

Multistate iterative qubit coupled cluster (MS-iQCC): a quantum-inspired, state-averaged approach to ground- and excited-state energies

Robert A. Lang,^{1,2} Shashank G. Mehendale,^{1,2} Ilya G. Ryabinkin,³ and Artur F. Izmaylov^{1,2, a)}

¹⁾*Chemical Physics Theory Group, Department of Chemistry, University of Toronto, Toronto, Ontario, M5S 3H6, Canada*

²⁾*Department of Physical and Environmental Sciences, University of Toronto Scarborough, Toronto, Ontario, M1C 1A4, Canada*

³⁾*OTI Lumionics Inc., 3415 American Drive Unit 1, Mississauga, Ontario L4V 1T4, Canada*

(Dated: November 4, 2025)

We introduce the multistate iterative qubit coupled cluster (MS-iQCC) method, a quantum-inspired algorithm that runs efficiently on classical hardware and is designed to predict both ground and excited electronic states of molecules. Accurate excited-state energetics are essential for interpreting spectroscopy and chemical reactivity, but standard electronic structure methods are either too computationally expensive for larger systems or lose reliability in the presence of strong electron correlation. MS-iQCC addresses this challenge by simultaneously optimizing multiple electronic states in a single, state-averaged procedure that treats ground and excited states on equal footing. This removes the energetic bias that is introduced when excited states are computed one at a time and constrained to remain orthogonal to previously optimized states. The approach supports multireference electronic structure by allowing multideterminantal initial guesses and by adaptively building a compact exponential ansatz from a pool of qubit excitation generators. We apply MS-iQCC to H_4 , H_2O , N_2 , and C_2 , including strongly correlated geometries, and observe robust convergence of all targeted state energies to chemically meaningful accuracy across their potential energy surfaces.

I. INTRODUCTION

The accurate determination of ground and excited electronic states is essential for understanding and predicting molecular properties and reactivity. Many physical and chemical phenomena depend critically on excited-state energetics and potential energy surfaces^{1,2}. Classical computational chemistry provides a variety of approaches for these tasks, but their computational cost grows rapidly with system size. In particular, while multiconfigurational and perturbative methods such as MC-SCF^{3–6}, CASPT2^{7–9}, and DMRG^{10–14} can provide highly accurate results, they become prohibitively expensive for larger molecules. More computationally affordable methods, such as time-dependent density functional theory (TDDFT)^{15–17}, offer broader applicability but often lack the accuracy required to reliably describe complex excited-state phenomena. Therefore, in this work, we focus on developing an alternative approach motivated by quantum computational concepts but executable efficiently on classical hardware.

Recent developments in quantum algorithms, especially the variational quantum eigensolver (VQE)^{18–21} and qubit coupled cluster (QCC)^{22,23} frameworks, have demonstrated new ways to approximate molecular eigenstates. However, near-term quantum devices (NISQ) remain limited by coherence times and circuit depth constraints, making it challenging to implement deep circuits required for accurate state preparation²⁴ and

measurement²⁵. Consequently, quantum-inspired approaches—algorithms leveraging quantum formalisms but implemented on classical computers—offer an attractive intermediate step. Yet, another alternative that became popular recently is quantum sampling-based diagonalization (QSD) techniques^{26,27}. QSD uses an unoptimized unitary ansatz to produce Slater determinants by sampling the quantum state on a quantum computer. These Slater determinants are used to build a subspace for solving the Hamiltonian eigenvalue problem on a classical computer.

Our method, the multistate iterative qubit coupled cluster (MS-iQCC), belongs to this class of quantum-inspired techniques. Although originally motivated by the variational quantum eigensolver and iterative qubit coupled cluster (iQCC)²² algorithms, MS-iQCC is designed to run entirely on classical hardware. It retains the adaptive, exponential ansatz structure of iQCC but extends it to simultaneously describe multiple electronic states in a state-averaged, unbiased manner. This makes MS-iQCC particularly suited for classical computation today, while remaining directly relevant to future fault-tolerant quantum computing.

The adaptive exponential ansatz scheme used in MS-iQCC ensures separability for non-interacting fragments—an essential property shared with coupled cluster theory²⁸—but combines this with the variational flexibility of modern adaptive approaches^{29,30}. This separability also makes MS-iQCC conceptually preferable to selected configuration interaction methods^{31–34}. Compared to quantum excited-state methods such as the variational quantum deflation (VQD)^{35,36} and orthogonal state reduction variational eigensolver (OSRVE)³⁷,

^{a)}Electronic mail: artur.izmaylov@utoronto.ca

MS-iQCC avoids the bias that arises from sequential optimization of excited states. Instead, our state-averaged formulation treats ground and excited states on an equal footing, analogous to the recently proposed multistate-objective, Ritz-eigenspectral ADAPT-VQE (MORE-ADAPT-VQE)³⁸.

We further describe how the MS-iQCC algorithm generalizes iQCC to multiple states by iteratively selecting generators from an unrestricted operator pool while maintaining orthogonality between evolving state vectors. Finally, while MS-iQCC is designed for efficient classical computation, it naturally produces compact unitary transformations suitable for preparing approximate eigenstates. These transformations can serve as high-quality inputs for future fault-tolerant quantum phase estimation (QPE)^{39–44} algorithms, enabling accurate and scalable excited-state calculations once quantum hardware matures.

The rest of this paper is organized as follows. Section II details the theoretical formulation of MS-iQCC. Section III presents numerical demonstrations for several molecular systems, including H₄, H₂O, N₂, and C₂. Finally, Section IV summarizes the main findings and discusses directions for future development.

II. THEORY

1. The MS-iQCC formalism

Given a Hamiltonian $\hat{H}^{(0)}$ written in terms of qubit operators, and a set of orthogonal reference states $\{|I\rangle\}_{I=1}^{N_s}$, the goal of MS-iQCC is to iteratively construct a parametrized unitary operator $\hat{U}$ such that,

$$\hat{H}_{qcc} = \hat{U} \hat{H}^{(0)} \hat{U}^\dagger, \quad (1)$$

$$\hat{H}_{qcc} |I\rangle = E_I |I\rangle, \quad (2)$$

where $\hat{H}_{qcc}$ is the resulting Hamiltonian from MS-iQCC algorithm, and E_I is the I^{th} lowest eigenenergy of $\hat{H}^{(0)}$. Equivalently, in the Schrödinger picture, $\hat{U}$ maps the reference state $|I\rangle$ to the I^{th} lowest eigenstate of $\hat{H}^{(0)}$. The reference states $|I\rangle$ could potentially be multi-reference in nature, which is necessary when targeting excited multiplet states,²⁸ or excited singlets with open shell character.⁴⁵

$\hat{H}^{(0)}$ is obtained by applying a fermion-to-qubit mapping to the second-quantized molecular Hamiltonian. The MS-iQCC procedure then starts with constructing a density operator out of reference states $\{|I\rangle\}_{I=1}^{N_s}$,

$$\hat{\rho} = \sum_{I=1}^{N_s} w_I |I\rangle\langle I|, \quad (3)$$

where w_I are probabilities. To avoid state-specific bias we set all $w_I = 1/N_s$. The underlying principle behind

MS-iQCC is the state-ensemble variational principle.⁴⁶ At any iteration K , a parametrized unitary transformation is applied to the Hamiltonian from the previous iteration, $\hat{H}^{(K-1)}$, to yield

$$\hat{H}^{(K)}(\tau_K) = \hat{U}_K(\tau_K) \hat{H}^{(K-1)} \hat{U}_K(\tau_K)^\dagger, \quad (4)$$

which gives the state-averaged energy

$$E_{SA}(\tau_K) = \text{Tr}(\hat{H}^{(K)}(\tau_K) \hat{\rho}), \quad (5)$$

where τ_K is a parameter vector. $E_{SA}(\tau_K)$ is then minimized as a function of the parameters to yield the optimal parameter vector τ_K^* and the optimal Hamiltonian at iteration K , $\hat{H}^{(K)} = \hat{H}^{(K)}(\tau_K^*)$. Because of the variational principle, we can be certain that

$$E_{SA}(\tau_K^*) \geq \frac{1}{N_s} \sum_{I=1}^{N_s} E_I \quad (6)$$

which is the average of the true N_s lowest energies of the Hamiltonian $\hat{H}^{(0)}$. Thus by adding a suitably chosen unitary $\hat{U}_K$ at each iteration, we can bring $E_{SA}(\tau_K^*)$ closer to its lower bound. A subtlety unique to MS-iQCC (absent in GS-iQCC) is that, even when equality in Eq. (6) holds, the state-averaged value $\langle I | \hat{H}^{(K)} | I \rangle$ need not necessarily equal E_I . This is a direct consequence of choosing all $w_I = 1/N_s$, since any unitary that acts non trivially only on the reference subspace, leaves the density operator $\hat{\rho}$, and hence the state-averaged energy E_{SA} , unchanged. In order to get the state-specific energy at the end of an iteration, the Hamiltonian $\hat{H}^{(K)}$ needs to be diagonalized within the subspace spanned by the reference states $\{|I\rangle\}_{I=1}^{N_s}$. The algorithm can be terminated when $\Delta E_{SA} = |E_{SA}^{(K)} - E_{SA}^{(K-1)}|$ falls below a threshold. The overview of the algorithm is given in Fig. 1.

One of the main advantages of MS-iQCC is in the choice of unitaries $\hat{U}(\tau_K)$. At each iteration K , one chooses a set of N_g unitaries generated by the basis elements of the $\mathfrak{su}(2^{N_q})$ algebra. These basis elements are the $4^{N_q} - 1$ possible N_q -qubit ‘‘Pauli terms’’ omitting the N_q -qubit identity. Pauli terms are defined as the N_q -fold tensor products of the Pauli operations $\hat{x}$, $\hat{y}$, $\hat{z}$, along with the single qubit identity,

$$\hat{T}_\alpha = \bigotimes_{p=1}^{N_q} \hat{\sigma}_p^{(\alpha)}, \quad \hat{\sigma} \in \{\hat{x}, \hat{y}, \hat{z}, \hat{1}\} \quad (7)$$

where $\hat{\sigma}_p$ is a single qubit operation $\hat{\sigma}$ applied to qubit p . A layer of N_g generators is expressed as

$$\hat{U}_K(\tau_K) = \prod_{\alpha=1}^{N_g} \exp(-i\tau_\alpha^{(K)} \hat{T}_\alpha^{(K)}/2), \quad (8)$$

where $\tau_K = (\tau_1^{(K)}, \tau_2^{(K)}, \dots, \tau_{N_g}^{(K)})$ are N_g real variational parameters referred to as *amplitudes*.

Due to the exponential number of possible generators to include in $\hat{U}_K$, the MS-iQCC algorithm, much like the GS-iQCC, relies on heuristic approaches to find N_g generators which will have the most impact on variationally lowering the energy. We use the magnitude of the energy derivative at $\tau_\alpha = 0$ as an importance measure; the procedure selects N_g generators with the largest g_α given by,

$$g_\alpha = \left| \frac{\partial}{\partial \tau_\alpha} \text{Tr} \left(e^{i\tau_\alpha \hat{T}_\alpha/2} \hat{H}^{(K-1)} e^{-i\tau_\alpha \hat{T}_\alpha/2} \hat{\rho} \right) \right|_{\tau_\alpha=0},$$

$$= \frac{1}{N_s} \left| \sum_{I=1}^{N_s} \text{Im} \left(\langle I | \hat{H}^{(K-1)} \hat{T}_\alpha | I \rangle \right) \right|. \quad (9)$$

Critically, the algorithm does *not* employ a predefined operator pool from which candidate Pauli terms are screened. This fact differentiates the algorithm from other adaptive schemes, which employ operator pools restricted to polynomially scaling sizes to facilitate efficient generator screening. Instead, the effective Hamiltonian $\hat{H}^{(K-1)}$ directly informs us on the structure of Pauli terms $\hat{T}_\alpha^{(K)}$ with non-zero gradient. From Eq. (9), a screening algorithm efficiently checks which $\hat{T}_\alpha^{(K)}$ generate excited configuration $|I_\alpha\rangle \equiv \hat{T}_\alpha^{(K)} |I\rangle$ which can be coupled to $|I\rangle$ through $\hat{H}^{(K-1)}$. This is reminiscent of the selection criterion used in heat-bath CI (HCI).^{47,48} The space generated by the screened Pauli operators $\hat{T}_\alpha$ is called the ‘‘Direct Interaction Space’’ (DIS). Due to its importance, we dedicate Section II 2 to describe its construction, and provide further details in Appendix A.

Another important feature of the algorithm presented here is the ability to classically obtain $\hat{H}^{(K)}$ from $\hat{H}^{(K-1)}$ efficiently, which relies on the self-inverse property of any Pauli term;

$$e^{i\tau_\alpha \hat{T}_\alpha/2} \hat{H} e^{-i\tau_\alpha \hat{T}_\alpha/2} = \hat{H} - \frac{i}{2} \sin(\tau_\alpha) [\hat{H}, \hat{T}_\alpha]$$

$$+ \frac{1}{2} (1 - \cos(\tau_\alpha)) (\hat{T}_\alpha \hat{H} \hat{T}_\alpha - \hat{H}). \quad (10)$$

By repeating this for each of the N_g generators and over K iterations, the number of Pauli terms in the resulting Hamiltonian $\hat{H}^{(K)}$ formally scales as $O((3/2)^{N_g K})$.²² In practice, the proliferation of terms in the Hamiltonian is often drastically below the theoretical scaling.¹⁷ Numerically, the proliferation of terms is mitigated by ‘compressing’ the effective Hamiltonian after each application of Eq. (10), where Pauli terms with coefficient magnitude less than a compression threshold ε_c are pruned.

2. Multi-state direct interaction space

To understand the construction of DIS for MS-iQCC, we start with the simpler case of single determinant ref-

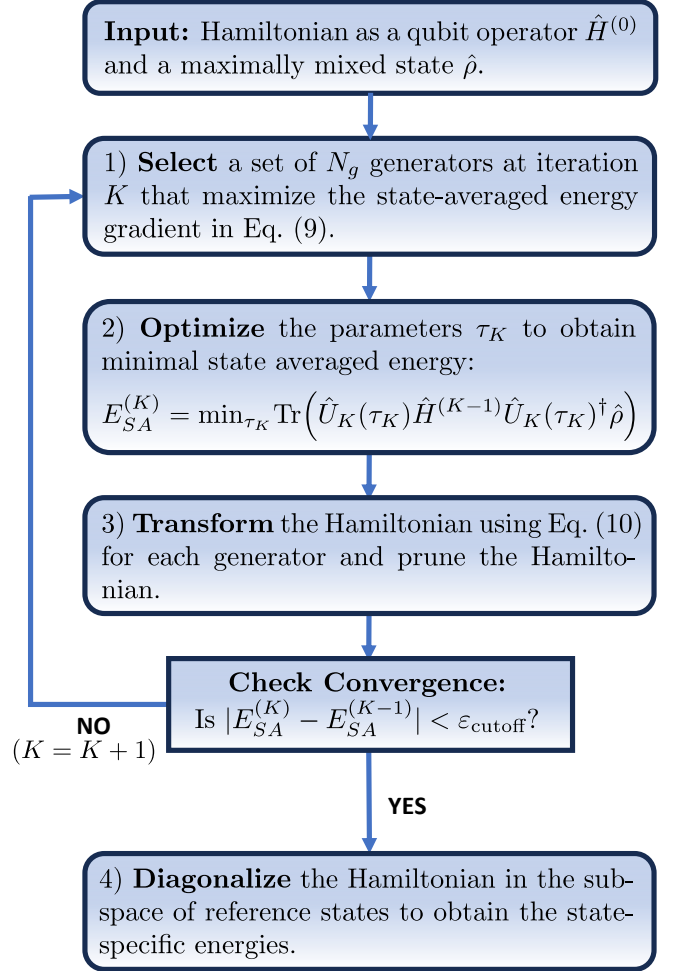


Figure 1. A step-by-step description of the MS-iQCC algorithm. The input mixed state $\hat{\rho}$ is defined in Eq. (3). The diagonalization in step 4 marks the end of the algorithm and the resulting Hamiltonian is the $\hat{H}_{qcc}$ defined in Eq. (1). The diagonalization can also be performed at the end of each iteration to track the convergence of state-specific energies instead of state-averaged energy. If one chooses $N_g = 1$, then step 4 becomes irrelevant and the whole procedure reduces to GS-iQCC.

erence states, $\{|I\rangle\}_{I=1}^{N_s} = \{|\phi_I\rangle\}_{I=1}^{N_s}$. The case of multi-configurational reference states is considered in Appendix A. The ensemble is written as

$$\hat{\rho} = \frac{1}{N_s} \sum_I |\phi_I\rangle\langle\phi_I|.$$

Before maximizing the gradient in Eq. (9), we can eliminate many Pauli words from consideration as they will certainly yield zero gradient. Start by writing the Hamiltonian in Ising factorized form²²,

$$\hat{H} = \sum_i \eta_i \hat{P}_i = \sum_j \hat{D}_j \hat{X}_j, \quad (11)$$

where $\hat{P}_i$ is a Pauli term, and $\hat{D}_j = \sum_k \eta_k^{(j)} \hat{Z}_k^{(j)}$, with $\eta_k^{(j)}$ as real numbers, $\hat{X}$ and $\hat{Z}$ as tensor products of strictly $\hat{x}$ and $\hat{z}$ operations respectively (up to inclusion of identity). For real valued electronic Hamiltonians such as those in the absence of external magnetic field, the Hamiltonian can be equivalently written as $\hat{H} = \sum_j \hat{X}_j \hat{D}_j$. We also express Pauli word $\hat{T}_\alpha$ as $\hat{T}_\alpha = \theta_\alpha \hat{X}_\alpha \hat{Z}_\alpha$, where $\theta_\alpha \in \{1, -1, i, -i\}$. An important point to note here is that θ_α is imaginary if and only if there is an odd number of Pauli $\hat{y}$ operators in the tensor product defining $\hat{T}_\alpha$. This implies that the operators $\hat{x}$ and $\hat{z}$ defining $\hat{X}_\alpha$ and $\hat{Z}_\alpha$ overlap on odd number of qubits. In such case, we say that $\hat{X}_\alpha$ and $\hat{Z}_\alpha$ have *odd overlapping support*.

With this representation of the Hamiltonian and Pauli words, the gradient can be rewritten as,

$$g_\alpha = \frac{1}{N_s} \left| \sum_{I=1}^{N_s} \text{Im} \left(\langle \phi_I | \sum_j \hat{D}_j \hat{X}_j \cdot \theta_\alpha \hat{X}_\alpha \hat{Z}_\alpha | \phi_I \rangle \right) \right|.$$

As $|\phi_I\rangle$ are single Slater determinants, we can reduce this expression to

$$g_\alpha = \frac{|\text{Im}(\theta_\alpha)|}{N_s} \left| \sum_{I=1}^{N_s} \lambda_I^{(\alpha)} \Xi_I^{(\alpha)} \right|, \quad (12)$$

where $\Xi_I^{(\alpha)} = \sum_j D_j^{(I)} \langle \phi_I | \hat{X}_j \hat{X}_\alpha | \phi_I \rangle$ with $D_j^{(I)} = \langle \phi_I | \hat{D}_j | \phi_I \rangle$, and $\lambda_I^{(\alpha)} = \langle \phi_I | \hat{Z}_\alpha | \phi_I \rangle$. Based on Eq. (12), we can identify $\hat{T}_\alpha$'s that do not contribute to the gradient as follows.

- $g_\alpha = 0$ if $\hat{X}_\alpha$ and $\hat{Z}_\alpha$ have even overlapping support, as this would imply $\text{Im}(\theta_\alpha) = 0$.
- $g_\alpha = 0$ if $\hat{X}_\alpha \neq \hat{X}_j$ for any j , as this would imply $\Xi_I^{(\alpha)} = 0$ for all I .

This is the key result underlying the use of an unrestricted pool of unitary generators in MS-iQCC. Procedurally, the DIS is constructed by first choosing $\hat{X}_\alpha = \hat{X}_j$ for each j from Eq. (11), followed by finding $\hat{Z}_\alpha$ such that $\hat{X}_\alpha$ and $\hat{Z}_\alpha$ have odd overlapping support. For these Pauli words, the gradient can be simplified to,

$$g_\alpha = \frac{1}{N_s} \left| \sum_{I=1}^{N_s} \lambda_I^{(\alpha)} D_\alpha^{(I)} \right|. \quad (13)$$

where $\hat{D}_\alpha$ equals one of the $\hat{D}_j$ in Eq. (11), and $\lambda_I^{(\alpha)}$ depends on the choice of $\hat{Z}_\alpha$. We call maximization of Eq. (13) the *phase-alignment problem*: for each fixed $\hat{D}_\alpha$ choose $\hat{Z}_\alpha$ to align the phases $\lambda_I^{(\alpha)} = \pm 1$ so as to maximize g_α . The problem amounts to an N_q -dimensional binary optimization problem and solving for the optimal solution is equivalent to finding the global optimum

to a weighted MAX-SAT problem.⁴⁹ Appendix B details the phase-alignment problem, presenting an optimal approach (OPT) and a greedy approximation (GreedySAT). Note, unlike MS-iQCC, GS-iQCC would have a single term instead of a sum in Eq. (13), which would imply that $\lambda_I^{(\alpha)}$ is an irrelevant global phase, avoiding the problem of phase-alignment entirely. In Appendix A, we extend the DIS construction to state-averages of generally multiconfigurational reference states. We show that this general DIS construction is conceptually similar to the one presented in this section, with gradients retaining the general form of Eq. (12), but with modification of the $\Xi^{(\alpha)}$ components.

III. NUMERICAL RESULTS

In this section we perform simultaneous ground and excited state calculations using the MS-iQCC procedure on a set of modestly sized, albeit strongly correlated, chemical systems. Prior to their fermion-to-qubit encodings, all second quantized Hamiltonians were obtained in the restricted Hartree Fock (RHF) orbital basis. All energy errors reported throughout are calculated via the absolute difference of the state-specific energy and the target state FCI/CASCI energy computed via the PySCF library.⁵⁰ The Hamiltonian growth factors reported throughout refer to the proportionality in number of Pauli terms between an iQCC effective Hamiltonian and the initial qubit-mapped Hamiltonian. The $\hat{H}$ growth factor at iteration K is

$$G_K = \frac{M_K}{M_1}, \quad (14)$$

where M_K and M_1 are the number of Pauli terms in the K^{th} iteration effective and initial Hamiltonians, respectively. We also report the error in target state fidelities (overlaps) defined as:

$$F_I = 1 - \left| \langle E_I | \hat{U}_{\text{iQCC}} | I \rangle \right|^2, \quad (15)$$

computed using the `Openfermion` library⁵¹, to showcase convergence of states along with energies, which could be of independent interest to the community. We use a compression threshold to numerically suppress the proliferation of terms in a transformed iQCC Hamiltonian. Utilizing a threshold of ε_c involves truncating the effective Hamiltonian as

$$\hat{H}_{\varepsilon_c}^{(K)} = \sum_{|\eta_i| \geq \varepsilon_c} \eta_i^{(K)} \hat{P}_i^{(K)}, \quad (16)$$

where such a truncation is performed directly following each step of single-generator dressing.

When describing RHF electronic configurations, we represent them as Fock occupation vectors in the spin orbital basis. The fermionic modes are enumerated by ascending orbital energy, e.g., the HF configuration for an arbitrary system is expressed as $|1, \dots, 1, 0, \dots, 0\rangle$.

A. H_4

Herein, MS-iQCC is applied to multiple state determination for the linear equidistant H_4 chain in the STO-3G basis set. We assess the algorithm in determining the four lowest lying states of the molecule with equidistant H–H separation of twice the equilibrium, $2r_e = 1.9$ Å. The qubit Hamiltonian is obtained through the Jordan-Wigner map, resulting in an $N_q = 8$ qubit Hamiltonian. The MS-iQCC algorithm is employed in determining the four lowest lying eigenstates. A model space of $L = 8$ configurations was employed, consisting of the Fock vectors

$$\begin{aligned} |\phi_1\rangle &= |11110000\rangle, |\phi_2\rangle = |11001100\rangle, |\phi_3\rangle = |11100100\rangle, \\ |\phi_4\rangle &= |11011000\rangle, |\phi_5\rangle = |10110100\rangle, |\phi_6\rangle = |01111000\rangle, \\ |\phi_7\rangle &= |00111100\rangle, |\phi_8\rangle = |11000011\rangle. \end{aligned} \quad (17)$$

The multiconfigurational references are obtained by diagonalizing the 8-dimensional Hamiltonian matrix in the subspace of Eq. (17) and picking the four lowest eigenstates. Energetic errors, target state fidelity errors, and Hamiltonian growth factors for the $N_s = 4$ state determination are presented in Fig. 2. We perform the multi-state determination using $N_g = 1$ generator per iteration, and include the results for two different compression thresholds, ε_c . In Fig. 2, it is seen that an aggressive compression of $\varepsilon_c = 10^{-4}$ causes the state-specific energies to drift away at later iterations instead of converging to their target values. However, the target state fidelity trajectories over iterations are almost identical to those obtained using $\varepsilon_c = 10^{-8}$, with both compression thresholds providing all four state fidelities of > 0.9 by $K \approx 500$. For $\varepsilon_c = 10^{-8}$, all four state-specific energies achieve chemical accuracy after ~ 500 iterations of a single generator at a time. The effective iQCC Hamiltonian grows rapidly until iteration $K \approx 15$, at which point the Hamiltonians reach the so-called “saturation point”,^{22,23} where the set of Pauli terms in the Hamiltonian becomes algebraically closed with respect to commutator with the selected generators. Notably, the effective Hamiltonians using $\varepsilon_c = 10^{-4}$ exhibit a maximal number of terms in vicinity to when $\varepsilon_c = 10^{-8}$ achieves saturation, however the number of terms begins to slightly decrease past this point. The rapid early saturation, and near-saturation of the aggressively compressed Hamiltonians, can be attributed to the low Hilbert space dimensionality. The scaling of effective Hamiltonians in the MS-iQCC procedure along with the effect of compression is revisited for the larger problems of N_2 and C_2 in Sections III C and III D respectively.

B. H_2O

Herein the MS-iQCC procedure is applied to a CAS(4e, 4o) model of the stretched H_2O molecule in the 6-31G

basis set. A distance of 2.35 Å was used for both O–H bonds, with a H–O–H bond angle of 107.6°. At such a geometry, the two lowest lying singlet states, S_0 and S_1 , exhibit a high degree of multiconfigurational character, making their simultaneous estimation a challenging problem. In order to ensure the S_1 estimate does not converge to a lower energy high-spin solution, such as T_1 , a spin-penalized Hamiltonian

$$\hat{H}_s = \hat{H} + \mu \hat{S}^2 \quad (18)$$

is utilized for generator screening and amplitude optimization, where $\mu = 0.25$ a.u. We assess the performance of the MS-iQCC procedure utilizing only $N_g = 1$, i.e., a single generator per iteration, with varying reference states and phase-alignment procedures utilized. The MS-iQCC is applied to the S_0 and S_1 estimation using a model space comprised of a total of $L = 4$ configurations,

$$\begin{aligned} |\phi_1\rangle &= |11110000\rangle, |\phi_2\rangle = |11001100\rangle, \\ |\phi_3\rangle &= |10110100\rangle, |\phi_4\rangle = |01111000\rangle. \end{aligned} \quad (19)$$

The multiconfigurational references are obtained by diagonalizing the 4-dimensional Hamiltonian matrix in the subspace of Eq. (19) and picking the two lowest eigenstates. In Fig. 3, we compare the two phase-alignment procedures (see Appendix B), the optimal OPT with the heuristic **GreedySAT** approach. We observe that the OPT phase-alignment yields accelerated convergence of the state-specific energies and state-averaged energy compared to the **GreedySAT** phase-alignment. Convergence to chemical accuracy for both singlet states was accomplished using **GreedySAT** with roughly 50 additional iterations. The slower convergence can be attributed to the heuristic nature of **GreedySAT**, which may potentially miss the true maximal gradient generators.

Notably, both phase-alignment methods achieve > 0.9 state fidelities for S_0 and S_1 with $K \sim 100$, despite the initial S_1 overlap being merely 0.19. To assess the benefit of increasing the model space, we perform an identical MS-iQCC calculation in Fig. 4, except with an additional 2 electronic configurations included in the S_1 reference state, given by

$$|\phi_5\rangle = |11100001\rangle, |\phi_6\rangle = |11010010\rangle. \quad (20)$$

Diagonalization in the 6-dimensional model space yields a multiconfigurational reference for the S_1 state with an improved initial target fidelity of 0.39. It is seen that the MS-iQCC algorithm using either of the phase-alignment techniques requires ~ 50 fewer iterations to achieve state-specific energies within chemical accuracy, yielding $\sim 33\%$ and $\sim 25\%$ reductions for the OPT and **GreedySAT** phase-alignments respectively.

C. N_2

We assess the MS-iQCC procedure in performing simultaneous estimation of the ground (S_0) and first excited (T_1) states of N_2 in a CAS(6e, 6o) active space

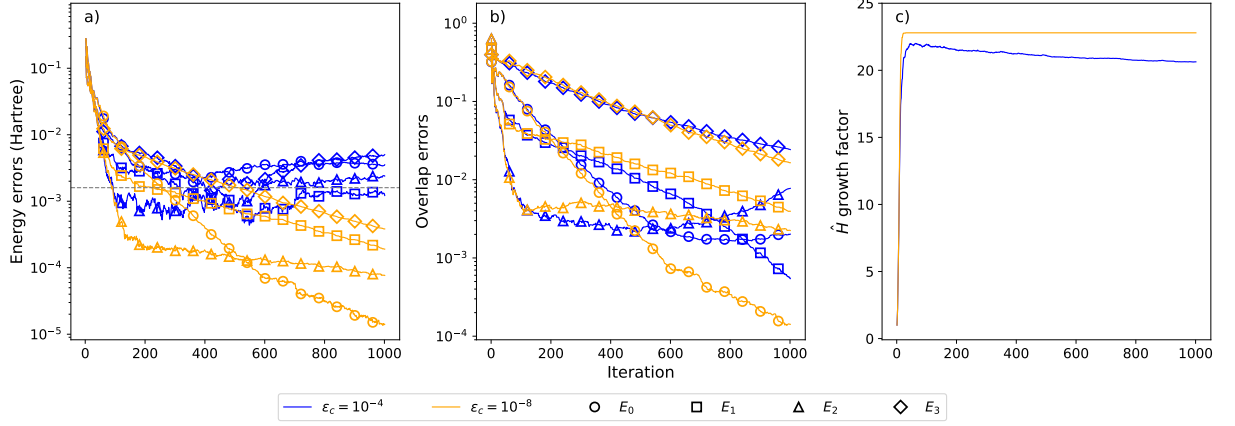


Figure 2. MS-iQCC applied to the determination of the four lowest energy eigenstates for the H_4 chain at an equidistant separation of $2r_e$ in the STO-3G basis set. The ϵ_c denotes the compression threshold used. a) Errors in the state-specific energies relative to the associated FCI target energies. b) Errors in the squared wavefunction overlaps of the iQCC-rotated trial states with their corresponding target eigenstates (see Eq. (15)). c) The ratio of number of terms in the effective iQCC Hamiltonian and the initial qubit-mapped Hamiltonian.

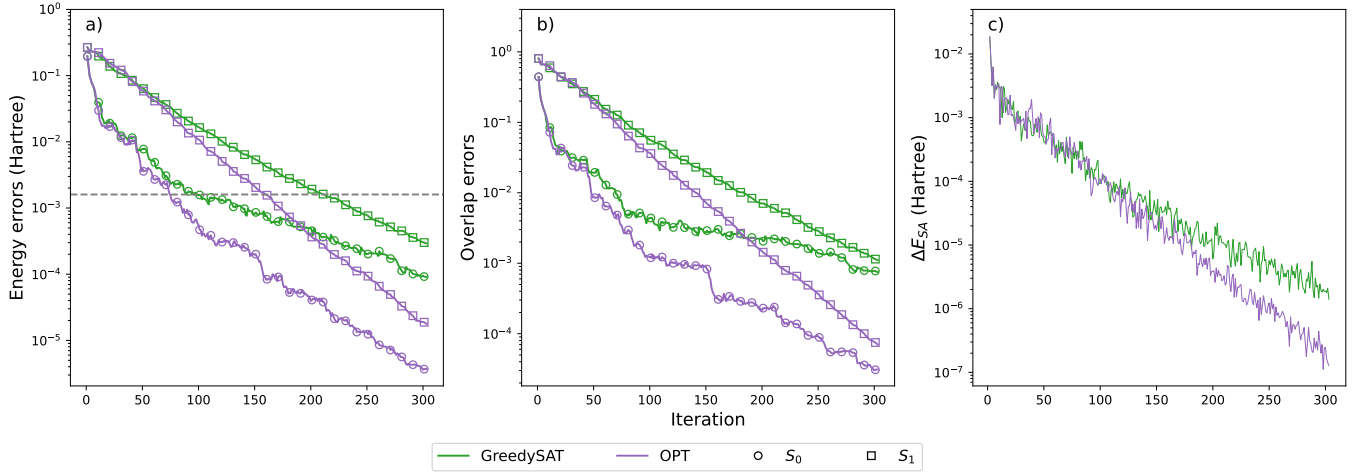


Figure 3. The MS-iQCC procedure with $N_g = 1$ applied to the simultaneous determination of the S_0 and S_1 states of the CAS(4e, 4o) model of stretched H_2O in the 6-31G basis set. The $L = 4$ configurations in Eq. (19) were used to define the model space. The global optimization described in Appendix B 1 was utilized for phase-alignment. a) The errors of the MS-iQCC S_0 and S_1 estimates taken with respect to the exact S_0 and S_1 solutions within the CAS(4e, 4o) space. b) Errors in the squared overlaps of the MS-iQCC trial states with respect to the associated target state. c) ΔE_{SA} at iteration K is given by the difference between the optimized K^{th} and $(K - 1)^{\text{th}}$ MS-iQCC state-averaged energy.

in the STO-6G basis set. We benchmark the algorithm at both equilibrium ($r_e = 1.0975 \text{ \AA}$) and extended ($2r_e$) bond distances. The qubit Hamiltonian is obtained under the parity mapping,⁵² and a two qubit tapering procedure⁵³ is employed to obtain an $N_q = 10$ qubit Hamiltonian. To obtain the multiconfigurational reference states for S_0 and T_1 , we utilize a small set of $L = 7$ configurations from the singles-and-doubles manifold from Hartree-Fock. For the calculation performed at equilibrium geometry, we utilize electronic configurations

with the following representation as Fock vectors

$$\begin{aligned}
 |\phi_1\rangle &= |111111000000\rangle, |\phi_2\rangle = |001111110000\rangle, \\
 |\phi_3\rangle &= |110011001100\rangle, |\phi_4\rangle = |101111010000\rangle, \\
 |\phi_5\rangle &= |011111110000\rangle, |\phi_6\rangle = |111011000100\rangle, \\
 |\phi_7\rangle &= |110111001000\rangle.
 \end{aligned} \tag{21}$$

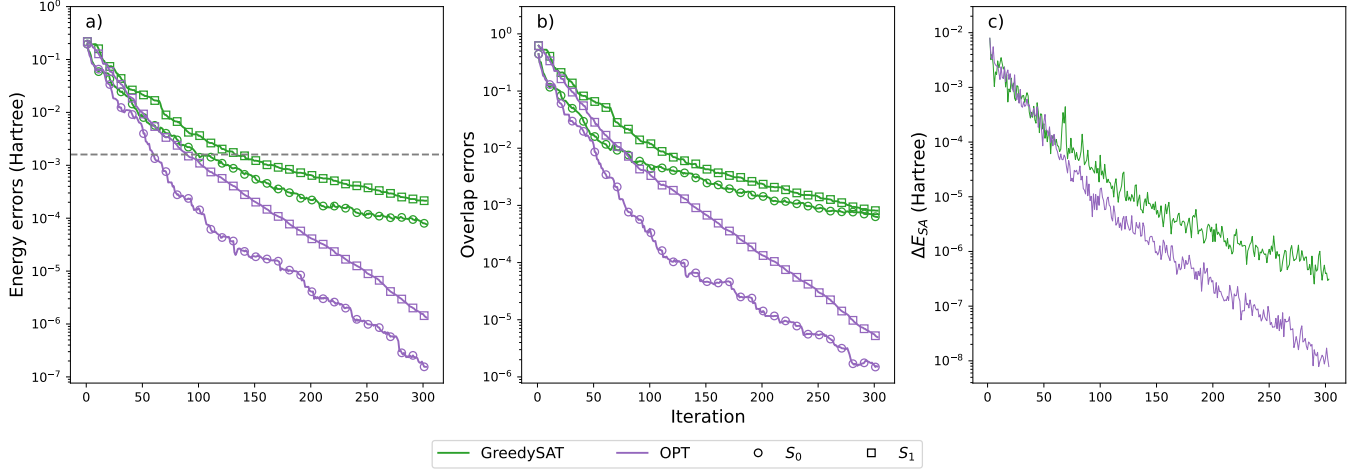


Figure 4. The same as Fig. 3, however employing the slightly larger model space of dimension $K = 6$, with extra configurations contributing to the S_1 trial state reference given in Eq. (20).

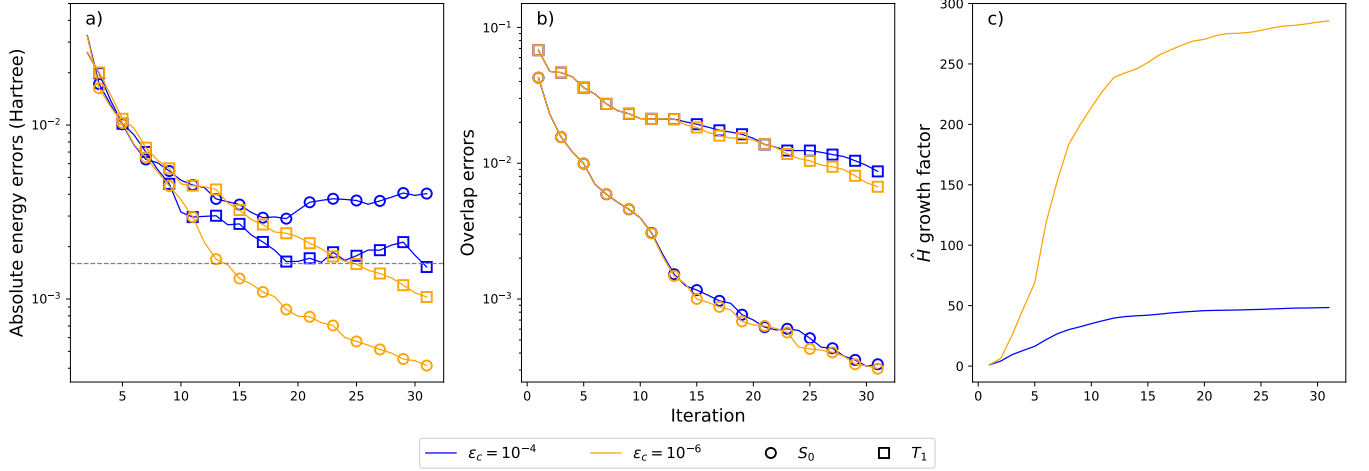


Figure 5. The MS-iQCC procedure applied to simultaneous determination of the S_0 and T_1 states of the CAS(6e, 6o) model of N_2 molecule at $r_e = 1.0975$ Å bond distance in the STO-6G basis set. A total of $L = 7$ computational basis states were used, with $N_g = 5$ generators used at each iteration, and OPT strategy utilized for phase-alignment. The subplots a, b, and c correspond to energy errors, fidelity errors, and the growth factors, respectively.

At the extended bond length of $2r_e$, the model space used is

$$\begin{aligned}
 |\phi_1\rangle &= |111111000000\rangle, |\phi_2\rangle = |110011110000\rangle, \\
 |\phi_3\rangle &= |111100001100\rangle, |\phi_4\rangle = |111011010000\rangle, \\
 |\phi_5\rangle &= |110111100000\rangle, |\phi_6\rangle = |111110000100\rangle, \\
 |\phi_7\rangle &= |111101001000\rangle.
 \end{aligned} \quad (22)$$

The choice of model space configurations was made by observing the highest weighted determinants in the S_0 and T_1 CISD solutions at both geometries, with $|\phi_i\rangle$ for $i \in \{1, 2, 3\}$ being the top three contributing configurations to the S_0 solution, and the remaining four configurations coming from the T_1 solution. The Hamiltonian was diagonalized in the subspace of Eq. (21) and Eq. (22) and the two lowest lying solutions

were taken as the multiconfigurational reference states for S_0 and T_1 states. Figure 5 provides numerics for the MS-iQCC procedure applied to the two state determination at equilibrium geometry, using $N_g = 5$ generators each iteration, for two different regimes of compression. Similarly to the situation seen for H_4 , crudely compressing to a threshold of 10^{-4} is seen to provide unreliable convergence of state-specific energies, albeit providing target state fidelities on par with a compression of 10^{-6} . The latter less severe compression is seen to converge both S_0 and T_1 trial energies to within chemical accuracy at $K = 25$. Notably, the number of terms in the iQCC effective Hamiltonian is seen to be much smaller for $\varepsilon_c = 10^{-4}$ compared to $\varepsilon_c = 10^{-6}$. Such behaviour can be explained by the estimation at equilibrium geometry being dominated by

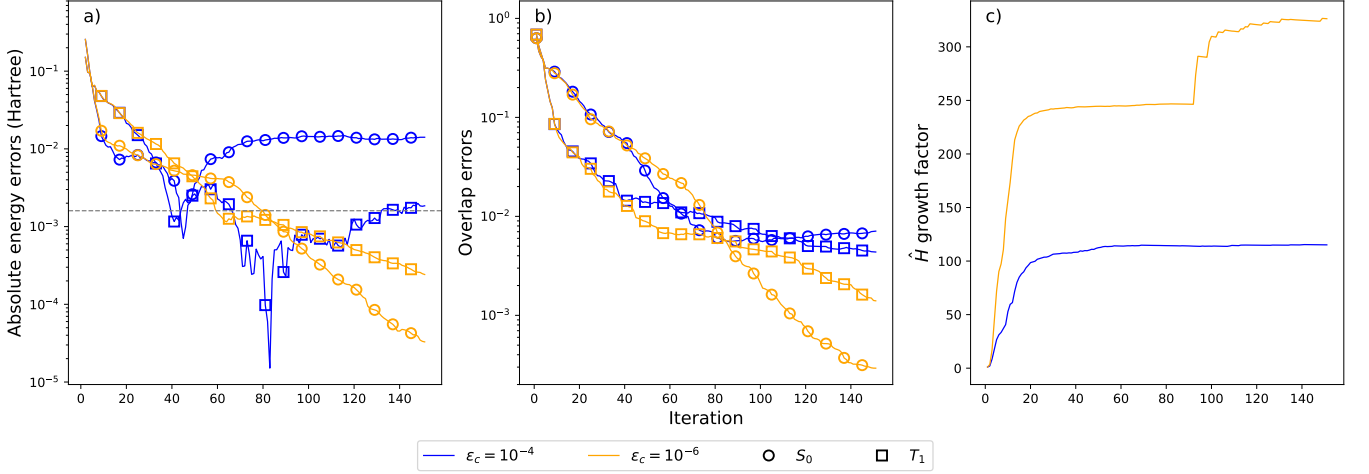


Figure 6. The same as Fig. 5 but performed at a N–N bond distance of $2r_e$.

dynamical/weak correlation, evident by the high overlaps of the starting references. In the non-strongly correlated regime, the optimized iQCC amplitudes τ_α are generally small. From Eq. (10), new terms entering the updated Hamiltonian carry coefficient $\sin(\tau_\alpha)$. As the Hamiltonian is iteratively dressed using small amplitudes, terms are being continually suppressed by $\sin(\tau_\alpha) \ll 1$ factors, leading to many terms being pruned during compression.

At the bond distance of $2r_e$, the reference states possess squared overlaps of ~ 0.37 and ~ 0.31 with the exact S_0 and T_1 states, respectively. The low overlaps of the reference states with their target states present this example as an interesting and challenging problem for simultaneous state estimation. Furthermore, following T_1 , the next low-lying excited states are quintet and septet states. Since iQCC generators do not conserve spin symmetries, energetically quasidegenerate high spin states can be problematic. If the algorithm finds generators which maximally lower energy at early iterations, yet introduce large amounts of spin contamination, there is a risk of convergence to one such high spin eigenstate. A less severe scenario, yet still unfavorable, is that many iQCC iterations are required to approximately restore the desired spin quantum numbers to that of the lower-lying target state. Such behaviour is a clear manifestation of Löwdin’s symmetry dilemma.⁵⁴ To avoid large amounts of spin contamination entering the trial states at early iterations, we again employ the spin-penalized Hamiltonian of Eq. (18) for operator screening and optimization, with a smaller penalty $\mu = 0.025$ a.u., to avoid overly penalizing triplet energies.

In Fig. 6, numerics are provided for the S_0 and T_1 determinations at $2r_e$ bond length. Utilizing $\epsilon_c = 10^{-6}$ is seen to provide systematic convergence towards the state-specific S_0 and T_1 energies, with chemical accuracy achieved by $K = 80$ iterations. The poor fidelities of the reference states are rapidly improved, with both

achieving ~ 0.9 squared overlap with their target states by $K = 25$. Using $\epsilon_c = 10^{-4}$ resulted in both the S_0 and T_1 energy estimates becoming non-variational, resulting in the sharp negative peaks in the trajectories of their absolute energy errors. The breaking of the variational behaviour is possible when compressing the effective Hamiltonian to low precision. Aggressively pruning terms of high coefficient magnitude can lead to significant spectral perturbations, and hence should be avoided when accurate energies are desired as the output of the MS-iQCC algorithm.

D. C_2

We demonstrate the usage of MS-iQCC procedure for simultaneous determination of S_0 state and two degenerate S_1 states of C_2 in the complete active space of 6 electrons and 6 orbitals of the cc-pVDZ basis, at a bond length of $r_e = 1.2 \text{ \AA}$. Here as well, the qubit Hamiltonian is obtained under the parity mapping, and a two qubit tapering procedure is employed to obtain an $N_q = 10$ qubit Hamiltonian. All target states have spin projection along the z-axis equal to zero. We choose the reference determinants by looking at the top contributors to the target states in the CISD space. They are given by,

$$\begin{aligned} |\phi_1\rangle &= |111111000000\rangle, & |\phi_2\rangle &= |001111110000\rangle, \\ |\phi_3\rangle &= |111101100000\rangle, & |\phi_4\rangle &= |111110010000\rangle, \\ |\phi_5\rangle &= |111011010000\rangle, & |\phi_6\rangle &= |110111100000\rangle. \end{aligned}$$

We diagonalize the spin penalized Hamiltonian of Eq. (18) with $\mu = 0.025$ in the subspace spanned by these determinants and pick the three lowest lying states as the multiconfigurational references. We run the algorithm for 150 iterations with number of generators in each iteration set to 6. GreedySAT algorithm is utilized for solving the phase-alignment problem.

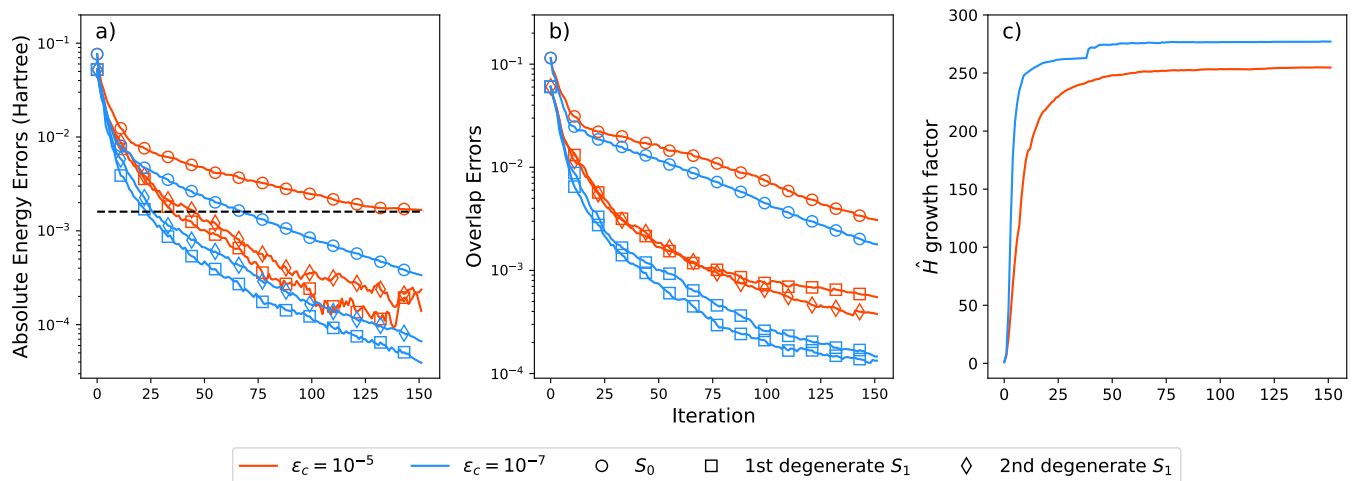


Figure 7. The MS-iQCC procedure applied to simultaneous determination of S_0 and the two degenerate S_1 states of the CAS(6e, 6o) model of C_2 molecule at $r_e = 1.2$ Å bond distance in the cc-pVDZ basis. Each iteration utilized $N_g = 6$ generators, and GreedySAT strategy was utilized for phase-alignment. Two different compression factors were tested. The subplots a, b, and c correspond to energy errors, fidelity errors, and growth factors, respectively.

We study the results under two different compression thresholds. In Fig. 7 a), we see that for $\epsilon_c = 10^{-5}$, convergence of S_0 state to chemical accuracy is extremely slow, taking nearly 150 iterations. For S_1 states, chemical accuracy is reached within 50 iterations, although the convergence is substantially perturbed towards the end. This could be attributed to the choice of large compression threshold which might prune many important Pauli words off the Hamiltonian, breaking the variational nature of the problem. To recover a more systematic convergence for all states, we rerun the calculation with a smaller compression threshold of $\epsilon_c = 10^{-7}$. We see that all three states converge systematically to the target energies and the chemical accuracy is reached around 25 iterations for excited states, and 75 iterations for the ground state. This suggests that the ground state is possibly more strongly correlated than the excited states, requiring more Slater determinants in its expression and hence more number of iQCC generators.

Unlike in the case of N_2 at equilibrium bond length (see Fig. 5), we can see in Fig. 7 b) that reducing the compression threshold improves the target state fidelities. In subplot c) Hamiltonian growth factor, as expected, displays a rise in number of terms as ϵ_c is lowered. But compared to other molecules, we can see that the rise is less significant. The smaller compression threshold leads to a drastic rise in the growth factor at the beginning as most terms generated by unitaries will be retained, and the Hamiltonian terms become algebraically closed with respect to the dominant generators fairly quickly. In case of a more aggressive pruning, many iterations are needed to boost the coefficients of the important Pauli words to reach algebraic closure.

IV. CONCLUSIONS AND OUTLOOK

In this work, we have introduced the MS-iQCC algorithm for simultaneous and unbiased determination of ground and excited state energies of qubit mapped Hamiltonians. We tested the algorithm on several strongly correlated molecules and found that, under moderate Hamiltonian compression thresholds, it demonstrated robust convergence of all state-specific energies to within chemical accuracy. In this regime, MS-iQCC functions as a fully classical multi-state solver. We have also shown that one can use spin penalized Hamiltonians to target excited states in different symmetry sectors. The efficacy of the MS-iQCC algorithm can be largely attributed to the unrestricted use of the full $\mathfrak{su}(2^{N_q})$ algebra acting on the N_q -qubit Hilbert space in selecting unitary generators. This results in an adaptively growing *Hamiltonian*, in analogy to the use of an adaptively growing configuration space as used in iterative selected CI techniques.^{47,48} Interestingly, for all the cases studied, the number of terms in the effective iQCC Hamiltonian converges rapidly, showing that the Pauli words in the Hamiltonian become algebraically closed with respect to the dominant generators being used. This property supports the observation that the growth in the number of terms of the iQCC Hamiltonian is substantially below the theoretical scaling²². We also showed that along with state specific energies, MS-iQCC displays a systematic convergence of target state fidelities. This could be of independent interest to the quantum computing community where state preparation is an important subroutine in extracting spectral properties using the Quantum Phase Estimation algorithm. Due to the compact nature of MS-iQCC unitaries and the ability to generate multiple excited

states, it is worth exploring the benefits and limitations of using MS-iQCC as a quantum state preparation algorithm.

MS-iQCC has tunable parameters, such as the compression threshold and the number of generators per iteration, which control the trade-off between cost and accuracy. More consideration is needed in choosing their values, since their rigorous connection with target accuracy is not apparent. Another important component of the algorithm is the solution to phase-alignment problem. Since finding an optimal solution is prohibitively expensive for large systems, quality of greedy solutions strongly influences the number of iQCC iterations needed to reach a target accuracy. We believe further research is needed in improving these approximate solutions, or modifying steps in MS-iQCC that could suggest better strategies for tackling the phase-alignment problem.

V. ACKNOWLEDGEMENTS

We thank Junan Lin and Christian Schilling for stimulating discussions. A.F.I. acknowledges financial support from the Natural Sciences and Engineering Council of Canada (NSERC). This research was partly enabled by the support of Compute Ontario (computeontario.ca) and the Digital Research Alliance of Canada (alliance-can.ca). Part of the computations were performed on the Niagara and Trillium supercomputers at the SciNet HPC Consortium, and the NARVAL and RORQUAL supercomputers under the Calcul Quebec Consortium. SciNet is funded by Innovation, Science, and Economic Development Canada, the Digital Research Alliance of Canada, the Ontario Research Fund: Research Excellence, and the University of Toronto.

REFERENCES

- ¹W. Domcke, D. R. Yarkony, and H. Köppel, “Conical Intersections,” *Advanced Series in Physical Chemistry* **17** (2011).
- ²D. R. Yarkony, “Nonadiabatic quantum chemistry—past, present, and future,” *Chem. Rev.* **112**, 481–498 (2012).
- ³B. O. Roos, P. R. Taylor, and P. E. Sigbahn, “A complete active space scf method (casscf) using a density matrix formulated super-ci approach,” *Chem. Phys.* **48**, 157–173 (1980).
- ⁴J. Olsen, B. O. Roos, P. Jørgensen, and H. J. A. Jensen, “Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces,” *J. Chem. Phys.* **89**, 2185–2192 (1988).
- ⁵P. A. Malmqvist, A. Rendell, and B. O. Roos, “The restricted active space self-consistent-field method, implemented with a split graph unitary group approach,” *J. Phys. Chem.* **94**, 5477–5482 (1990).
- ⁶G. Li Manni, F. Aquilante, and L. Gagliardi, “Strong correlation treated via effective hamiltonians and perturbation theory,” *J. Chem. Phys.* **134**, 034114 (2011).
- ⁷K. Andersson, P. A. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, “Second-order perturbation theory with a casscf reference function,” *J. Phys. Chem.* **94**, 5483–5488 (1990).
- ⁸K. Andersson, P. Malmqvist, and B. O. Roos, “Second-order perturbation theory with a complete active space self-consistent field reference function,” *J. Chem. Phys.* **96**, 1218–1226 (1992).
- ⁹B. O. Roos, K. Andersson, M. P. Fülscher, P.-å. Malmqvist, L. Serrano-Andrés, K. Pierloot, and M. Merchán, “Multiconfigurational perturbation theory: Applications in electronic spectroscopy,” *Advances in chemical physics: new methods in computational quantum mechanics* **93**, 219–331 (1996).
- ¹⁰G. K.-L. Chan and S. Sharma, “The density matrix renormalization group in quantum chemistry,” *Annu Rev Phys Chem* **62**, 465–481 (2011).
- ¹¹M. Chandross and J. C. Hicks, “Density-matrix renormalization-group method for excited states,” *Phys. Rev. B* **59**, 9699–9702 (1999).
- ¹²W. Hu and G. K.-L. Chan, “Excited-state geometry optimization with the density matrix renormalization group, as applied to polyenes,” *J. Chem. Theory Comput.* **11**, 3000–3009 (2015).
- ¹³L. Freitag and M. Reiher, “The density matrix renormalization group for strong correlation in ground and excited states,” in *Quantum Chemistry and Dynamics of Excited States* (John Wiley & Sons, Ltd, 2020) Chap. 7, pp. 205–245.
- ¹⁴W. Barford, R. J. Bursill, and M. Y. Lavrentiev, “Density-matrix renormalization-group calculations of excited states of linear polyenes,” *Phys. Rev. B* **63**, 195108 (2001).
- ¹⁵R. Sarkar, M. Casanova-Páez, F. Neese, and R. Izsák, “Benchmarking td-dft and wave function methods for excitation energies,” *Journal of Chemical Theory and Computation* **17**, 6612–6629 (2021).
- ¹⁶T. Froitzheim, L. Kunze, S. Grimme, J. M. Herbert, and J.-M. Mewes, “Benchmarking charge-transfer excited states in tadm emitters: Δ dft outperforms td-dft for emission energies,” *Journal of Physical Chemistry A* **128**, 6324–6335 (2024).
- ¹⁷S. N. Genin, I. G. Ryabinkin, N. R. Paisley, S. O. Whelan, M. G. Helander, and Z. M. Hudson, “Estimating phosphorescent emission energies in iriii complexes using large-scale quantum computing simulations**,” *Angewandte Chemie International Edition* **61**, e202116175 (2022).
- ¹⁸A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O’Brien, “A variational eigenvalue solver on a photonic quantum processor,” *Nature Communications* **5**, 4213 (2014).
- ¹⁹J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, “The theory of variational hybrid quantum-classical algorithms,” *New Journal of Physics* **18**, 023023 (2016).
- ²⁰M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, “Variational quantum algorithms,” *Nature Reviews Physics* **3**, 625–644 (2021).
- ²¹J. Tilly, H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. H. Booth, and J. Tennyson, “The variational quantum eigensolver: A review of methods and best practices,” *Physics Reports* **986**, 1–128 (2022), the Variational Quantum Eigensolver: a review of methods and best practices.
- ²²I. G. Ryabinkin, R. A. Lang, S. N. Genin, and A. F. Izmaylov, “Iterative qubit coupled cluster approach with efficient screening of generators,” *J. Chem. Theory Comput.* **16**, 1055–1063 (2020).
- ²³R. A. Lang, I. G. Ryabinkin, and A. F. Izmaylov, “Unitary transformation of the electronic Hamiltonian with an exact quadratic truncation of the Baker-Campbell-Hausdorff expansion,” *J. Chem. Theory Comput.* **17**, 66–78 (2021).
- ²⁴A. Anand, P. Schleich, S. Alperin-Lea, P. Jensen, S. Sim, M. Díaz-Tinoco, J. Kottmann, M. Degroote, A. Izmaylov, and A. Aspuru-Guzik, “A quantum computing view on unitary coupled cluster theory,” *Chem. Soc. Rev.* **51**, 1659–1684 (2022).
- ²⁵S. Patel, P. Jayakumar, T.-C. Yen, and A. F. Izmaylov, “Quantum measurement for quantum chemistry on a quantum computer,” *Chemical Reviews* **125**, 7490–7524 (2025).
- ²⁶S. Barison, J. Robledo Moreno, and M. Motta, “Quantum-centric computation of molecular excited states with extended sample-based quantum diagonalization,” *Quantum Science and Technol-*

- ogy **10**, 025034 (2025).
- ²⁷S. Piccinelli, A. Baiardi, M. Rossmannek, A. C. Vazquez, F. Tacchino, S. Mensa, E. Altamura, A. Alavi, M. Motta, J. Robledo-Moreno, W. Kirby, K. Sharma, A. Mezzacapo, and I. Tavernelli, “Quantum chemistry with provable convergence via randomized sample-based quantum diagonalization,” (2025), arXiv:2508.02578 [quant-ph].
 - ²⁸T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (John Wiley & Sons, Ltd, Chichester, UK, 2000).
 - ²⁹H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, “An adaptive variational algorithm for exact molecular simulations on a quantum computer,” *Nat. Commun.* **10**, 3007 (2019).
 - ³⁰H. L. Tang, V. Shkolnikov, G. S. Barron, H. R. Grimsley, N. J. Mayhall, E. Barnes, and S. E. Economou, “Qubit-adapt-vqe: An adaptive algorithm for constructing hardware-efficient ansätze on a quantum processor,” *PRX Quantum* **2**, 020310 (2021).
 - ³¹B. Huron, J.-P. Malrieu, and P. Rancurel, “Configuration interaction by perturbatively selected iterative method,” *The Journal of Chemical Physics* **58**, 5745–5759 (1973).
 - ³²A. A. Holmes, N. M. Tubman, and C. J. Umrigar, “An efficient selected configuration interaction plus perturbation theory algorithm,” *Journal of Chemical Theory and Computation* **12**, 3674–3680 (2016).
 - ³³F. A. Evangelista, “Adaptive configuration interaction: A multi-configurational wave function method inspired by machine learning,” *The Journal of Chemical Physics* **140**, 124114 (2014).
 - ³⁴S. Sharma, A. A. Holmes, G. Jeanmairet, A. Alavi, and C. J. Umrigar, “Semistochastic heat-bath configuration interaction method: Selected configuration interaction with perturbation theory using a semistochastic approach,” *Journal of Chemical Theory and Computation* **13**, 1595–1604 (2017).
 - ³⁵O. Higgott, D. Wang, and S. Brierley, “Variational Quantum Computation of Excited States,” *Quantum* **3**, 156 (2019).
 - ³⁶J. Wen, D. Lv, M.-H. Yung, and G.-L. Long, “Variational quantum packaged deflation for arbitrary excited states,” *Quantum Eng.* **3**, e80 (2021).
 - ³⁷Q.-X. Xie, S. Liu, and Y. Zhao, “Orthogonal state reduction variational eigensolver for the excited-state calculations on quantum computers,” *Journal of Chemical Theory and Computation* **18**, 3737–3746 (2022).
 - ³⁸H. R. Grimsley and F. A. Evangelista, “Challenging excited states from adaptive quantum eigensolvers: subspace expansions vs. state-averaged strategies,” *Quantum Science and Technology* **10**, 025003 (2025).
 - ³⁹A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, “Simulated quantum computation of molecular energies,” *Science* **309**, 1704–1707 (2005).
 - ⁴⁰R. B. Griffiths and C.-S. Niu, “Semiclassical fourier transform for quantum computation,” *Phys. Rev. Lett.* **76**, 3228–3231 (1996).
 - ⁴¹L. Lin and Y. Tong, “Heisenberg-limited ground-state energy estimation for early fault-tolerant quantum computers,” *PRX Quantum* **3**, 010318 (2022).
 - ⁴²G. Wang, D. S. França, R. Zhang, S. Zhu, and P. D. Johnson, “Quantum algorithm for ground state energy estimation using circuit depth with exponentially improved dependence on precision,” *Quantum* **7**, 1167 (2023).
 - ⁴³Z. Ding and L. Lin, “Even shorter quantum circuit for phase estimation on early fault-tolerant quantum computers with applications to ground-state energy estimation,” *PRX Quantum* **4**, 020331 (2023).
 - ⁴⁴K. Gratsea, J. S. Kottmann, P. D. Johnson, and A. A. Kunitsa, “Comparing Classical and Quantum Ground State Preparation Heuristics,” arXiv e-prints, arXiv:2401.05306 (2024), arXiv:2401.05306 [quant-ph].
 - ⁴⁵A. Balková and R. J. Bartlett, “Coupled-cluster method for open-shell singlet states,” *Chem. Phys. Lett.* **193**, 364–372 (1992).
 - ⁴⁶E. K. U. Gross, L. N. Oliveira, and W. Kohn, “Rayleigh-ritz variational principle for ensembles of fractionally occupied states,” *Phys. Rev. A* **37**, 2805–2808 (1988).
 - ⁴⁷A. A. Holmes, N. M. Tubman, and C. J. Umrigar, “Heat-bath configuration interaction: An efficient selected configuration interaction algorithm inspired by heat-bath sampling,” *J. Chem. Theory Comput.* **12**, 3674–3680 (2016).
 - ⁴⁸S. Sharma, A. A. Holmes, G. Jeanmairet, A. Alavi, and C. J. Umrigar, “Semistochastic heat-bath configuration interaction method: Selected configuration interaction with semistochastic perturbation theory,” *J. Chem. Theory Comput.* **13**, 1595–1604 (2017).
 - ⁴⁹P. Mills and E. Tsang, “Guided local search for solving SAT and weighted MAX-SAT problems,” *Journal of Automated Reasoning* **24**, 205–223 (2000).
 - ⁵⁰Q. Sun, X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui, J. J. Eriksen, Y. Gao, S. Guo, J. Hermann, M. R. Hermes, K. Koh, P. Koval, S. Lehtola, Z. Li, J. Liu, N. Mardirossian, J. D. McClain, M. Motta, B. Mussard, H. Q. Pham, A. Pulkin, W. Purwanto, P. J. Robinson, E. Ronca, E. R. Sayfutyarova, M. Scheurer, H. F. Schurkus, J. E. T. Smith, C. Sun, S.-N. Sun, S. Upadhyay, L. K. Wagner, X. Wang, A. White, J. D. Whitfield, M. J. Williamson, S. Wouters, J. Yang, J. M. Yu, T. Zhu, T. C. Berkelbach, S. Sharma, A. Y. Sokolov, and G. K.-L. Chan, “Recent developments in the PySCF program package,” *J. Chem. Phys.* **153**, 024109 (2020).
 - ⁵¹J. R. McClean, N. C. Rubin, K. J. Sung, I. D. Kivlichan, X. Bonet-Monroig, Y. Cao, C. Dai, E. S. Fried, C. Gidney, B. Gimby, P. Gokhale, T. Häner, T. Hardikar, V. Havlíček, O. Higgott, C. Huang, J. Izaac, Z. Jiang, X. Liu, S. McArdle, M. Neeley, T. O’Brien, B. O’Gorman, I. Ozfidan, M. D. Radin, J. Romero, N. P. D. Sawaya, B. Senjean, K. Setia, S. Sim, D. S. Steiger, M. Steudtner, Q. Sun, W. Sun, D. Wang, F. Zhang, and R. Babbush, “Openfermion: the electronic structure package for quantum computers,” *Quantum Sci. Technol.* **5**, 034014 (2020).
 - ⁵²S. B. Bravyi and A. Y. Kitaev, “Fermionic quantum computation,” *Annals of Physics* **298**, 210–226 (2002).
 - ⁵³S. Bravyi, J. M. Gambetta, A. Mezzacapo, and K. Temme, “Tapering off qubits to simulate fermionic hamiltonians,” arXiv:1701.08213 (2017), 1701.08213.
 - ⁵⁴P. Lykos and G. W. Pratt, “Discussion on the Hartree-Fock approximation,” *Rev. Mod. Phys.* **35**, 496–501 (1963).

Appendix A: DIS construction for multiconfigurational reference states

Herein, the DIS construction for an ensemble of multiconfigurational reference states is provided. In Section II 2, it is shown that the ensemble-averaged energy gradient for candidate generator $\hat{T}_\alpha$ has the form Eq. (12) for ensembles over determinantal references, i.e., $\hat{\rho} = \sum_i |\phi_i\rangle \langle \phi_i| / N_s$. In the multiconfigurational case, we utilize the more general ensemble $\rho = \sum_I |I\rangle \langle I| / N_s$, with

$$|I\rangle = \sum_i c_i^{(I)} |\phi_i\rangle. \quad (\text{A1})$$

It is shown here that we can obtain a form of the multiconfigurational state-averaged gradient:

$$g_\alpha = \frac{1}{N_s} \left| \sum_{I=1}^{N_s} \text{Im} \left(\langle I | \hat{H} \hat{T}_\alpha | I \rangle \right) \right| \quad (\text{A2})$$

which retains the general form of Eq. (12), but with modifications to $\Xi^{(\alpha)}$ components. Firstly, let $\hat{\Omega}_i^{(I)}$ be the *wave operator* which achieves $\hat{\Omega}_i^{(I)} |\phi_i\rangle = |I\rangle$. From Eq. (A1), such operator can be written in the form

$$\hat{\Omega}_i^{(I)} = \sum_j c_j^{(I)} \hat{X}_{ij}, \quad (\text{A3})$$

where $|\phi_j\rangle = \hat{X}_{ij} |\phi_i\rangle$, hence $\hat{X}_{ii} = \hat{1}$. Note that $\hat{\Omega}_i^{(I)}$ is real-valued when $|I\rangle$ is real, as assumed here. In Eq. (A2), we can explicitly expand $|I\rangle$, and substitute $\langle I |$ with $\langle \phi_i | \hat{\Omega}_i^{(I)}$ to obtain

$$\begin{aligned} g_\alpha &= \frac{1}{N_s} \left| \sum_{I=1}^{N_s} \sum_{i=1}^L c_i^{(I)} \text{Im} \left(\langle \phi_i | \hat{\Omega}_i^{(I)} \hat{H} \hat{T}_\alpha | \phi_i \rangle \right) \right| \\ &= \frac{1}{N_s} \left| \text{Im} \left(\sum_{i=1}^L \langle \phi_i | \hat{H}_i^{(SA)} \hat{T}_\alpha | \phi_i \rangle \right) \right| \end{aligned} \quad (\text{A4})$$

where we have defined

$$\hat{H}_i^{(SA)} = \sum_{I=1}^{N_s} c_i^{(I)} \hat{\Omega}_i^{(I)} \hat{H}, \quad (\text{A5})$$

which is a linear combination of $\hat{H}$ left-multiplied by $\hat{X}_{ij}$ operators. L is the total number of unique Slater determinants used across all multireference states $|I\rangle$. Factorizing candidate Pauli term $\hat{T}_\alpha$ as $\theta_\alpha \hat{X}_\alpha \hat{Z}_\alpha$ leads to

$$g_\alpha = \frac{|\text{Im}(\theta_\alpha)|}{N_s} \left| \sum_{i=1}^L \lambda_i^{(\alpha)} \Xi_i^{(\alpha)} \right|, \quad (\text{A6})$$

with $\lambda_i^{(\alpha)} = \langle \phi_i | \hat{Z}_\alpha | \phi_i \rangle$, and $\Xi_i^{(\alpha)} = \langle \phi_i | \hat{H}_i^{(SA)} \hat{X}_\alpha | \phi_i \rangle$. From Eq. (A6), the multiconfigurational ensemble energy gradient exhibits the same decoupling of the $\hat{X}_\alpha$

and $\hat{Z}_\alpha$ degrees of freedom in $\hat{T}_\alpha$, up to θ_α , as for the case of ensembles of single determinant references in Section II 2. The $\hat{Z}_\alpha$ choice fully dictates the relative phases $\{\lambda_i^{(\alpha)}\}_i$, whereas $\hat{X}_\alpha$'s role is in expectation values $\Xi_i^{(\alpha)}$. Such expectation values can be connected to expectation values of the generalized Ising parts in $\hat{H}$ in Eq. (11) following simple rules, which we now derive.

By expanding $\hat{\Omega}_i^{(I)}$ in Eq. (A5), and inserting Ising factorized form of $\hat{H} = \sum_k \hat{D}_k \hat{X}_k = \sum_k \hat{X}_k \hat{D}_k$ (with second equality holding due to assumed realness of $\hat{H}$), we obtain

$$\hat{H}_i^{(SA)} = \sum_{I=1}^{N_s} c_i^{(I)} \sum_{j=1}^L \sum_k c_j^{(I)} \hat{X}_{ij} \hat{X}_k \hat{D}_k \quad (\text{A7})$$

$$= \sum_{I=1}^{N_s} c_i^{(I)} \sum_{j=1}^L \sum_k c_j^{(I)} \hat{X}_k^{(ij)} \hat{D}_k, \quad (\text{A8})$$

where we have defined $\hat{X}_k^{(ij)} = \hat{X}_{ij} \hat{X}_k$. This leads to

$$\begin{aligned} \Xi_i^{(\alpha)} &= \sum_{I=1}^{N_s} \sum_{j=1}^L \sum_k c_i^{(I)} c_j^{(I)} \langle \phi_i | \hat{X}_k^{(ij)} \hat{X}_\alpha | \phi_i \rangle \\ &\quad \times \langle \hat{X}_k^{(ij)} \phi_i | \hat{D}_k | \hat{X}_k^{(ij)} \phi_i \rangle, \end{aligned} \quad (\text{A9})$$

where $|\hat{X}_k^{(ij)} \phi_i\rangle = \hat{X}_k^{(ij)} |\phi_i\rangle$. Hence, $\Xi_i^{(\alpha)}$ is zero unless there exists at least one instance of $\hat{X}_k^{(ij)} = \hat{X}_\alpha$, which additively contributes $w_{IC_i^{(I)}} c_j^{(I)} \langle \hat{X}_k^{(ij)} \phi_i | \hat{D}_k | \hat{X}_k^{(ij)} \phi_i \rangle$ to $\Xi_i^{(\alpha)}$. To efficiently screen the $\hat{X}_\alpha$'s which lead to non-zero g_α , we consider all possible $\hat{X}_k^{(ij)}$'s, i.e., possible products from coupling $\hat{X}_{ij}$ operators in the wave operators $\hat{\Omega}_i^{(I)}$'s, and $\hat{X}_k$'s appearing in the Ising factorized form of $\hat{H}$, Eq. (11).

Appendix B: Phase-alignment procedures

For general ensembles of reference states, the energy gradient magnitude expression for generator $\hat{T}_\alpha$ takes the form of

$$g_\alpha = \frac{|\text{Im}(\theta_\alpha)|}{N_s} \left| \sum_{i=1}^L \lambda_i^{(\alpha)} \Xi_i^{(\alpha)} \right|, \quad (\text{B1})$$

Recall that for a given $\hat{T}_\alpha$ in the direct interaction space, $\lambda_i^{(\alpha)}$ depends only on the choice of $\hat{Z}_\alpha$, and $\Xi_i^{(\alpha)}$ depends only on $\hat{X}_\alpha$. We represent the operators $\hat{Z}_\alpha$ and $\hat{X}_\alpha$ as tensor products

$$\hat{Z}_\alpha = \prod_{p=1}^{N_q} \hat{z}_p^{\nu_p^{(\alpha)}} \quad (\text{B2})$$

$$\hat{X}_\alpha = \prod_{p=1}^{N_q} \hat{x}_p^{\mu_p^{(\alpha)}} \quad (\text{B3})$$

where $\mu_p^{(\alpha)}, \nu_p^{(\alpha)} \in \{0, 1\}$ are the p^{th} elements of the vectors $\vec{\mu}^{(\alpha)}$ and $\vec{\nu}^{(\alpha)}$ respectively. In this Appendix, we describe how to select $\hat{Z}_\alpha$ which maximizes g_α , for a given $\hat{X}_\alpha$. We formulate this problem in the domain of $\hat{Z}_\alpha$'s binary representation, $\vec{\nu}^{(\alpha)}$. We describe the strategy for finding optimal $\vec{\nu}^{(\alpha)}$ in Appendix B 1, and a heuristic yet efficient strategy in Appendix B 2.

Recall $g_\alpha = 0$ unless $\theta_\alpha \in \{i, -i\}$, which is the case when $\hat{X}_\alpha$ and $\hat{Z}_\alpha$ have odd overlapping support. In terms of the binary vectors, this leads to the requirement

$$\vec{\mu}^{(\alpha)} \cdot \vec{\nu}^{(\alpha)} \pmod{2} = 1. \quad (\text{B4})$$

The relative phases $\lambda_i^{(\alpha)}$ are obtained as,

$$\lambda_i^{(\alpha)} = \langle \phi_i | \hat{Z}_\alpha | \phi_i \rangle = \prod_{p=1}^{N_q} \langle \phi_i^{(p)} | \hat{z}_\alpha^{\nu_p^{(\alpha)}} | \phi_i^{(p)} \rangle \in \{1, -1\}, \quad (\text{B5})$$

where $|\phi_i^{(p)}\rangle \in \{|0\rangle, |1\rangle\}$. Let us introduce an N_q -bit binary vector $\vec{\phi}^{(i)} = (\phi_1^{(i)}, \dots, \phi_{N_q}^{(i)})$ representing the computational basis state $|\phi_i\rangle$, that is

$$\phi_p^{(i)} = \begin{cases} 1 & \text{if } |\phi_i^{(p)}\rangle = |1\rangle \\ 0 & \text{if } |\phi_i^{(p)}\rangle = |0\rangle \end{cases}. \quad (\text{B6})$$

Phase $\lambda_i^{(\alpha)} = -1$ if and only if there are an odd number of instances where $\nu_p^{(\alpha)} = \phi_p^{(i)} = 1$, otherwise $\lambda_i^{(\alpha)} = 1$. Hence, $\lambda_i^{(\alpha)}$ is written as a function of binary vectors $\vec{\nu}^{(\alpha)}$ and $\vec{\phi}^{(i)}$ as

$$\lambda_i^{(\alpha)} = 1 - 2 \left(\vec{\phi}^{(i)} \cdot \vec{\nu}^{(\alpha)} \pmod{2} \right). \quad (\text{B7})$$

Thus we can express the gradient (up to normalization by N_s) as a cost function in terms of $\nu^{(\alpha)}$ and formulate the problem as an N_q bit constrained binary optimization

$$\begin{aligned} & \max_{\vec{\nu}^{(\alpha)}} C(\vec{\nu}^{(\alpha)}) \\ & \text{subject to: } \vec{\mu}^{(\alpha)} \cdot \vec{\nu}^{(\alpha)} \pmod{2} = 1 \end{aligned} \quad (\text{B8})$$

where

$$C(\vec{\nu}^{(\alpha)}) = \left| \sum_{i=1}^L \Xi_i^{(\alpha)} \left[1 - 2 \left(\vec{\phi}^{(i)} \cdot \vec{\nu}^{(\alpha)} \pmod{2} \right) \right] \right|. \quad (\text{B9})$$

We will look at two different strategies at solving this binary optimization problem.

1. Optimal strategy

To find optimal solution to (B8), note that the cost function, Eq. (B9), consists of L distinct N_q -bit clauses,

and hence is generally exponentially hard to find the optimal value of $\vec{\nu}^{(\alpha)}$. Since $\Xi_i^{(\alpha)}$ has been precomputed, evaluation of Eq. (B9) is computationally fast, with scaling $O(LN_q)$. We find the optimal solution by brute force search of the constrained space of binary vectors. Since this is implausible for sufficiently large N_q , heuristic binary optimization strategies can be employed. In the next section, we discuss one such heuristic method.

2. Greedy satisfiability

It turns out one can formulate the question, is there a $\vec{\nu}^{(\alpha)}$ which aligns signs of *all* terms $\{\lambda_i^{(\alpha)} \Xi_i^{(\alpha)}\}_{i=1}^L$ in Eq. (B1), as an efficiently solvable satisfiability problem. If satisfied, the algorithm returns the satisfying $\vec{\nu}^{(\alpha)}$. If unsatisfiable, one can remove constraints existing in the problem and check again for satisfiability. We refer to this procedure as the **GreedySAT** phase-alignment routine, and explain it in further detail below.

The idealized scenario occurs when for all nonzero $\Xi_i^{(\alpha)}$,

$$\text{sgn}(\lambda_i^{(\alpha)} \Xi_i^{(\alpha)}) = \text{sgn}(\lambda_i^{(\alpha)} \Xi_i^{(\alpha)}), \quad (\text{B10})$$

where $\text{sgn}(x)$ is the signum function for $x \in \mathbb{R}$,

$$\text{sgn}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0. \end{cases} \quad (\text{B11})$$

To satisfy Eq. (B10), two equivalent assignments exist,

$$\lambda_i^{(\alpha)} = (\pm) \text{sgn}(\Xi_i^{(\alpha)}), \quad (\text{B12})$$

for all considered $\Xi_i^{(\alpha)}$, where the choice of $(\pm)$ is held fixed for all i .

Attempting to find $\vec{\nu}^{(\alpha)}$ which satisfies Eq. (B12) for all $L' \leq L$ non-zero $\Xi_i^{(\alpha)}$'s, along with the requirement of Eq. (B4), leads to a system of $L' + 1$ equations on the N_q binary variables in $\vec{\nu}^{(\alpha)}$. To see this, substitute Eq. (B12) in Eq. (B7) to get

$$\vec{\phi}^{(i)} \cdot \vec{\nu}^{(\alpha)} \pmod{2} = \frac{1 - (\pm) \text{sgn}(\Xi_i^{(\alpha)})}{2} \quad (\text{B13})$$

which are L' set of equations for $i \in \{1 \dots L'\}$, and one more equation comes from the constraint in Eq. (B4). We can then formulate the system of equations as

$$\mathbf{M} \vec{\nu}^{(\alpha)} = \vec{\mathbf{b}} \quad \text{over } \mathbb{F}_2, \quad (\text{B14})$$

where $\mathbf{M}$ is a $(L' + 1) \times N_q$ matrix,

$$\mathbf{M} = \begin{pmatrix} \vec{\mu}_1^{(\alpha)} & \dots & \vec{\mu}_{N_q}^{(\alpha)} \\ \vec{\phi}_1^{(1)} & \dots & \vec{\phi}_{N_q}^{(1)} \\ \vdots & & \vdots \\ \vec{\phi}_1^{(L')} & \dots & \vec{\phi}_{N_q}^{(L')} \end{pmatrix}, \quad (\text{B15})$$

and $\vec{b}$ is a $L' + 1$ dimensional vector,

$$\vec{b} = \begin{pmatrix} 1 \\ \left[1 - (\pm)\text{sgn}(\Xi_1^{(\alpha)})\right]/2 \\ \vdots \\ \left[1 - (\pm)\text{sgn}(\Xi_{L'}^{(\alpha)})\right]/2 \end{pmatrix}. \quad (\text{B16})$$

The restriction of matrix arithmetic to the space $\mathbb{F}_2$ effectively ensures the *modulo 2* operations in equations (B4) and (B13). We can solve Eq. (B14) by binary Gaussian elimination, for instance. In our numerical examples, we utilize the **SageMath** package to this end.

We now describe the complete **GreedySAT** procedure below, which includes a prescription for when no solution to Eq. (B14) can be found. Essentially, if no solution can be found for considering the phase-alignment of all L' terms via satisfying all Eq. (B13), we remove consideration of a specific instance of Eq. (B13) associated with the lowest valued $|\Xi_k^{(\alpha)}|$, and attempt to solve the system of fewer equations. This removal of least-important constraints is iteratively performed until a solution $\vec{\nu}^{(\alpha)}$ has been found, and the true value of the associated g_α is computed via Eq. (B9). The procedure can be summarized as follows:

1. For a given $\vec{\mu}^{(\alpha)}$ corresponding to an element of the direct interaction space, evaluate $\vec{\Xi}^{(\alpha)} = (\Xi_1^{(\alpha)}, \Xi_2^{(\alpha)} \dots \Xi_{L'}^{(\alpha)})$, where $\vec{\Xi}^{(\alpha)}$ has only non zero elements, arranged in decreasing order of their absolute values.
2. Attempt to find solution to Eq. (B14) for the (+) assignment in Eq. (B16). If no solution exists, attempt to solve the same system with the (−) assignment. If a solution has been found for either (±) assignments, go to Step 4 with solution $\vec{\nu}^{(\alpha)}$ to Eq. (B14), otherwise, enter Step 3.
3. Redefine $\vec{\Xi}^{(\alpha)} = (\Xi_1^{(\alpha)}, \Xi_2^{(\alpha)} \dots \Xi_{L'}^{(\alpha)})$ with $L' \rightarrow L' - 1$, by dropping the smallest absolute element of $\vec{\Xi}^{(\alpha)}$ and re-enter step 2.
4. Once a solution $\vec{\nu}^{(\alpha)}$ has been found, the corresponding Pauli term $\hat{T}_\alpha$ has $\hat{X}_\alpha$ and $\hat{Z}_\alpha$ parts given by $\vec{\mu}^{(\alpha)}$ and $\vec{\nu}^{(\alpha)}$ respectively, and its gradient magnitude g_α is given by inserting the found $\vec{\nu}^{(\alpha)}$ into

Eq. (B9). This resulting $\hat{T}_\alpha$ represents the highest g_α candidate found by the **GreedySAT** routine.

3. Numerical assessment of GreedySAT

To assess the performance of the **GreedySAT**, we benchmark it against the **OPT** strategy in producing the highest g_α [Eq. (B1)] for a fixed $\hat{X}_\alpha$, over a class of uniformly random phase-alignment problems. Such problems are obtained by the following prescription:

- The L configurations $\{|\phi_i\rangle\}_{i=1}^L$ defining the model space are independently uniformly sampled from the set of N_q -bit strings, $\{0, 1\}^{\otimes N_q}$.
- Similarly, the $\hat{X}_\alpha$'s binary representation, $\vec{\mu}^{(\alpha)}$, is also uniformly sampled from the set of N_q -bit strings.
- The values of $\{\Xi_i^{(\alpha)}\}_{i=1}^L$ are independently uniformly sampled from within the interval $[-1, 1]$.

To denote a specific instance, we label the problem by $g_\alpha^{\text{SAT/OPT}}(\{|\phi_i\rangle\}, \{\Xi_i^{(\alpha)}\})$ where the superscript refers to which of GreedySAT and OPT was used. To quantify the performance we take their ratios,

$$R_\alpha(\{|\phi_i\rangle\}, \{\Xi_i^{(\alpha)}\}) = \frac{g_\alpha^{\text{SAT}}(\{|\phi_i\rangle\}, \{\Xi_i^{(\alpha)}\})}{g_\alpha^{\text{OPT}}(\{|\phi_i\rangle\}, \{\Xi_i^{(\alpha)}\})}.$$

By performing many uniformly random samples of the phase-alignment problem, we empirically approximate the average case ratios:

$$R_{\text{avg}} = \mathbb{E}_{\{|\phi_i\rangle\}, \{\Xi_i^{(\alpha)}\}, \vec{\mu}^{(\alpha)}} R_\alpha(\{|\phi_i\rangle\}, \{\Xi_i^{(\alpha)}\}).$$

In Table I we report the values of R_{avg} for a range of phase-alignment problem instances. Each reported mean is obtained from 100 uniformly sampled phase-alignment problems of L -dimensional model spaces defined on N_q qubits. Since MS-iQCC does not require large model spaces for the multiconfigurational reference states, we restrict $L \leq N_q$ in this analysis. The mean of means of ratios obtained across all (N_q, L) 's considered is 0.995 ± 0.009 , i.e., it is nearly indistinguishable in average performance from the **OPT** solution. While the performance may begin to suffer in the $L > N_q$ regime, where the number of equations is larger than number of free variables, this is generally not considered problematic in the context of MS-iQCC where relatively simple CSFs may be used as reference states, leading to small model spaces.

Table I. Empirical average gradient ratios R_{avg} for the **GreedySAT** routine applied to uniformly random phase-alignment problems of N_q qubits and L -dimensional model space, obtained using 100 samples of problems. One standard deviation σ is included as $\pm\sigma$.

N_q	L									
	2	4	6	8	10	12	14	16	18	20
2	1.0 ± 0.0									
4	1.0 ± 0.0	0.974 ± 0.139								
6	1.0 ± 0.0	0.99 ± 0.064	0.975 ± 0.111							
8	1.0 ± 0.0	0.99 ± 0.095	0.979 ± 0.092	0.944 ± 0.18						
10	1.0 ± 0.0	1.0 ± 0.0	0.997 ± 0.021	0.997 ± 0.023	0.985 ± 0.052					
12	1.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.0	0.996 ± 0.036	0.989 ± 0.097	0.985 ± 0.055				
14	1.0 ± 0.0	1.0 ± 0.0	0.997 ± 0.031	1.0 ± 0.0	1.0 ± 0.0	0.999 ± 0.006	0.986 ± 0.063			
16	1.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.0	0.997 ± 0.027	0.999 ± 0.006	0.993 ± 0.02		
18	1.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.0	0.999 ± 0.012	1.0 ± 0.005	1.0 ± 0.0	1.0 ± 0.001	0.995 ± 0.016	
20	1.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.0	1.0 ± 0.003	1.0 ± 0.0	0.999 ± 0.005	0.992 ± 0.025