [S.l.], 2021. p. 27–34.

364 LYKOV, D. et al. “Tensor network quantum simulator with step-dependent parallelization”. In: IEEE. *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*. [S.l.], 2022. p. 582–593.

365 BERQUIST, W. et al. “Stochastic approach for simulating quantum noise using tensor networks”. In: IEEE. *2022 IEEE/ACM Third International Workshop on Quantum Computing Software (QCS)*. [S.l.], 2022. p. 107–113.

366 SHAH, M. et al. “Gpu-accelerated error-bounded compression framework for quantum circuit simulations”. In: IEEE. *2023 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*. [S.l.], 2023. p. 757–767.

367 BARTOLOMEO, G. D. et al. “Noisy gates for simulating quantum computers”. *Phys. Rev. Res.*, American Physical Society, v. 5, p. 043210, Dec 2023. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevResearch.5.043210>.

368 SAWAYA, N. P. et al. “HamLib: A library of Hamiltonians for benchmarking quantum algorithms and hardware”. In: *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*. [S.l.: s.n.], 2023. v. 02, p. 389–390.

369 REIHER, M. et al. “Elucidating reaction mechanisms on quantum computers”. *Proceedings of the national academy of sciences*, National Acad Sciences, v. 114, n. 29, p. 7555–7560, 2017. Availabe at: <https://www.pnas.org/content/114/29/7555.short>.

370 Tubman, N. M. et al. “Postponing the orthogonality catastrophe: efficient state preparation for electronic structure simulations on quantum devices”. *arXiv e-prints*, p. arXiv:1809.05523, set. 2018.

371 ELFVING, V. E. et al. “How will quantum computers provide an industrially relevant computational advantage in quantum chemistry?” *arXiv preprint arXiv:2009.12472*, 2020.

372 GOINGS, J. J. et al. “Reliably assessing the electronic structure of cytochrome p450 on today’s classical computers and tomorrow’s quantum computers”. *Proceedings of the National Academy of Sciences*, v. 119, n. 38, p. e2203533119, 2022. Availabe at: <https://www.pnas.org/doi/abs/10.1073/pnas.2203533119>.

373 MEJUTO-ZAERA, C. et al. “Are multi-quasiparticle interactions important in molecular ionization?” *The Journal of Chemical Physics*, v. 154, n. 12, p. 121101, 03 2021. ISSN 0021-9606. Availabe at: <https://doi.org/10.1063/5.0044060>.

374 SHISHKIN, M.; KRESSE, G. “Self-consistent *gw* calculations for semiconductors and insulators”. *Phys. Rev. B*, American Physical Society, v. 75, p. 235102, Jun 2007. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevB.75.235102>.

375 PERDEW, J. P. “Density functional theory and the band gap problem”. *International Journal of Quantum Chemistry*, Wiley Online Library, v. 28, n. S19, p. 497–523, 1985.

376 COHEN, A. J.; MORI-SÁNCHEZ, P.; YANG, W. “Challenges for density functional theory”. *Chemical reviews*, ACS Publications, v. 112, n. 1, p. 289–320, 2012.

377 SPARTA, M.; NEESE, F. “Chemical Applications Carried Out by Local Pair Natural Orbital Based Coupled-Cluster Methods”. *Chem. Soc. Rev.*, Royal Society of Chemistry, v. 43, n. 14, p. 5032–5041, 2014. Availabe at: <https://pubs.rsc.org/en/content/articlelanding/2014/cs/c4cs00050a#!divAbstract>.

378 HAIT, D. et al. “What levels of coupled cluster theory are appropriate for transition metal systems? a study using near-exact quantum chemical values for 3d transition metal binary compounds”. *Journal of Chemical Theory and Computation*, v. 15, n. 10, p. 5370–5385, 2019. PMID: 31465217. Availabe at: <https://doi.org/10.1021/acs.jctc.9b00674>.

379 MOLTVED, K. A.; KEPP, K. P. “Performance of density functional theory for transition metal oxygen bonds”. *ChemPhysChem*, Wiley Online Library, v. 20, n. 23, p. 3210–3220, 2019.

380 SHARMA, S. et al. “Low-Energy Spectrum of Iron–Sulfur Clusters Directly from Many-Particle Quantum Mechanics”. *Nature Chemistry*, Nature Publishing Group, v. 6, n. 10, p. 927–933, 2014. Availabe at: <https://www.nature.com/articles/nchem.2041>.

381 LI, Z. “*Active-space-model-for-Iron-Sulfur-Clusters*”. [S.l.]: GitHub, 2020. <https://github.com/zhendongli2008/Active-space-model-for-Iron-Sulfur-Clusters>.

382 LI, Z.; CHAN, G. K.-L. "Spin-projected matrix product states: Versatile tool for strongly correlated systems". *Journal of Chemical Theory and Computation*, ACS Publications, v. 13, n. 6, p. 2681–2695, 2017.

383 NOODLEMAN, L. et al. "Insights into Properties and Energetics of Iron–Sulfur Proteins from Simple Clusters to Nitrogenase". *Curr. Opin. Chem. Biol.*, Elsevier, v. 6, n. 2, p. 259–273, 2002. Availabe at: <https://www.sciencedirect.com/science/article/pii/S1367593102003095>.

384 LI, Z. [S.l.]: GitHub, 2020. <https://github.com/zhendongli2008/Active-space-model-for-FeMoco>.

385 LI, Z. et al. "The Electronic Complexity of the Ground-State of the FeMo Cofactor of Nitrogenase as Relevant to Quantum Simulations". *JCP*, AIP Publishing, v. 150, n. 2, p. 024302–024306, 2019. Availabe at: <https://aip.scitation.org/doi/10.1063/1.5063376>.

386 LI, Z. et al. "Electronic Landscape of the P-Cluster of Nitrogenase as Revealed Through Many-Electron Quantum Wavefunction Simulations". *Nature Chemistry*, Nature Publishing Group, v. 11, p. 1026–1033, 2019. Availabe at: <https://www.nature.com/articles/s41557-019-0337-3?draft=marketing>.

387 CADY, C. W.; CRABTREE, R. H.; BRUDVIG, G. W. "Functional Models for the Oxygen-Evolving Complex of Photosystem II". *Coord. Chem. Rev.*, Elsevier, v. 252, n. 3-4, p. 444–455, 2008. Availabe at: <https://sofi-northwestern.pure.elsevier.com/en/publications/functional-models-for-the-oxygen-evolving-complex-of-photosystem->.

388 KURASHIGE, Y.; CHAN, G. K.-L.; YANAI, T. "Entangled Quantum Electronic Wavefunctions of the Mn_4CaO_5 Cluster in Photosystem II". *Nature Chemistry*, Nature Publishing Group, v. 5, n. 8, p. 660–666, 2013. Availabe at: <https://www.nature.com/articles/nchem.1677>.

389 CHILKURI, V. G.; DEBEER, S.; NEESE, F. "Ligand Field Theory and Angular Overlap Model Based Analysis of the Electronic Structure of Homovalent Iron–Sulfur Dimers". *Inorg. Chem.*, ACS Publications, v. 59, p. 984–995, 2020. Availabe at: <https://pubs.acs.org/doi/10.1021/acs.inorgchem.9b00974>.

390 CAO, L.; CALDARARU, O.; RYDE, U. "Protonation and Reduction of the FeMo Cluster in Nitrogenase Studied by Quantum Mechanics/Molecular Mechanics (QM/MM) Calculations". *Journal of Chemical Theory and Computation*, ACS Publications, v. 14, n. 12, p. 6653–6678, 2018. Availabe at: <https://pubs.acs.org/doi/abs/10.1021/acs.jctc.8b00778>.

391 SIEGBAHN, P. E. "Structures and Energetics for O_2 Formation in Photosystem II". *Acc. Chem. Res.*, ACS Publications, v. 42, n. 12, p. 1871–1880, 2009. Availabe at: <https://pubs.acs.org/doi/abs/10.1021/ar900117k>.

392 LANCASTER, K. M. et al. "X-ray Emission Spectroscopy Evidences a Central Carbon in the Nitrogenase Iron-Molybdenum Cofactor". *Science*, American Association for the Advancement of Science, v. 334, n. 6058, p. 974–977, 2011. Availabe at: <https://science.sciencemag.org/content/334/6058/974>.

393 BATOOL, J.; HAHN, T.; PEDERSON, M. R. "Magnetic Signatures of Hydroxyl-and Water-Terminated Neutral and Tetra-Anionic Mn12-Acetate". *J. Comput. Chem.*, Wiley Online Library, v. 40, n. 26, p. 2301–2308, 2019. Availabe at: <https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.26008>.

394 SCHIMKA, L. et al. "Accurate Surface and Adsorption Energies from Many-Body Perturbation Theory". *Nat. Mater.*, Nature Publishing Group, v. 9, n. 9, p. 741–744, 2010. Availabe at: <https://www.nature.com/articles/nmat2806>.

395 CAPDEVILA-CORTADA, M.; ŁODZIANA, Z.; LÓPEZ, N. "Performance of DFT+U Approaches in the Study of Catalytic Materials". *ACS Catalysis*, American Chemical Society, v. 6, p. 8370–8379, 2016. Availabe at: <https://pubs.acs.org/doi/abs/10.1021/acscatal.6b01907>.

396 NØRSKOV, J. K. et al. "Density Functional Theory in Surface Chemistry and Catalysis". *PNAS*, National Acad Sciences, v. 108, n. 3, p. 937–943, 2011. Availabe at: <https://www.pnas.org/content/108/3/937.short>.

397 WODTKE, A. M. "Electronically Non-Adiabatic Influences in Surface Chemistry and Dynamics". *Chem. Soc. Rev.*, Royal Society of Chemistry, v. 45, n. 13, p. 3641–3657, 2016. Availabe at: <https://pubs.rsc.org/en/content/articlelanding/2016/cs/c6cs00078a#!divAbstract>.

398 NØRSKOV, J. K. et al. "Towards the Computational Design of Solid Catalysts". *Nature Chemistry*, Nature Publishing Group, v. 1, n. 1, p. 37–46, 2009. Availabe at: <https://www.nature.com/articles/nchem.121>.

399 VANDEVONDELE, J.; BORSTNIK, U.; HUTTER, J. "Linear Scaling Self-Consistent Field Calculations with Millions of Atoms in the Condensed Phase". *Journal of Chemical Theory and Computation*, ACS Publications, v. 8, n. 10, p. 3565–3573, 2012. Availabe at: <https://pubs.acs.org/doi/abs/10.1021/ct200897x>.

400 STETINA, T. F. et al. "Simulating effective qed on quantum computers". *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 6, p. 622, 2022.

401 KEIMER, B. et al. "From Quantum Matter to High-Temperature Superconductivity in Copper Oxides". *Nature*, Nature Publishing Group, v. 518, n. 7538, p. 179–186, 2015.

402 LEE, S.-S. "Recent Developments in Non-Fermi Liquid Theory". *Annu. Rev. Condens. Matter Phys.*, Annual Reviews, v. 9, p. 227–244, 2018. Availabe at: <https://www.annualreviews.org/doi/abs/10.1146/annurev-conmatphys-031016-025531>.

403 FRADKIN, E.; KIVELSON, S. A.; TRANQUADA, J. M. "Colloquium: Theory of Intertwined Orders in High-Temperature Superconductors". *RMP*, American Physical Society, v. 87, p. 457–482, May 2015. Availabe at: <https://link.aps.org/doi/10.1103/RevModPhys.87.457>.

404 BALENTS, L. "Spin Liquids in Frustrated Magnets". *Nature*, Nature Publishing Group, v. 464, n. 7286, p. 199–208, 2010. Availabe at: <https://www.nature.com/articles/nature08917>.

405 ROSSMANNEK, M. et al. “Quantum hf/dft-embedding algorithms for electronic structure calculations: Scaling up to complex molecular systems”. *The Journal of Chemical Physics*, AIP Publishing, v. 154, n. 11, mar. 2021. ISSN 1089-7690. Availabe at: <http://dx.doi.org/10.1063/5.0029536>.

406 ROSSMANNEK, M. et al. “Quantum embedding method for the simulation of strongly correlated systems on quantum computers”. *The Journal of Physical Chemistry Letters*, American Chemical Society (ACS), v. 14, n. 14, p. 3491–3497, abr. 2023. Availabe at: <https://doi.org/10.1021%2Facs.jpclett.3c00330>.

407 GAO, Y. et al. “On-surface synthesis of a doubly anti-aromatic carbon allotrope: Cyclo[16]carbon”. American Chemical Society (ACS), mar. 2023. Availabe at: <https://doi.org/10.26434%2Fchemrxiv-2023-w124b>.

408 RONČEVIĆ, I. et al. “Aromaticity reversal induced by vibrations in cyclo[16]carbon”. American Chemical Society (ACS), ago. 2023. Availabe at: <https://doi.org/10.26434/chemrxiv-2023-8kc5t>.

409 ANISIMOV, V. I. et al. “First-principles calculations of the electronic structure and spectra of strongly correlated systems: Dynamical mean-field theory”. *Journal of Physics: Condensed Matter*, v. 9, p. 7359, 1997.

410 LICHTENSTEIN, A. I.; KATSNELSON, M. I. “Ab initio calculations of quasiparticle band structure in correlated systems: LDA++ approach”. *Physical Review B*, v. 57, p. 6884–6895, 1998.

411 SUN, P.; KOTLIAR, G. “Extended dynamical mean-field theory and GW method”. *Physical Review B*, v. 66, p. 085120, 2002.

412 BIERMANN, S.; ARYASETIAWAN, F.; GEORGES, A. “First-principles approach to the electronic structure of strongly correlated systems: Combining the GW approximation and dynamical mean-field theory”. *Physical Review Letters*, v. 90, p. 086402, 2003.

413 BIERMANN, S. “Dynamical screening effects in correlated electron materials—a progress report on combined many-body perturbation and dynamical mean field theory: GW+DMFT”. *Journal of Physics: Condensed Matter*, v. 26, p. 173202, 2014.

414 BOEHNKE, L. et al. “When strong correlations become weak: Consistent merging of gw and dmft”. *Physical Review B*, v. 94, p. 201106, 2016.

415 CHOI, S. et al. “First-principles treatment of Mott insulators: Linearized QSGW+DMFT approach”. *npj Quantum Materials*, v. 1, p. 1–6, 2016.

416 NILSSON, F. et al. “Multitier self-consistent GW+EDMFT”. *Physical Review Materials*, v. 1, p. 043803, 2017.

417 LAN, T. N.; ZGID, D. “Generalized self-energy embedding theory”. *The Journal of Physical Chemistry Letters*, v. 8, p. 2200–2205, 2017.

418 ZGID, D.; GULL, E. “Finite temperature quantum embedding theories for correlated systems”. *New Journal of Physics*, v. 19, p. 023047, 2017.

419 RUSAKOV, A. A. et al. “Self-energy embedding theory (SEET) for periodic systems”. *Journal of Chemical Theory and Computation*, v. 15, p. 229–240, 2019.

420 MA, H. et al. "First-principles studies of strongly correlated states in defect spin qubits in diamond". *Physical Chemistry Chemical Physics*, Royal Society of Chemistry, v. 22, n. 44, p. 25522–25527, 2020.

421 MA, H. et al. "Quantum embedding theory for strongly correlated states in materials". *Journal of Chemical Theory and Computation*, ACS Publications, v. 17, n. 4, p. 2116–2125, 2021.

422 RUNGGER, I. et al. "*Dynamical mean field theory algorithm and experiment on quantum computers*". 2020.

423 KEEN, T. et al. "Quantum-classical simulation of two-site dynamical mean-field theory on noisy quantum hardware". *Quantum Science and Technology*, v. 5, p. 035001, 2020.

424 KAWASHIMA, Y. et al. "Optimizing electronic structure simulations on a trapped-ion quantum computer using problem decomposition". *Communications Physics*, v. 4, n. 1, p. 245, 2021. Availabe at: <https://doi.org/10.1038/s42005-021-00751-9>.

425 YAO, Y. et al. "Gutzwiller hybrid quantum-classical computing approach for correlated materials". *Physical Review Research*, v. 3, p. 013184, 2021.

426 TILLY, J. et al. "Reduced density matrix sampling: Self-consistent embedding and multiscale electronic structure on current generation quantum computers". *Physical Review Research*, v. 3, p. 033230, 2021.

427 VORWERK, C. et al. "Quantum embedding theories to simulate condensed systems on quantum computers". *Nature Computational Science*, v. 2, p. 424–432, 2022.

428 HUANG, B.; GOVONI, M.; GALLI, G. "Simulating the electronic structure of spin defects on quantum computers". *PRX Quantum*, v. 3, p. 010339, 2022.

429 MITRA, A. et al. "Periodic Density Matrix Embedding for CO Adsorption on the MgO(001) Surface". *J. Phys. Chem. Lett.*, American Chemical Society, v. 13, n. 32, p. 7483–7489, aug 2022. Availabe at: <https://doi.org/10.1021/acs.jpclett.2c01915>.

430 PAVARINI, E. et al. "Band-structure trend in hole-doped cuprates and correlation with t c max". *Physical review letters*, APS, v. 87, n. 4, p. 047003, 2001.

431 ARYASETIAWAN, F. et al. "Frequency-dependent local interactions and low-energy effective models from electronic structure calculations". *Physical Review B*, APS, v. 70, n. 19, p. 195104, 2004.

432 JESCHKE, H. O.; SALVAT-PUJOL, F.; VALENTÍ, R. "First-principles determination of heisenberg hamiltonian parameters for the spin-1 2 kagome antiferromagnet zncu 3 (oh) 6 cl 2". *Physical Review B*, APS, v. 88, n. 7, p. 075106, 2013.

433 HAULE, K. "Exact double counting in combining the dynamical mean field theory and the density functional theory". *Physical review letters*, APS, v. 115, n. 19, p. 196403, 2015.

434 TEN-NO, S. "Stochastic determination of effective hamiltonian for the full configuration interaction solution of quasi-degenerate electronic states". *The Journal of chemical physics*, AIP Publishing, v. 138, n. 16, 2013.

435 ZHOU, S.; CEPERLEY, D. "Construction of localized wave functions for a disordered optical lattice and analysis of the resulting hubbard model parameters". *Physical Review A*, APS, v. 81, n. 1, p. 013402, 2010.

436 WHITE, S. R. "Numerical canonical transformation approach to quantum many-body problems". *The Journal of chemical physics*, American Institute of Physics, v. 117, n. 16, p. 7472–7482, 2002.

437 YANAI, T.; CHAN, G. K. "Canonical transformation theory for multireference problems". *The Journal of chemical physics*, AIP Publishing, v. 124, n. 19, 2006.

438 SAVARY, L.; BALENTS, L. "Quantum spin liquids: a review". *Reports on Progress in Physics*, IOP Publishing, v. 80, n. 1, p. 016502, 2016.

439 ZHOU, Y.; KANODA, K.; NG, T.-K. "Quantum spin liquid states". *Reviews of Modern Physics*, APS, v. 89, n. 2, p. 025003, 2017.

440 CHAMORRO, J. R.; MCQUEEN, T. M.; TRAN, T. T. "Chemistry of quantum spin liquids". *Chemical Reviews*, ACS Publications, v. 121, n. 5, p. 2898–2934, 2020.

441 SEARS, J. A. et al. "Phase diagram of α- rucl 3 in an in-plane magnetic field". *Physical Review B*, APS, v. 95, n. 18, p. 180411, 2017.

442 JAHIN, A. et al. "Fermionic approach to variational quantum simulation of kitaev spin models". *Phys. Rev. A*, American Physical Society, v. 106, p. 022434, Aug 2022. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevA.106.022434>.

443 LI, A. C. et al. "Benchmarking variational quantum eigensolvers for the square-octagon-lattice kitaev model". *Physical Review Research*, APS, v. 5, n. 3, p. 033071, 2023.

444 PLUMB, K. et al. "α- rucl 3: A spin-orbit assisted mott insulator on a honeycomb lattice". *Physical Review B*, APS, v. 90, n. 4, p. 041112, 2014.

445 KITAEV, A. "Anyons in an exactly solved model and beyond". *Annals of Physics*, Elsevier, v. 321, n. 1, p. 2–111, 2006.

446 JACKELI, G.; KHALIULLIN, G. "Mott insulators in the strong spin-orbit coupling limit: from heisenberg to a quantum compass and kitaev models". *Physical review letters*, APS, v. 102, n. 1, p. 017205, 2009.

447 BANERJEE, A. et al. "Proximate kitaev quantum spin liquid behaviour in a honeycomb magnet". *Nature materials*, Nature Publishing Group UK London, v. 15, n. 7, p. 733–740, 2016.

448 SEARS, J. A. et al. "Magnetic order in α- rucl 3: A honeycomb-lattice quantum magnet with strong spin-orbit coupling". *Physical Review B*, APS, v. 91, n. 14, p. 144420, 2015.

449 SANDILANDS, L. J. et al. "Scattering continuum and possible fractionalized excitations in α- rucl 3". *Physical review letters*, APS, v. 114, n. 14, p. 147201, 2015.

450 MAKSIMOV, P.; CHERNYSHEV, A. "Rethinking α- rucl 3". *Physical Review Research*, APS, v. 2, n. 3, p. 033011, 2020.

451 YADAV, R. et al. "Kitaev exchange and field-induced quantum spin-liquid states in honeycomb α-rucl3". *Scientific reports*, Nature Publishing Group UK London, v. 6, n. 1, p. 37925, 2016.

452 WINTER, S. M. et al. "Challenges in design of kitaev materials: Magnetic interactions from competing energy scales". *Physical Review B*, APS, v. 93, n. 21, p. 214431, 2016.

453 WINTER, S. M. et al. "Models and materials for generalized kitaev magnetism". *Journal of Physics: Condensed Matter*, IOP Publishing, v. 29, n. 49, p. 493002, 2017.

454 GOHLKE, M. et al. "Dynamics of the kitaev-heisenberg model". *Physical review letters*, APS, v. 119, n. 15, p. 157203, 2017.

455 LAURELL, P.; OKAMOTO, S. "Dynamical and thermal magnetic properties of the kitaev spin liquid candidate α-rucl3". *npj quantum materials*, Nature Publishing Group UK London, v. 5, n. 1, p. 2, 2020.

456 IMADA, M.; FUJIMORI, A.; TOKURA, Y. "Metal-insulator transitions". *Reviews of modern physics*, APS, v. 70, n. 4, p. 1039, 1998.

457 SPALDIN, N. A. "*Magnetic materials: fundamentals and applications*". [S.l.]: Cambridge university press, 2010.

458 ORENSTEIN, J.; MILLIS, A. "Advances in the physics of high-temperature superconductivity". *Science*, American Association for the Advancement of Science, v. 288, n. 5465, p. 468–474, 2000.

459 DAGOTTO, E. "Correlated electrons in high-temperature superconductors". *Reviews of Modern Physics*, APS, v. 66, n. 3, p. 763, 1994.

460 HUBBARD, J. "Electron correlations in narrow energy bands. ii. the degenerate band case". *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, The Royal Society London, v. 277, n. 1369, p. 237–259, 1964.

461 KANAMORI, J. "Electron correlation and ferromagnetism of transition metals". *Progress of Theoretical Physics*, Oxford University Press, v. 30, n. 3, p. 275–289, 1963.

462 GUTZWILLER, M. C. "Effect of correlation on the ferromagnetism of transition metals". *Physical Review Letters*, APS, v. 10, n. 5, p. 159, 1963.

463 AROVAS, D. P. et al. "The hubbard model". *Annual review of condensed matter physics*, Annual Reviews, v. 13, p. 239–274, 2022.

464 LIEB, E. H.; WU, F.-Y. "Absence of mott transition in an exact solution of the short-range, one-band model in one dimension". *Physical Review Letters*, APS, v. 20, n. 25, p. 1445, 1968.

465 METZNER, W.; VOLLHARDT, D. "Correlated lattice fermions in $d = \infty$ dimensions". *Physical review letters*, APS, v. 62, n. 3, p. 324, 1989.

466 OITMAA, J.; HAMER, C.; ZHENG, W. "*Series expansion methods for strongly interacting lattice models*". [S.l.]: Cambridge University Press, 2006.

467 LEBLANC, J. P. et al. “Solutions of the two-dimensional hubbard model: Benchmarks and results from a wide range of numerical algorithms”. *Physical Review X*, APS, v. 5, n. 4, p. 041041, 2015.

468 ZHENG, B.-X. et al. “Stripe order in the underdoped region of the two-dimensional hubbard model”. *Science*, American Association for the Advancement of Science, v. 358, n. 6367, p. 1155–1160, 2017.

469 EMERY, V. “Theory of high-t c superconductivity in oxides”. *Physical Review Letters*, APS, v. 58, n. 26, p. 2794, 1987.

470 GREITER, M.; THOMALE, R. “No evidence for spontaneous orbital currents in numerical studies of three-band models for the cuo planes of high-temperature superconductors”. *Physical review letters*, APS, v. 99, n. 2, p. 027005, 2007.

471 HUANG, E. W. et al. “Numerical evidence of fluctuating stripes in the normal state of high-t_c cuprate superconductors”. *Science*, American Association for the Advancement of Science, v. 358, n. 6367, p. 1161–1164, 2017.

472 VITALI, E. et al. “Metal-insulator transition in the ground state of the three-band hubbard model at half filling”. *Physical Review B*, APS, v. 99, n. 16, p. 165116, 2019.

473 CUI, Z.-H. et al. “Ground-state phase diagram of the three-band hubbard model from density matrix embedding theory”. *Physical Review Research*, APS, v. 2, n. 4, p. 043259, 2020.

474 FAUQUÉ, B. et al. “Magnetic order in the pseudogap phase of high-t_c superconductors”. *Physical Review Letters*, APS, v. 96, n. 19, p. 197001, 2006.

475 HYBERTSEN, M. S.; SCHLÜTER, M.; CHRISTENSEN, N. E. “Calculation of coulomb-interaction parameters for la_2cuo_4 using a constrained-density-functional approach”. *Physical Review B*, APS, v. 39, n. 13, p. 9028, 1989.

476 MARTIN, R. L. “Electronic localization in the cuprates”. *Physical Review B*, APS, v. 53, n. 23, p. 15501, 1996.

477 HANKE, W. et al. “The 3-band hubbard-model versus the 1-band model for the high-t c cuprates: Pairing dynamics, superconductivity and the ground-state phase diagram”. *The European Physical Journal Special Topics*, Springer, v. 188, p. 15–32, 2010.

478 DAMASCELLI, A. “Probing the Electronic Structure of Complex Systems by ARPES”. *Physica Scripta*, IOP Publishing, T109, p. 61–74, 2004. Availabe at: <https://doi.org/10.1238%2Fphysica.topical.109a00061>.

479 ZUEHLSDORFF, T. J. et al. “Optical spectra in the condensed phase: Capturing anharmonic and vibronic features using dynamic and static approaches”. *The Journal of Chemical Physics*, v. 151, n. 7, p. 074111, 08 2019. ISSN 0021-9606. Availabe at: <https://doi.org/10.1063/1.5114818>.

480 WIETHORN, Z. R. et al. “Beyond the Condon limit: Condensed phase optical spectra from atomistic simulations”. *The Journal of Chemical Physics*, v. 159, n. 24, p. 244114, 12 2023. ISSN 0021-9606. Availabe at: <https://doi.org/10.1063/5.0180405>.

481 PATTERSON, B. D. et al. “Coherent Science at the SwissFEL X-Ray Laser”. *NJP*, IOP Publishing, v. 12, n. 3, p. 035012–035028, 2010. Availabe at: <https://iopscience.iop.org/article/10.1088/1367-2630/12/3/035012/meta>.

482 WEATHERSBY, S. P. et al. “Mega-Electron-Volt Ultrafast Electron Diffraction at SLAC National Accelerator Laboratory”. *Rev. Sci. Instrum*, AIP Publishing, v. 86, n. 7, p. 073702–073709, 2015. Availabe at: <https://aip.scitation.org/doi/10.1063/1.4926994>.

483 FISCHER, M. C. et al. “Invited Review Article: Pump-Probe Microscopy”. *Rev. Sci. Instrum*, AIP Publishing, v. 87, n. 3, p. 031101–031122, 2016. Availabe at: <https://aip.scitation.org/doi/10.1063/1.4943211>.

484 BUZZI, M. et al. “Probing Dynamics in Quantum Materials with Femtosecond X-rays”. *Nat. Rev. Mater*, Nature Publishing Group, v. 3, n. 9, p. 299–311, 2018. Availabe at: <http://dx.doi.org/10.1038/s41578-018-0024-9>.

485 OLLITRAULT, P. J. et al. “Hardware efficient quantum algorithms for vibrational structure calculations”. *Chem. Sci.*, v. 11, p. 6842–6855, 2020.

486 TRENEV, D. et al. “*Refining resource estimation for the quantum computation of vibrational molecular spectra through Trotter error analysis*”. 2023.

487 SAWAYA, N. P. D.; PAESANI, F.; TABOR, D. P. “Near- and long-term quantum algorithmic approaches for vibrational spectroscopy”. *Phys. Rev. A*, v. 104, p. 062419, 2021.

488 STOBER, S. T. et al. “Considerations for evaluating thermodynamic properties with hybrid quantum-classical computing work flows”. *Phys. Rev. A*, American Physical Society, v. 105, p. 012425, Jan 2022. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevA.105.012425>.

489 DEMMIG-ADAMS, B.; ADAMS, W. “Photosynthesis and partitioning | photoinhibition”. In: THOMAS, B. (Ed.). *Encyclopedia of Applied Plant Sciences*. Oxford: Elsevier, 2003. p. 707–714. ISBN 978-0-12-227050-5.

490 HIGGINS, J. S. et al. “Photosynthesis tunes quantum-mechanical mixing of electronic and vibrational states to steer exciton energy transfer”. *Proceedings of the National Academy of Sciences*, v. 118, n. 11, p. e2018240118, 2021.

491 RITZ, T.; DAMJANOVIć, A.; SCHULTEN, K. “The quantum physics of photosynthesis”. *ChemPhysChem*, v. 3, n. 3, p. 243–248, 2002.

492 COX, N.; PANTAZIS, D. A.; LUBITZ, W. “Current understanding of the mechanism of water oxidation in photosystem ii and its relation to xfel data”. *Annual Review of Biochemistry*, v. 89, n. 1, p. 795–820, 2020. PMID: 32208765. Availabe at: <https://doi.org/10.1146/annurev-biochem-011520-104801>.

493 JADERBERG, B. et al. “Recompilation-enhanced simulation of electron–phonon dynamics on ibm quantum computers”. *New Journal of Physics*, IOP Publishing, v. 24, n. 9, p. 093017, sep 2022. Availabe at: <https://dx.doi.org/10.1088/1367-2630/ac8a69>.

494 SCIENCE, N.; (US), T. C. “*Materials genome initiative for global competitiveness*”. [S.l.]: Executive Office of the President, National Science and Technology Council, 2011.

495 LONG, D.-B. et al. “Two-dimensional bimetal-embedded expanded phthalocyanine monolayers: A class of multifunctional materials with fascinating properties”. *Advanced Functional Materials*, n/a, n. n/a, p. 2313171. Availabe at: <https://onlinelibrary.wiley.com/doi/abs/10.1002/adfm.202313171>.

496 SMITH, D. R.; PENDRY, J. B.; WILTSHIRE, M. C. K. “Metamaterials and negative refractive index”. *Science*, v. 305, n. 5685, p. 788–792, 2004. Availabe at: <https://www.science.org/doi/abs/10.1126/science.1096796>.

497 CUMMER, S. A.; CHRISTENSEN, J.; ALÙ, A. “Controlling sound with acoustic metamaterials”. *Nature Reviews Materials*, v. 1, p. 16001, 2016. Availabe at: <https://www.nature.com/articles/natrevmats20161>.

498 BERTOLDI, K. et al. “Flexible mechanical metamaterials”. *Nature Reviews Materials*, v. 2, p. 17066, 2017. Availabe at: <https://www.nature.com/articles/natrevmats201766>.

499 TIAN, X. et al. “Wireless body sensor networks based on metamaterial textiles”. *Nature Electronics*, v. 2, p. 243–251, 2019. Availabe at: <https://www.nature.com/articles/s41928-019-0257-7>.

500 BOIXO, S. et al. “Evidence for quantum annealing with more than one hundred qubits”. *Nature Physics*, v. 10, p. 218–224, 2014. Availabe at: <https://www.nature.com/articles/nphys2900>.

501 RENDLE, S. “Factorization machines”. In: *2010 IEEE International Conference on Data Mining*. [S.l.: s.n.], 2010. p. 995–1000.

502 KITAI, K. et al. “Designing metamaterials with quantum annealing and factorization machines”. *Phys. Rev. Res.*, American Physical Society, v. 2, p. 013319, Mar 2020. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevResearch.2.013319>.

503 IZAWA, S. et al. “Continuous black-box optimization with an ising machine and random subspace coding”. *Phys. Rev. Res.*, American Physical Society, v. 4, p. 023062, Apr 2022. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevResearch.4.023062>.

504 HARTREE, D. R. “The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part II. Some Results and Discussion”. *Mathematical Proceedings of the Cambridge Philosophical Society*, v. 24, n. 1, p. 111–132, 1928. ISSN 1469-8064. Availabe at: <https://www.cambridge.org/core/article/wave-mechanics-of-an-atom-with-a-noncoulomb-central-field-part-ii-some-results-and-discussion/5916E7A0DEC0A051B435688BE2ACD57E>.

505 SLATER, J. C. “Note on hartree’s method”. *Phys. Rev.*, American Physical Society, v. 35, p. 210–211, Jan 1930. Availabe at: <https://link.aps.org/doi/10.1103/PhysRev.35.210.2>.

506 FOCK, V. “„selfconsistent field“ mit austausch für natrium”. *Zeitschrift für Physik*, v. 62, n. 11, p. 795–805, Nov 1930. ISSN 0044-3328. Availabe at: <https://doi.org/10.1007/BF01330439>.

507 SLATER, J. C. “The theory of complex spectra”. *Phys. Rev.*, American Physical Society, v. 34, p. 1293–1322, Nov 1929. Availabe at: <https://link.aps.org/doi/10.1103/PhysRev.34.1293>.

508 BORN, M.; OPPENHEIMER, R. "Zur quantentheorie der molekeln". *Annalen der Physik*, v. 389, n. 20, p. 457–484, 1927. Availabe at: <https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.19273892002>.

509 SLATER, J. C. "The self consistent field and the structure of atoms". *Phys. Rev.*, American Physical Society, v. 32, p. 339–348, Sep 1928. Availabe at: <https://link.aps.org/doi/10.1103/PhysRev.32.339>.

510 Weiss, Pierre. "L'hypothèse du champ moléculaire et la propriété ferromagnétique". *J. Phys. Theor. Appl.*, v. 6, n. 1, p. 661–690, 1907. Availabe at: <https://doi.org/10.1051/jphystap:019070060066100>.

511 CHAIKIN, P. M.; LUBENSKY, T. C. "*Principles of condensed matter physics*". [S.l.]: Cambridge University Press, 2013.

512 HEAD-GORDON, M. et al. "A doubles correction to electronic excited states from configuration interaction in the space of single substitutions". *Chemical Physics Letters*, Elsevier BV, v. 219, n. 1–2, p. 21–29, Mar 1994. Availabe at: <http://dx.doi.org/10.1016/0009-2614(94)00070-0>.

513 MØLLER, C.; PLESSET, M. S. "Note on an approximation treatment for many-electron systems". *Phys. Rev.*, American Physical Society, v. 46, n. 7, p. 618–622, Oct 1934. Availabe at: <https://link.aps.org/doi/10.1103/PhysRev.46.618>.

514 PURVIS, G. D.; BARTLETT, R. J. "A full coupled-cluster singles and doubles model: The inclusion of disconnected triples". *The Journal of Chemical Physics*, v. 76, n. 4, p. 1910–1918, 1982. Availabe at: <https://doi.org/10.1063/1.443164>.

515 PERDEW, J. P. et al. "Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation". *Phys. Rev. B*, American Physical Society, v. 46, p. 6671–6687, Sep 1992. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevB.46.6671>.

516 BECKE, A. D. "Perspective: Fifty years of density-functional theory in chemical physics". *The Journal of Chemical Physics*, v. 140, n. 18, p. 18A301, 04 2014. ISSN 0021-9606. Availabe at: <https://doi.org/10.1063/1.4869598>.

517 HEHRE, W. J.; DITCHFIELD, R.; POPLE, J. A. "Self-consistent molecular orbital methods. xii. further extensions of gaussian-type basis sets for use in molecular orbital studies of organic molecules". *The Journal of Chemical Physics*, v. 56, p. 2257–2261, 1972.

518 HARIHARAN, P. C.; POPLE, J. A. "The influence of polarization functions on molecular orbital hydrogenation energies". *Theoretica Chimica Acta*, v. 28, p. 213–222, 1973.

519 FRANCL, M. M. et al. "Self-consistent molecular orbital methods. xxiii. a polarization-type basis set for second-row elements". *The Journal of Chemical Physics*, v. 77, p. 3654–3665, 1982.

520 DUNNING THOM H., J. "Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen". *The Journal of Chemical Physics*, v. 90, p. 1007–1023, 1989.

521 KENDALL, R. A.; DUNNING T. H., J.; HARRISON, R. J. "Electron affinities of the first-row atoms revisited. systematic basis sets and wave functions". *The Journal of Chemical Physics*, v. 96, p. 6796–6806, 1992.

522 WOON, D. E.; DUNNING T. H., J. "Gaussian basis sets for use in correlated molecular calculations. v. core-valence basis sets for boron through neon". *The Journal of Chemical Physics*, v. 103, p. 4572–4585, 1995.

523 WEIGEND, F.; AHLRICHS, R. "Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for h to rn: Design and assessment of accuracy". *Physical Chemistry Chemical Physics*, v. 7, p. 3297–3305, 2005.

524 WHITTEN, J. L. "Coulombic potential energy integrals and approximations". *The Journal of Chemical Physics*, v. 58, p. 4496–4501, 1973.

525 DUNLAP, B. I.; CONNOLLY, J. W. D.; SABIN, J. R. "On some approximations in applications of xα theory". *The Journal of Chemical Physics*, v. 71, p. 3396–3402, 1979.

526 FEYEREISEN, M.; FITZGERALD, G.; KOMORNICKI, A. "Use of approximate integrals in ab initio theory: An application in mp2 energy calculations". *Chemical Physics Letters*, v. 208, p. 359–363, 1993.

527 WEIGEND, F. "Accurate coulomb-fitting basis sets for h to rn". *Physical Chemistry Chemical Physics*, v. 8, p. 1057–1065, 2006.

528 BEEBE, N. H. F.; LINDERBERG, B. "Simplifications in the generation and transformation of two-electron integrals in molecular calculations". *International Journal of Quantum Chemistry*, v. 12, p. 683–705, 1977.

529 RAGHAVACHARI, K. et al. "A fifth-order perturbation comparison of electron correlation theories". *Chemical Physics Letters*, v. 157, p. 479–483, 1989.

530 HAY, P. J.; WADT, W. R. "Ab initio effective core potentials for molecular calculations. potentials for the transition metal atoms sc to hg". *The Journal of Chemical Physics*, v. 82, p. 270–283, 1985.

531 WADT, W. R.; HAY, P. J. "Ab initio effective core potentials for molecular calculations. ii. the main group elements na to bi". *The Journal of Chemical Physics*, v. 82, p. 284–298, 1985.

532 ANDRAE, D. et al. "Energy-adjusted ab initio pseudopotentials for the second and third row transition elements". *Theoretica Chimica Acta*, v. 77, p. 123–141, 1990.

533 HELGAKER, T. et al. "Basis-set convergence of correlated calculations on water". *The Journal of Chemical Physics*, v. 106, p. 9639–9646, 1997.

534 HALKIER, A. et al. "Basis-set convergence in correlated calculations on ne, n_2, and h_2o". *Chemical Physics Letters*, v. 286, p. 243–252, 1998.

535 FELLER, D. "Application of systematic sequences of wave functions to the water molecule". *The Journal of Chemical Physics*, v. 98, p. 7059–7071, 1993.

536 WOON, D. E.; DUNNING THOM H., J. "Gaussian basis sets for use in correlated molecular calculations. iii. the atoms aluminum through argon". *The Journal of Chemical Physics*, v. 98, n. 2, p. 1358–1371, 01 1993. ISSN 0021-9606. Availabe at: <https://doi.org/10.1063/1.464303>.

537 BOYS, S.; BERNARDI, F. "The calculation of small molecular interactions by the differences of separate total energies. some procedures with reduced errors". *Molecular Physics*, Taylor & Francis, v. 19, n. 4, p. 553–566, 1970. Availabe at: <https://doi.org/10.1080/00268977000101561>.

538 CONDON, E. U. "The theory of complex spectra". *Physical Review*, v. 36, n. 7, p. 1121–1133, 1930.

539 DAVIDSON, E. R. "The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices". *Journal of Computational Physics*, v. 17, n. 1, p. 87–94, 1975.

540 SCHAEFER, H. F. "*Methods of electronic structure theory*". New York, NY: Springer, 2013. (Modern Theoretical Chemistry).

541 OLSEN, J. et al. "Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces". *Journal of Chemical Physics*, v. 89, p. 2185–2192, 1988.

542 SHAVITT, I.; BARTLETT, R. J. "*Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*". Cambridge: Cambridge University Press, 2009. ISBN 978-0-521-87088-0.

543 LANGHOFF, S. R.; DAVIDSON, E. R. "Configuration interaction calculations on the nitrogen molecule". *International Journal of Quantum Chemistry*, v. 8, n. 1, p. 61–72, 1974.

544 POPLE, J. A.; HEAD-GORDON, M.; RAGHAVACHARI, K. "Quadratic configuration interaction. a general technique for determining electron correlation energies". *The Journal of Chemical Physics*, v. 87, n. 10, p. 5968–5975, 1987.

545 PALDUS, J. "Linked-cluster expansions and the coupled-cluster method in quantum chemistry". In: KUCHITSU, K. (Ed.). *New Horizons of Quantum Chemistry*. Berlin, Heidelberg: Springer, 1989, (Lecture Notes in Chemistry, v. 52). p. 99–194.

546 COESTER, F. "Bound states of a many-particle system". *Nucl. Phys.*, v. 7, p. 421–424, 1958.

547 COESTER, F.; KÜMMEL, H. "Short-range correlations in nuclear wave functions". *Nucl. Phys.*, v. 17, p. 477–485, 1960.

548 ČÍŽEK, J. "On the correlation problem in atomic and molecular systems. calculation of wavefunction components in ursell-type expansion using quantum-field theoretical methods". *J. Chem. Phys.*, v. 45, n. 11, p. 4256–4266, 1966.

549 WATTS, J. D.; GAUSS, J.; BARTLETT, R. J. "Coupled-cluster methods with noniterative triple excitations for restricted open-shell hartree–fock and other general single determinant reference functions. energies and analytical gradients". *J. Chem. Phys.*, v. 98, n. 11, p. 8718–8733, 1993.

550 LEE, T. J.; TAYLOR, P. R. "A diagnostic for determining the quality of single-reference electron correlation methods". *Int. J. Quantum Chem.*, v. 36, n. S23, p. 199–207, 1989.

551 JANSSEN, C. L.; NIELSEN, I. M. B. "New diagnostics for coupled-cluster and møller–plesset perturbation theory". *Chemical Physics Letters*, v. 290, p. 423–430, 1998.

552 CREMER, D. "Møller–plesset perturbation theory: From small molecules to many-electron systems of chemical interest". *WIREs Computational Molecular Science*, v. 1, n. 4, p. 509–530, 2011.

553 CHAN, G. K.-L.; SHARMA, S. "The density matrix renormalization group in quantum chemistry". *Annual Review of Physical Chemistry*, v. 62, p. 465–481, 2011.

554 SCHOLLWÖCK, U. "The density-matrix renormalization group in the age of matrix product states". *Annals of Physics*, v. 326, n. 1, p. 96–192, 2011.

555 AND. "Large-scale parallel tensor network quantum simulator". *Submitted to a conference*, 2020.

556 ROOS, B. O. "The complete active space self-consistent field method and its applications in electronic structure calculations". In: LAWLEY, K. P. (Ed.). *Advances in Chemical Physics: Ab Initio Methods in Quantum Chemistry II*. New York: John Wiley & Sons, 1987. v. 69, p. 399–445.

557 WERNER, H.-J.; KNOWLES, P. J. "An efficient internally contracted multiconfiguration–reference configuration interaction method". *The Journal of Chemical Physics*, v. 89, n. 9, p. 5803–5814, 1988.

558 FOULKES, W. M. C. et al. "Quantum Monte Carlo simulations of solids". *Reviews of Modern Physics*, v. 73, n. 1, p. 33–83, 2001.

559 ZHANG, S.; KRAKAUER, H. "Quantum Monte Carlo method using phase-free random walks with Slater determinants". *Physical Review Letters*, v. 90, n. 13, p. 136401, 2003.

560 ROMERO, J. et al. "Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz". *Quantum Science and Technology*, v. 4, p. 014008, 10 2018. ISSN 2058-9565. Availabe at: <https://iopscience.iop.org/article/10.1088/2058-9565/aad3e4>.

561 BARKOUTSOS, P. K. et al. "Quantum algorithms for electronic structure calculations: Particle-hole hamiltonian and optimized wave-function expansions". *Physical Review A*, American Physical Society (APS), v. 98, n. 2, ago. 2018. ISSN 2469-9934. Availabe at: <http://dx.doi.org/10.1103/PhysRevA.98.022322>.

562 WHITFIELD, J. D.; BIAMONTE, J.; ASPURU-GUZIK, A. "Simulation of electronic structure hamiltonians using a quantum computer". *Mol. Phys.*, v. 109, n. 5, p. 735–750, 2011.

563 TRANTER, A. et al. "The Bravyi–Kitaev transformation: Properties and implementation". *International Journal of Quantum Chemistry*, v. 115, n. 19, p. 1431–1441, 2015.

564 HUGGINS, W. et al. “Efficient and noise resilient measurements for quantum chemistry on near-term quantum computers”. *Journal of Chemical Theory and Computation*, v. 17, n. 10, p. 6502–6517, 2021. Section comparing JW, parity, and BK encodings.

565 HAVLÍČEK, V.; TROYER, M.; WHITFIELD, J. D. “Operator locality in quantum simulation of fermionic models”. *Physical Review A*, v. 95, n. 3, p. 032332, 2017.

566 SETIA, K. et al. “Reducing qubit requirements for quantum simulations using molecular point group symmetries”. *J. Chem. Theory Comput.*, v. 16, n. 10, p. 6091–6097, 2020.

567 VERTELETSKYI, V.; YEN, T.-C.; IZMAYLOV, A. F. “Measurement optimization in the variational quantum eigensolver using a minimum clique cover”. *The Journal of Chemical Physics*, v. 152, n. 12, p. 124114, 2020.

568 GRIMSLEY, H. R. et al. “An adaptive variational algorithm for exact molecular simulations on a quantum computer”. *Nature Communications*, v. 10, n. 1, p. 3007, 2019. Availabe at: <https://doi.org/10.1038/s41467-019-10988-2>.

569 LEVINE, B. G. et al. “CAS without SCF—Why to use CASCI and where to get the orbitals”. *The Journal of Chemical Physics*, v. 154, n. 9, p. 090902, 03 2021. ISSN 0021-9606. Availabe at: <https://doi.org/10.1063/5.0042147>.

570 SAYFUTYAROVA, E. R. et al. “Automated construction of molecular active spaces from atomic valence orbitals”. *Journal of Chemical Theory and Computation*, American Chemical Society, v. 13, n. 9, p. 4063–4078, 09 2017. Availabe at: <https://doi.org/10.1021/acs.jctc.7b00128>.

571 LI, W. et al. “Toward practical quantum embedding simulation of realistic chemical systems on near-term quantum computers”. *Chemical Science*, Royal Society of Chemistry (RSC), v. 13, n. 31, p. 8953–8962, 2022. ISSN 2041-6539. Availabe at: <http://dx.doi.org/10.1039/d2sc01492k>.

572 METZ, B. et al. “Ipcc special report on carbon dioxide capture and storage”. 7 2005. Availabe at: <https://www.osti.gov/biblio/20740954>.

573 International Energy Agency,. “*Direct Air Capture: A Key Technology for Net Zero*”. OECD Publishing, 2022. Availabe at: <https://www.oecd-ilibrary.org/content/publication/bbd20707-en>.

574 SANZ-PÉREZ, E. S. et al. “Direct capture of co2 from ambient air”. *Chemical reviews*, ACS Publications, v. 116, n. 19, p. 11840–11876, 2016.

575 SIEGELMAN, R. L.; KIM, E. J.; LONG, J. R. “Porous materials for carbon dioxide separations”. *Nature Materials*, v. 20, n. 8, p. 1060–1072, 2021. Availabe at: <https://doi.org/10.1038/s41563-021-01054-8>.

576 SABATINO, F. et al. “A comparative energy and costs assessment and optimization for direct air capture technologies”. *Joule*, Elsevier, v. 5, n. 8, p. 2047–2076, 2021.

577 KUMAR, A. et al. “Direct air capture of co2 by physisorbent materials”. *Angewandte Chemie International Edition*, Wiley Online Library, v. 54, n. 48, p. 14372–14377, 2015.

578 DECOSTE, J. B. et al. "The effect of water adsorption on the structure of the carboxylate containing metal–organic frameworks cu-btc, mg-mof-74, and uio-66". *Journal of Materials Chemistry A*, Royal Society of Chemistry, v. 1, n. 38, p. 11922–11932, 2013.

579 SUMIDA, K. et al. "Carbon dioxide capture in metal–organic frameworks". *Chemical Reviews*, American Chemical Society, v. 112, n. 2, p. 724–781, 02 2012. Availabe at: <https://doi.org/10.1021/cr2003272>.

580 YAO, Z. et al. "Inverse design of nanoporous crystalline reticular materials with deep generative models". *Nature Machine Intelligence*, Nature Publishing Group UK London, v. 3, n. 1, p. 76–86, 2021.

581 WILMER, C. E. et al. "Large-scale screening of hypothetical metal–organic frameworks". *Nature Chemistry*, v. 4, n. 2, p. 83–89, 2012. Availabe at: <https://doi.org/10.1038/nchem.1192>.

582 BOYD, P. G. et al. "Data-driven design of metal–organic frameworks for wet flue gas co2 capture". *Nature*, v. 576, n. 7786, p. 253–256, 2019. Availabe at: <https://doi.org/10.1038/s41586-019-1798-7>.

583 MOGHADAM, P. Z. et al. "Development of a Cambridge Structural Database Subset: A Collection of Metal–Organic Frameworks for Past, Present, and Future". *Chemistry of Materials*, American Chemical Society, v. 29, n. 7, p. 2618–2625, 3 2017. Availabe at: <https://acs.figshare.com/articles/dataset/Development_of_a_Cambridge_Structural_Database_Subset_A_Collection_of_Metal_Organic_Frameworks_for_Past_Present_and_Future/4794040>.

584 DAHALE, G. R. "*Quantum Simulations for Carbon Capture on Metal-Organic Frameworks*". 2023. Availabe at: <https://arxiv.org/abs/2311.12411>.

585 GREENE-DINIZ, G. et al. "Modelling carbon capture on metal-organic frameworks with quantum computing". *EPJ Quantum Technology*, Springer Berlin Heidelberg, v. 9, n. 1, p. 37, 2022.

586 MCCLEAN, J. R. et al. "The theory of variational hybrid quantum-classical algorithms". *New Journal of Physics*, v. 18, n. 2, p. 023023–023023, 2016. Availabe at: <https://iopscience.iop.org/article/10.1088/1367-2630/18/2/023023>.

587 SUN, Q. et al. "Pyscf: the python-based simulations of chemistry framework". *WIREs Computational Molecular Science*, John Wiley & Sons, Ltd, v. 8, n. 1, p. e1340, 2023/11/22 2018. Availabe at: <https://doi.org/10.1002/wcms.1340>.

588 SUN, Q. et al. "Recent developments in the PySCF program package". *The Journal of Chemical Physics*, v. 153, n. 2, p. 024109, 07 2020. ISSN 0021-9606. Availabe at: <https://doi.org/10.1063/5.0006074>.

589 IBM. "*Qiskit: An Open-source Framework for Quantum Computing*". 2023. <https://qiskit.org/>. <https://www.ibm.com/quantum/qiskit>.

590 JAVADI-ABHARI, A. et al. "*Quantum computing with Qiskit*". 2024. Availabe at: <https://arxiv.org/abs/2405.08810>.

591 LINSTROM, P. J.; MALLARD, E. W. "*NIST Chemistry WebBook, NIST Standard Reference Database Number 69*". [S.l.]: National Institute of Standards and Technology, Gaithersburg MD, 20899, 1997.

592 EVANGELISTA, F. A.; CHAN, G. K.-L.; SCUSERIA, G. E. "Exact parameterization of fermionic wave functions via unitary coupled cluster theory". *The Journal of Chemical Physics*, v. 151, n. 24, p. 244112, 12 2019. ISSN 0021-9606. Availabe at: <https://doi.org/10.1063/1.5133059>.

593 GRIMSLEY, H. R. et al. "Is the trotterized uccsd ansatz chemically well-defined?" *Journal of Chemical Theory and Computation*, American Chemical Society (ACS), v. 16, n. 1, p. 1–6, dez. 2019. ISSN 1549-9626. Availabe at: <http://dx.doi.org/10.1021/acs.jctc.9b01083>.

594 TILLY, J. et al. "The variational quantum eigensolver: a review of methods and best practices". *Physics Reports*, Elsevier, v. 986, p. 1–128, 2022.

595 BERG, E. van den; MINEV, Z. K.; TEMME, K. "Model-free readout-error mitigation for quantum expectation values". *Phys. Rev. A*, American Physical Society, v. 105, p. 032620, Mar 2022. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevA.105.032620>.

596 MORSE, P. M. "Diatomic molecules according to the wave mechanics. ii. vibrational levels". *Phys. Rev.*, American Physical Society, v. 34, p. 57–64, Jul 1929. Availabe at: <https://link.aps.org/doi/10.1103/PhysRev.34.57>.

597 KöKçAM-DEMIR, Ü. et al. "Coordinatively unsaturated metal sites (open metal sites) in metal–organic frameworks: design and applications". *Chem. Soc. Rev.*, The Royal Society of Chemistry, v. 49, p. 2751–2798, 2020. Availabe at: <http://dx.doi.org/10.1039/C9CS00609E>.

598 FEDOROV, D. A. et al. "VQE method: a short survey and recent developments". *Mater. Theory*, Springer Science and Business Media LLC, v. 6, n. 1, jan. 2022.

599 PICCINELLI, S. et al. "*Quantum chemistry with provable convergence via randomized sample-based quantum diagonalization*". 2025. Availabe at: <https://arxiv.org/abs/2508.02578>.

600 DANILOV, D. et al. "*Enhancing the accuracy and efficiency of sample-based quantum diagonalization with phaseless auxiliary-field quantum Monte Carlo*". 2025. Availabe at: <https://arxiv.org/abs/2503.05967>.

601 KANNO, K. et al. "*Quantum-Selected Configuration Interaction: classical diagonalization of Hamiltonians in subspaces selected by quantum computers*". 2023. Availabe at: <https://arxiv.org/abs/2302.11320>.

602 REINHOLDT, P. et al. "Critical limitations in quantum-selected configuration interaction methods". *J. Chem. Theory Comput.*, American Chemical Society (ACS), v. 21, n. 14, p. 6811–6822, jul. 2025.

603 SMITH, T. et al. "Quantum-centric simulation of hydrogen abstraction by sample-based quantum diagonalization and entanglement forging". 2025.

604 LIEPUONIUTE, I. et al. "*Quantum-Centric Study of Methylene Singlet and Triplet States*". 2024. Availabe at: <https://arxiv.org/abs/2411.04827>.

605 EDDINS, A. et al. “Doubling the size of quantum simulators by entanglement forging”. *PRX Quantum*, American Physical Society (APS), v. 3, n. 1, jan. 2022. ISSN 2691-3399. Availabe at: <http://dx.doi.org/10.1103/PRXQuantum.3.010309>.

606 SHIRAKAWA, T. et al. “*Closed-loop calculations of electronic structure on a quantum processor and a classical supercomputer at full scale*”. 2025. Availabe at: <https://arxiv.org/abs/2511.00224>.

607 COHEN, A. J.; MORI-SáNCHEZ, P.; YANG, W. “Insights into current limitations of density functional theory”. *Science*, v. 321, n. 5890, p. 792–794, 2008. Availabe at: <https://www.science.org/doi/abs/10.1126/science.1158722>.

608 JONES, R. O.; GUNNARSSON, O. “The density functional formalism, its applications and prospects”. *Rev. Mod. Phys.*, American Physical Society, v. 61, p. 689–746, Jul 1989. Availabe at: <https://link.aps.org/doi/10.1103/RevModPhys.61.689>.

609 INCE, J. et al. “Overview of emerging hybrid and composite materials for space applications”. *Advanced Composites and Hybrid Materials*, v. 6, p. 130, 06 2023.

610 LIU, T. et al. “Reversible Discharge Products in Li-Air Batteries”. *Advanced Materials*, v. 35, 03 2023.

611 PAOLA, C. D. et al. “Platinum-based catalysts for oxygen reduction reaction simulated with a quantum computer”. *npj Computational Materials*, Springer Science and Business Media LLC, v. 10, n. 1, dez. 2024. ISSN 2057-3960. Availabe at: <http://dx.doi.org/10.1038/s41524-024-01460-x>.

612 KANG, C. et al. “Optimized quantum phase estimation for simulating electronic states in various energy regimes”. *Journal of Chemical Theory and Computation*, v. 18, n. 11, p. 6567–6576, 2022. PMID: 36201845. Availabe at: <https://doi.org/10.1021/acs.jctc.2c00577>.

613 CORTES, C. L. et al. “*Assessing the query complexity limits of quantum phase estimation using symmetry aware spectral bounds*”. 2024. Availabe at: <https://arxiv.org/abs/2403.04737>.

614 CLARY, J. M. et al. “Exploring the scaling limitations of the variational quantum eigensolver with the bond dissociation of hydride diatomic molecules”. *Int. J. Quantum Chem.*, Wiley, v. 123, n. 11, jun. 2023.

615 REIHER, M. et al. “Elucidating reaction mechanisms on quantum computers”. *Proceedings of the National Academy of Sciences*, v. 114, n. 29, p. 7555–7560, 2017. Availabe at: <https://www.pnas.org/doi/abs/10.1073/pnas.1619152114>.

616 WECKER, D. et al. “Gate-count estimates for performing quantum chemistry on small quantum computers”. *Phys. Rev. A*, American Physical Society, v. 90, p. 022305, Aug 2014. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevA.90.022305>.

617 HOLMES, Z. et al. “Connecting ansatz expressibility to gradient magnitudes and barren plateaus”. *PRX Quantum*, American Physical Society, v. 3, p. 010313, Jan 2022. Availabe at: <https://link.aps.org/doi/10.1103/PRXQuantum.3.010313>.

618 SINGKANIPA, P.; LIDAR, D. A. "Beyond unital noise in variational quantum algorithms: noise-induced barren plateaus and limit sets". *Quantum*, Verein zur Forderung des Open Access Publizierens in den Quantenwissenschaften, v. 9, n. 1617, p. 1617, jan. 2025.

619 MOTTA, M. et al. "Bridging physical intuition and hardware efficiency for correlated electronic states: the local unitary cluster jastrow ansatz for electronic structure". *Chem. Sci.*, The Royal Society of Chemistry, v. 14, p. 11213–11227, 2023. Availabe at: <http://dx.doi.org/10.1039/D3SC02516K>.

620 KRESSE, G.; HAFNER, J. "Ab initio molecular dynamics for liquid metals". *Phys. Rev. B*, American Physical Society, v. 47, p. 558–561, 1993.

621 KRESSE, G.; FURTHMÜLLER, J. "Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set". *Comput. Mater. Sci.*, Elsevier, v. 6, p. 15–50, 1996.

622 KRESSE, G.; JOUBERT, D. "From ultrasoft pseudopotentials to the projector augmented-wave method". *Phys. Rev. B*, American Physical Society, v. 59, p. 1758–1775, 1999.

623 GOEDECKER, S.; TETER, M.; HUTTER, J. "Separable dual-space gaussian pseudopotentials". *Phys. Rev. B*, American Physical Society, v. 54, p. 1703–1710, Jul 1996. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevB.54.1703>.

624 PERDEW, J. P.; BURKE, K.; ERNZERHOF, M. "Generalized gradient approximation made simple". *Phys. Rev. Lett.*, American Physical Society, v. 77, n. 18, p. 3865–3868, Oct 1996. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevLett.77.3865>.

625 M.W., C. J. et al. "*NIST-JANAF Thermochemical Tables*". National Institute of Standards and Technology, 1998. Availabe at: <https://janaf.nist.gov/>.

626 RADIN, M. D. et al. "Lithium peroxide surfaces are metallic, while lithium oxide surfaces are not". *Journal of the American Chemical Society*, v. 134, n. 2, p. 1093–1103, 2012. PMID: 22148314. Availabe at: <https://doi.org/10.1021/ja208944x>.

627 HUMMELSHØJ, J. S. et al. "Communications: Elementary oxygen electrode reactions in the aprotic li-air battery". *J. Chem. Phys.*, AIP Publishing, v. 132, n. 7, p. 071101, fev. 2010.

628 SERIANI, N. "Ab initio thermodynamics of lithium oxides: from bulk phases to nanoparticles". *Nanotechnology*, IOP Publishing, v. 20, n. 44, p. 445703, nov. 2009.

629 KELLER, S. et al. "Selection of active spaces for multiconfigurational wavefunctions". *J. Chem. Phys.*, AIP Publishing, v. 142, n. 24, p. 244104, jun. 2015.

630 STEIN, C. J.; REIHER, M. "Automated selection of active orbital spaces". *J. Chem. Theory Comput.*, American Chemical Society (ACS), v. 12, n. 4, p. 1760–1771, abr. 2016.

631 HIRJIBEHEDIN, C. F. et al. "Large magnetic anisotropy of a single atomic spin embedded in a surface molecular network". *Science*, American Association for the Advancement of Science (AAAS), v. 317, n. 5842, p. 1199–1203, ago. 2007.

632 RAU, I. G. et al. “Reaching the magnetic anisotropy limit of a 3d metal atom”. *Science*, American Association for the Advancement of Science (AAAS), v. 344, n. 6187, p. 988–992, maio 2014.

633 BAUMANN, S. et al. “Origin of perpendicular magnetic anisotropy and large orbital moment in fe atoms on MgO”. *Phys. Rev. Lett.*, American Physical Society (APS), v. 115, n. 23, p. 237202, dez. 2015.

634 ALBERTINI, O. R.; LIU, A. Y.; JONES, B. A. “Site-dependent magnetism of ni adatoms on MgO/Ag(001)”. *Phys. Rev. B Condens. Matter Mater. Phys.*, American Physical Society (APS), v. 91, n. 21, jun. 2015.

635 PIPEK, J.; MEZEY, P. G. “A fast intrinsic localization procedure applicable for ab-initio and semiempirical linear combination of atomic orbital wave functions”. *The Journal of Chemical Physics*, v. 90, n. 9, p. 4916–4926, 05 1989. ISSN 0021-9606. Availabe at: <https://doi.org/10.1063/1.456588>.

636 LEHTOLA, S.; JóNSSON, H. “Pipek–mezey orbital localization using various partial charge estimates”. *Journal of Chemical Theory and Computation*, v. 10, n. 2, p. 642–649, 2014. PMID: 26580041. Availabe at: <https://doi.org/10.1021/ct401016x>.

637 BOGUSLAWSKI, K.; MARTI, K. H.; REIHER, M. “Construction of casci-type wave functions for very large active spaces”. *The Journal of Chemical Physics*, AIP Publishing, v. 134, n. 22, jun. 2011. ISSN 1089-7690. Availabe at: <http://dx.doi.org/10.1063/1.3596482>.

638 The ffsim developers. “*ffsim: Faster simulations of fermionic quantum circuits.*” 2025. Availabe at: <https://github.com/qiskit-community/ffsim>.

639 NATION, P. D.; TREINISH, M. “*mapomatic*”. 2024. <https://github.com/qiskit-community/mapomatic>.

640 ROBERTSON, N. et al. “Approximate quantum compiling for quantum simulation: A tensor network based approach”. *ACM Transactions on Quantum Computing*, Association for Computing Machinery (ACM), v. 6, n. 3, p. 1–15, maio 2025. ISSN 2643-6817. Availabe at: <http://dx.doi.org/10.1145/3731251>.

641 GARRISON, J. R. et al. “*Qiskit addon: Approximate Quantum Compilation with Tensor Networks*”. 2024. <https://github.com/Qiskit/qiskit-addon-aqc-tensor>.

642 ARELLANO, E. et al. “*Qiskit addon: Dice eigensolver*”. 2024. <https://github.com/Qiskit/qiskit-addon-dice-solver>.

643 GALLOWAY, T. A. et al. “Oxygen reactions on pt{hkl} in a non-aqueous na+ electrolyte: site selective stabilisation of a sodium peroxy species”. *Chem. Sci.*, Royal Society of Chemistry (RSC), v. 10, n. 10, p. 2956–2964, mar. 2019.

644 CREMER, D. “From configuration interaction to coupled cluster theory: The quadratic configuration interaction approach”. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, Wiley, v. 3, n. 5, p. 482–503, set. 2013.

645 ANDERSSON, K.; MALMQVIST, P.-Å.; ROOS, B. O. “Second-order perturbation theory with a complete active space self-consistent field reference function”. *J. Chem. Phys.*, v. 96, p. 1218–1226, 1992.

646 FORSBERG, N.; MALMQVIST, P.-Å. "Multiconfiguration perturbation theory with imaginary level shift". *Chem. Phys. Lett.*, v. 274, p. 196–204, 1997.

647 GHIGO, G.; ROOS, B. O.; MALMQVIST, P.-Å. "A modified definition of the zeroth-order hamiltonian in multiconfigurational perturbation theory (caspt2): The ipea shift". *Chem. Phys. Lett.*, v. 396, p. 142–149, 2004.

648 SHARMA, S. et al. "Semistochastic heat-bath configuration interaction method: Selected configuration interaction with semistochastic perturbation theory". *J. Chem. Phys.*, v. 147, p. 034103, 2017.

649 LI, J. et al. "Fast semistochastic heat-bath configuration interaction". *J. Chem. Phys.*, v. 149, p. 214110, 2018.

650 BURKE, K.; WAGNER, L. O. "Dft in a nutshell". *International Journal of Quantum Chemistry*, v. 113, n. 2, p. 96–101, 2013.

651 PAVARINI, E. "Solving the strong-correlation problem in materials". *La Rivista del Nuovo Cimento*, v. 44, n. 11, p. 597–640, 2021.

652 SACHDEV, S. "Understanding correlated electron systems by a classification of mott insulators". *Annals of Physics*, v. 303, n. 1, p. 226–246, 2003. ISSN 0003-4916.

653 BUSSMANN-HOLDER, A.; KELLER, H. "High-temperature superconductors: underlying physics and applications". *Zeitschrift für Naturforschung B*, v. 75, n. 1-2, p. 3–14, 2020. Availabe at: <https://doi.org/10.1515/znb-2019-0103>.

654 SOLOVYEV, I. V. "Combining dft and many-body methods to understand correlated materials". *Journal of Physics: Condensed Matter*, v. 20, n. 29, p. 293201, jun 2008.

655 MANDAL, S. et al. "Systematic beyond-dft study of binary transition metal oxides". *npj Computational Materials*, v. 5, n. 1, p. 115, 2019.

656 SCHLüTER, M.; SHAM, L. "Density-functional theory of the band gap". In: LöWDIN, P.-O. (Ed.). *Density Functional Theory of Many-Fermion Systems*. [S.l.]: Academic Press, 1990, (Advances in Quantum Chemistry, v. 21). p. 97–112.

657 MA, J.; WANG, L.-W. "Using wannier functions to improve solid band gap predictions in density functional theory". *Scientific Reports*, v. 6, n. 1, p. 24924, 2016.

658 PERDEW, J. P.; SCHMIDT, K. "Jacob's ladder of density functional approximations for the exchange-correlation energy". *AIP Conference Proceedings*, v. 577, n. 1, p. 1–20, 07 2001. ISSN 0094-243X.

659 TAO, J. et al. "Climbing the density functional ladder: Nonempirical meta–generalized gradient approximation designed for molecules and solids". *Phys. Rev. Lett.*, American Physical Society, v. 91, p. 146401, Sep 2003.

660 SETTEN, M. J. van; WEIGEND, F.; EVERS, F. "The gw-method for quantum chemistry applications: Theory and implementation". *Journal of Chemical Theory and Computation*, v. 9, n. 1, p. 232–246, 2013.

661 ZHANG, Y.; CALLAWAY, J. "Extended hubbard model in two dimensions". *Phys. Rev. B*, American Physical Society, v. 39, p. 9397–9404, May 1989.

662 CAMPO, V. L.; COCOCCIONI, M. "Extended dft + u + v method with on-site and inter-site electronic interactions". *Journal of Physics: Condensed Matter*, v. 22, n. 5, p. 055602, jan 2010.

663 TIMROV, I.; MARZARI, N.; COCOCCIONI, M. "Hp – a code for the calculation of hubbard parameters using density-functional perturbation theory". *Computer Physics Communications*, v. 279, p. 108455, 2022. ISSN 0010-4655.

664 NAVEAS, N. et al. "First-principles calculations of hematite ($\alpha - fe_2o_3$) by self-consistent dft+u+v". *iScience*, Elsevier, v. 26, n. 2, 2023. ISSN 2589-0042. Availabe at: <https://doi.org/10.1016/j.isci.2023.106033>.

665 KIVELSON, S. et al. "Missing bond-charge repulsion in the extended hubbard model: Effects in polyacetylene". *Phys. Rev. Lett.*, American Physical Society, v. 58, p. 1899–1902, May 1987. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevLett.58.1899>.

666 CASTET, F.; FRITSCH, A.; DUCASSE, L. "Determination of the Coulombic Interaction Parameters of the Extended Hubbard Model in the Organic Conductors". *Journal de Physique I*, EDP Sciences, v. 6, n. 4, p. 583–597, 1996. Availabe at: <https://hal.science/jpa-00247204>.

667 LIAO, Y. D.; MENG, Z. Y.; XU, X. Y. "Valence bond orders at charge neutrality in a possible two-orbital extended hubbard model for twisted bilayer graphene". *Phys. Rev. Lett.*, American Physical Society, v. 123, p. 157601, Oct 2019. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevLett.123.157601>.

668 LUO, Z. et al. "High-tc superconductivity in la3ni2o7 based on the bilayer two-orbital t-j model". *npj Quantum Materials*, v. 9, n. 1, p. 61, 2024. Availabe at: <https://doi.org/10.1038/s41535-024-00668-w>.

669 LIANG, S.; PANG, H. "Approximate diagonalization using the density matrix renormalization-group method: A two-dimensional-systems perspective". *Phys. Rev. B*, American Physical Society, v. 49, p. 9214–9217, Apr 1994. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevB.49.9214>.

670 CEPERLEY, D. M.; ALDER, B. J. "Ground state of the electron gas by a stochastic method". *Phys. Rev. Lett.*, American Physical Society, v. 45, p. 566–569, Aug 1980. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevLett.45.566>.

671 TROYER, M.; WIESE, U.-J. "Computational complexity and fundamental limitations to fermionic quantum monte carlo simulations". *Phys. Rev. Lett.*, American Physical Society, v. 94, p. 170201, May 2005. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevLett.94.170201>.

672 PAN, G.; MENG, Z. Y. "The sign problem in quantum monte carlo simulations". In: _____. *Encyclopedia of Condensed Matter Physics*. Elsevier, 2024. p. 879–893. ISBN 9780323914086. Availabe at: <http://dx.doi.org/10.1016/B978-0-323-90800-9.00095-0>.

673 SOBCZAK, P. et al. “Magnetic properties and dmrg modeling of the 1d bimetallic thiocyanate bridged compound (CuL1)[Co(NCS)4] (l1=n-rac-5,12-me2-[14]-4,11-dienen4)”. *Polyhedron*, v. 28, n. 9, p. 1838–1841, 2009. ISSN 0277-5387. Proceedings of the 11th International Conference on Molecule-based Magnets (ICMM 2008). Availabe at: <https://www.sciencedirect.com/science/article/pii/S0277538709000734>.

674 MEZZACAPO, F. et al. “Ground-state properties of quantum many-body systems: entangled-plaquette states and variational monte carlo”. *New Journal of Physics*, IOP Publishing, v. 11, n. 8, p. 083026, ago. 2009. ISSN 1367-2630. Availabe at: <http://dx.doi.org/10.1088/1367-2630/11/8/083026>.

675 OHGOE, T. et al. “Demonstrating quantum computation for quasiparticle band structures”. *Phys. Rev. Res.*, American Physical Society, v. 6, p. L022022, Apr 2024. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevResearch.6.L022022>.

676 YOSHIOKA, N. et al. “Variational quantum simulation for periodic materials”. *Phys. Rev. Res.*, American Physical Society, v. 4, p. 013052, Jan 2022. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevResearch.4.013052>.

677 WU, D. et al. “Variational benchmarks for quantum many-body problems”. *Science*, American Association for the Advancement of Science (AAAS), v. 386, n. 6719, p. 296–301, out. 2024. ISSN 1095-9203. Availabe at: <http://dx.doi.org/10.1126/science.adg9774>.

678 ZHAO, G. et al. “Core mof db: A curated experimental metal-organic framework database with machine-learned properties for integrated material-process screening”. *Matter*, v. 8, n. 6, p. 102140, 2025. ISSN 2590-2385. Availabe at: <https://www.sciencedirect.com/science/article/pii/S2590238525001833>.

679 Buongiorno Nardelli, M. et al. “Paoflow: A utility to construct and operate on ab initio hamiltonians from the projections of electronic wavefunctions on atomic orbital bases, including characterization of topological materials”. *Computational Materials Science*, v. 143, p. 462–472, 2018. ISSN 0927-0256. Availabe at: <https://www.sciencedirect.com/science/article/pii/S0927025617306651>.

680 CERASOLI, F. T. et al. “Advanced modeling of materials with paoflow 2.0: New features and software design”. *Computational Materials Science*, v. 200, p. 110828, 2021. ISSN 0927-0256.

681 GIANNOZZI, P. et al. “Quantum espresso: a modular and open-source software project for quantum simulations of materials”. *Journal of Physics: Condensed Matter*, v. 21, n. 39, p. 395502 (19pp), 2009. Availabe at: <http://www.quantum-espresso.org>.

682 GIANNOZZI, P. et al. “Advanced capabilities for materials modelling with quantum espresso”. *Journal of Physics: Condensed Matter*, v. 29, n. 46, p. 465901, 2017. Availabe at: <http://stacks.iop.org/0953-8984/29/i=46/a=465901>.

683 GIANNOZZI, P. et al. “Quantum espresso toward the exascale”. *The Journal of Chemical Physics*, v. 152, n. 15, p. 154105, 2020. Availabe at: <https://doi.org/10.1063/5.0005082>.

684 TRUSHIN, E. et al. “Band gaps, ionization potentials, and electron affinities of periodic electron systems via the adiabatic-connection fluctuation-dissipation theorem”.

Phys. Rev. B, American Physical Society, v. 94, p. 075123, Aug 2016. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevB.94.075123>.

685 ROSEN, A. S. et al. "Machine learning the quantum-chemical properties of metal-organic frameworks for accelerated materials discovery". *Matter*, Elsevier, v. 4, n. 5, p. 1578–1597, 2021. ISSN 2590-2393.

686 ITO, T. et al. "Similarities in photoluminescence in hafnia and zirconia induced by ultraviolet photons". *Journal of Applied Physics*, v. 97, n. 5, p. 054104, 02 2005. ISSN 0021-8979. Availabe at: <https://doi.org/10.1063/1.1856220>.

687 KRÁLIK, B.; CHANG, E. K.; LOUIE, S. G. "Structural properties and quasiparticle band structure of zirconia". *Phys. Rev. B*, American Physical Society, v. 57, p. 7027–7036, Mar 1998. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevB.57.7027>.

688 ZANDIEHNADEM, F.; MURRAY, R.; CHING, W. "Electronic structures of three phases of zirconium oxide". *Physica B+C*, v. 150, n. 1, p. 19–24, 1988. ISSN 0378-4363. Availabe at: <https://www.sciencedirect.com/science/article/pii/037843638890099X>.

689 DASH, L. K. et al. "Electronic structure and electron energy-loss spectroscopy of zro_2 zirconia". *Phys. Rev. B*, American Physical Society, v. 70, p. 245116, Dec 2004. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevB.70.245116>.

690 HYBERTSEN, M. S.; LOUIE, S. G. "First-principles theory of quasiparticles: Calculation of band gaps in semiconductors and insulators". *Phys. Rev. Lett.*, American Physical Society, v. 55, p. 1418–1421, Sep 1985. Availabe at: <https://link.aps.org/doi/10.1103/PhysRevLett.55.1418>.