INDUSTRIAL AND SYSTEMS ENGINEERING

ISE

Unleashed from Constrained Optimization: Quantum Computing for Quantum Chemistry Employing Generator Coordinate Method

MUQING ZHENG[1,2], BO PENG[2], ANG LI[2], XIU YANG[1], AND KAROL KOWALSKI[2]

[1]Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, PA, USA

[2]Pacific Northwest National Laboratory, Richland, Washington, 99354, USA

ISE Technical Report 23T-028

LEHIGH UNIVERSITY

Unleashed from Constrained Optimization: Quantum Computing for Quantum Chemistry Employing Generator Coordinate Method

Muqing Zheng,[1,2] Bo Peng,[2,a)] Ang Li,[2] Xiu Yang,[1] and Karol Kowalski[2]
[1)]*Lehigh University, Bethlehem, Pennsylvania, 18015, USA*
[2)]*Pacific Northwest National Laboratory, Richland, Washington, 99354, USA*

(Dated: December 2023)

Highlights

- We categorize the challenges in hybrid quantum-classical approaches, such as the barren plateau and the exactness of the ansätze, as constrained optimization problems.
- We demonstrate the interconnection between constrained optimization and the generalized eigenvalue problem in quantum chemistry computations. This is achieved using a distinctive class of non-orthogonal and overcomplete basis sets generated by Givens rotation-type canonical transformations on a reference state.
- Using the generator coordinate approach, we represent the wave function with these basis sets. The resulting generalized eigenvalue problem establishes rigorous lower bounds on energy, surpassing the performance of the conventional variational quantum eigensolver (VQE) that utilizes the same canonical transformations in its ansätz.
- For practical applications, we introduce an adaptive scheme to select these transformations. This strategy features the linear expansion of the non-orthogonal basis set and bypasses the optimization, ensuring a balanced blend of accuracy and efficiency in hybrid quantum-classical simulations.
- Our methodology paves the way for broader applications in quantum chemistry, which encompasses simultaneous ground and excited state simulations and the generation of tailored projected/effective Hamiltonians for large scale simulations.

Abstract Hybrid quantum-classical approaches offer potential solutions for quantum chemistry problems, but they also introduce challenges such as the barren plateau and the exactness of the ansätze. These challenges often manifest as constrained optimization problems without a guarantee of identifying global minima. In this work, we highlight the interconnection between constrained optimization and generalized eigenvalue problems, using a unique class of non-orthogonal and overcomplete basis sets generated by Givens rotation-type canonical transformations on a reference state. Employing the generator coordinate approach, we represent the wave function in terms of these basis sets. The ensuing generalized eigenvalue problem yields rigorous lower bounds on energy, outperforming the conventional variational quantum eigensolver (VQE) that employs the same canonical transformations in its ansätze. Our approach effectively tackles the barren plateau issue and the heuristic nature of numerical minimizers in the standard VQE, making it ideal for intricate quantum chemical challenges. For practical applications, we propose an adaptive scheme for selecting these transformations, emphasizing the linear expansion of the non-orthogonal basis sets. This ensures a harmonious balance between accuracy and efficiency in hybrid quantum-classical simulations. Our analysis and suggested methodology further broaden the applications of quantum computing in quantum chemistry. Notably, they pave the way for alternative strategies in excited state computation and Hamiltonian downfolding, laying the groundwork for sophisticated quantum simulations in chemistry.

I. INTRODUCTION

Accurately obtaining ground and excited state energies, along with the corresponding many-body wave functions, is pivotal in comprehending diverse physical phenomena in molecules and materials. This ranges from high-temperature superconductivity in materials like cuprates[1] and bond-breaking chemical reactions to complex electronic processes in biological and synthetic catalysts with transition metals[2] or f-block atoms[3]. The associated spin, electronic properties, and dynamics are crucial for deciphering the structure-property-function correlation in various fields, including catalysis, sensors, and quantum materials. A prime example is the oxygen-evolving complex in photosystem II of green plants, essential for water molecule oxidation and photosynthesis[4,5]. This complex, involving manganese atoms in various oxidation states, facilitates electron-transfer reactions for water oxidation. However, this task becomes exceptionally challenging in the presence of non-trivial quantum effects, such as strong electron correlation, which influence the evolution of nuclei, electrons, and spins under external stimuli. Conventional wave function methodologies, like configuration interaction, coupled cluster, and many-body perturbation theory, are tailored for diverse electron correlation scenarios[6]. Still, they often fall short in handling complex cases or exhibit prohibitive scaling with increasing system size. Consequently, this has become a vigorous area of computational research, encompassing both classical and burgeoning quantum computing studies. The primary objective is to strike an optimal balance between accuracy and computational scalability.

In the realm of quantum computing, significant strides

[a)]Electronic mail: peng398@pnnl.gov

have been made towards promising near-term hybrid quantum-classical strategies, which includes variational quantum algorithm [7–21], quantum approximate optimization algorithm[22,23], quantum annealing[24,25], Gaussian boson sampling[26], analog quantum simulation[27,28], iterative quantum assisted eigensolver[19,29–31], and many others. These approaches typically delegate certain computational tasks to classical computers, thereby conserving quantum resources in contrast to exclusively to quantum methods. Within this framework, the variational quantum eigensolver (VQE) and its adaptive derivatives are seen as the frontrunners in leveraging near-future quantum advantages[7,16]. However, these anticipations are also potentially impeded by the heuristic nature inherent in the critical optimization processes. Issues such as the rigor of ansätze exactness, the challenge in navigating potential energy surfaces replete with numerous local minima, and the numerical nuances in minimization techniques, remain nebulous. These queries are further convoluted when considering the scalability of these methods concerning the number of operators or the depth of quantum circuits involved.

As an alternative to VQE, which employs a highly nonlinear parametrization of the wave function, other near-term strategies aim to construct and traverse a subspace within the Hilbert space to closely approximate the desired state and energy. Typical examples include the quantum subspace expansion[13,32–34], the hybrid and quantum Lanczos approaches[29,35,36], quantum computed moment approaches[31,37–40], and quantum equation-of-motion approach[41]. These strategies often draw inspiration from truncated configuration interaction approaches or explore the Krylov subspace.However, the comparison and interrelation of these methods, especially when contrasting subspace expansion with nonlinear optimization, frequently remain unclear. This uncertainty is exacerbated by the choice of ansätze, which rarely guarantees exactness.

Recently, we employed the generator coordinate method (GCM)[42–49] and nonorthogonal quantum eigensolver as an alternative near-term approach[50], potentially addressing the limitations identified in the VQE. Similar to its contemporaries, the GCM employs low-depth quantum circuits and efficiently utilizes existing ansätze to explore a subspace, targeting specific states and energy levels. Nonetheless, the original GCM demands a priori knowledge of the system to meticulously select both the ansätze and generator, thereby circumventing heuristic approaches but potentially affecting its scalability and efficiency.

In this study, we discerned that the Unitary Coupled Cluster (UCC) excitation generator, primarily implemented through Givens rotations, provides a framework for delineating the differences and connections between the VQE and GCM. This understanding illuminated why, given a consistent ansätze, solving a generalized eigenvalue problem within the GCM framework could establish a lower boundary compared to the customarily constrained optimization problem encountered in the VQE method. Equipped with this knowledge, we implemented a disentangled UCC ansätze and an adaptive generator selection scheme, aiming to overcome the challenges present in the original GCM. Our results suggest that the new ADAPT-GCM, guided by the subspace expansion using the employed UCC disentangled ans"atze, outperforms the ADAPT-VQE method. This superiority is evidenced by numerical results across various benchmark systems, particularly in instances of strong electron correlation.

II. GCM: A BRIEF OVERVIEW

Originating from studies of nuclear structure and nuclear reactions, including those occurring in heavy-ion collisions or astrophysical scenarios, the Generator Coordinate Method (GCM), proposed by Griffin-Hill-Wheeler (GHW)[42,43], approximates the many-body wave function $|\Psi\rangle$ of the Hamiltonian H by describing collective motion through the variation of one or more collective generator coordinates, denoted as $\boldsymbol{\alpha} = \{\alpha_i\}$, where $i = 1, \cdots, N_\alpha$ and $N_\alpha > 1$. This variation leads to the representation of the many-body wave function of a quantum system in terms of a set of basis states (also called generating functions) $\{|\phi(\boldsymbol{\alpha})\rangle\}$:

$$|\Psi\rangle \approx |\psi_{\mathrm{GCM}}\rangle = \int \mathrm{d}\boldsymbol{\alpha} |\phi(\boldsymbol{\alpha})\rangle f(\boldsymbol{\alpha}). \tag{1}$$

Here, $f(\boldsymbol{\alpha})$ is an unknown $\boldsymbol{\alpha}$-dependent weight function. The choice of these generator coordinates $\boldsymbol{\alpha}$ depends on the specific system under study. The fundamental concept of the GCM is to treat these generator coordinates as variational parameters, which are then used to approximate the target wave function as a superposition of $\{|\phi(\boldsymbol{\alpha})\rangle\}$. The GCM's ability to adapt to changing collective coordinates and its variational nature make it well-suited for describing complex states that break symmetries, allowing for a flexible and efficient description of properties such as ground states, excited states, and transitions between them.

Technically, to determine $f(\boldsymbol{\alpha})$, substituting (1) into the Schrödinger equation, and then the variational process leads to the so-called GHW integral equation:

$$\int \mathrm{d}\boldsymbol{\alpha}' \Big(\langle \phi(\boldsymbol{\alpha}) | H | \phi(\boldsymbol{\alpha}') \rangle - E \langle \phi(\boldsymbol{\alpha}) | \phi(\boldsymbol{\alpha}') \rangle \Big) f(\boldsymbol{\alpha}') = 0. \tag{2}$$

This integral equation can be discretized in a finite-dimensional subspace of the many-body Hilbert space, leading to a general eigenvalue problem:

$$\sum_{j=1}^{N_\alpha} \Big(\mathbf{H}_{ij} - E \mathbf{S}_{ij} \Big) f(\alpha_j) = 0, \tag{3}$$

where

$$\mathbf{H}_{ij} = \langle \phi(\alpha_i) | H | \phi(\alpha_j) \rangle,$$
$$\mathbf{S}_{ij} = \langle \phi(\alpha_i) | \phi(\alpha_j) \rangle.$$

While the GCM is a powerful method, it can be computationally demanding, especially when dealing with a large number of basis states or when symmetry-breaking effects are intricate. For strongly correlated molecular systems, these requirements are often associated with incorporating sufficient electron correlation effects into GCM

calculations, which typically involves using more sophisticated electronic structure methods than mean-field theory. Techniques such as configuration interaction (CI), coupled-cluster (CC) methods, and many-body perturbation theory (MBPT) can be employed to capture electron correlation effects accurately[6]. The basis functions generated from these methods should be capable of representing the electron correlation effects relevant to the quantum phenomena under investigation. Various levels of electron correlation can be included in GCM calculations, depending on the chosen basis functions and computational resources available. It is interesting to observe the GCM as a generalization of the non-orthogonal determinants (e.g., non-orthogonal configuration interaction). Even though non-entangled single-particle states were employed in the GCM for simplifying the implementation in our previous report[50], the entangled many-body states can also be employed in the framework to involve complex correlations among particles, making it suitable for describing systems with complex shapes and collective motion.

III. GCM VS. VQE: GENERAL EIGENVALUE PROCESS VS. CONSTRAINED OPTIMIZATION

Given the unique properties of GCM, we believe it holds the potential to address the prevalent issues identified in VQE practices. One primary advantage is that in GCM, variation occurs on the generating function, not directly on the scalar generator coordinate. This distinction allows the variational problem to be tackled through a straightforward eigenvalue process. This approach stands in contrast to the method of solving a constrained optimization problem, the benefits of which can even be illustrated with a simple toy model.

As depicted in Fig. 1, the toy model—comprising two electrons and four spin orbitals—employs a VQE approach with a Trotterized UCC single-type ansätze:

$$|\psi_{\mathrm{VQE}}(\vec{\theta})\rangle = G_{2,4}(\theta_2)G_{1,3}(\theta_1)|\phi_0\rangle, \tag{4}$$

where the single excitation Givens rotation here is defined as

$$G_{p_i,q_i}(\theta_i) = \exp[\theta_i(A_{p_i,q_i} - A^{\dagger}_{p_i,q_i})]. \tag{5}$$

with $A_{p,q} = a^{\dagger}_p a_q$. Essentially, this VQE approach explores a two-dimensional parameter subspace. The consecutive actions of $G_{2,4}$ and $G_{1,3}$ generate a total of four distinct configurations, including the reference. However, since the number of free parameters in this case is fewer than the number of distinct configurations required, the VQE approach is constrained by this limitation to be unable to fully explore the subspace to guarantee the most optimal solution within the subspace (which happened to be the exact target state), irrespective of the numerical minimizer used.

In contrast, GCM employs the same Givens rotations, separately or in the product form as discussed by Fukutome[51], to create four generating functions. These functions correspond to a subspace consisting of four non-orthogonal superpositioned states, providing a scope that is sufficiently expansive to encapsulate the target state. Here, each Givens rotation is essentially equivalent to a UCC single circuit, simplifying the quantum implementation compared to its double or higher excitation analogs. It's important to note that

- Even though the exactness of the disentangled UCC ansätz has been proven[52], the ansätze employed in GCM, or in the generalized eigenvalue process, need not be exact. The only prerequisite for the ansätze used in this process is their ability to generate sufficient superpositioned states so that a better approximate to the target state can be found in the corresponding subspace.

- The Givens rotations used in both VQE and GCM may originate from the same set, ensuring a comparable level of complexity in the quantum circuits involved (to achieve the same level of exactness, VQE would require a deeper circuit by incorporating additional Givens rotations). While GCM necessitates more measurements than a single VQE iteration, a direct comparison of the total quantum resources utilized between the two methods is not straightforward; it hinges on both the number of generating functions required in GCM and the number of iterations needed in VQE.

- In the context of non-orthogonal quantum eigensolver including GCM approaches, to cover a wide range of correlation, entangled basis sets are generated by applying UCC operators to the reference single determinant (e.g., the k-UpCCGSD ansätz[53], UCC-derived ansatz and a cluster-Jastrow modification[54]). These methods often require domain-specific knowledge available in classical quantum chemistry to construct high-quality wave functions, and rely on the modified Hadamard test of Ref. 14 for evaluating the off-diagonal elements at a polynomial cost. Nevertheless, how to automate the sophisticated ansätz construction process such that the deployment of the non-orthgonal quantum solver can be easier even for the non-experts is still a challenge.

These observations then lead us to the pivotal question in advancing a more efficient GCM-based algorithm for both classical and quantum computation: "*How can we guide the selection of generating functions, or the Givens rotations, to construct a subspace in the Hilbert space that encompasses the target state?*"

IV. THE SCALING OF NUMBER OF GENERATING FUNCTIONS: FROM EXPONENTIAL TO LINEAR

In constructing a non-trivial wave function ansätz for a molecular system that comprises n_e electrons in N spin orbitals, one might consider applying a sequence of $m \leq n_e$ Givens rotations to a reference state $|\phi_0\rangle$ to generate all the possible configurations. Each Givens rotation defined in Eq. (5) corresponds to a single excitation, and generates a

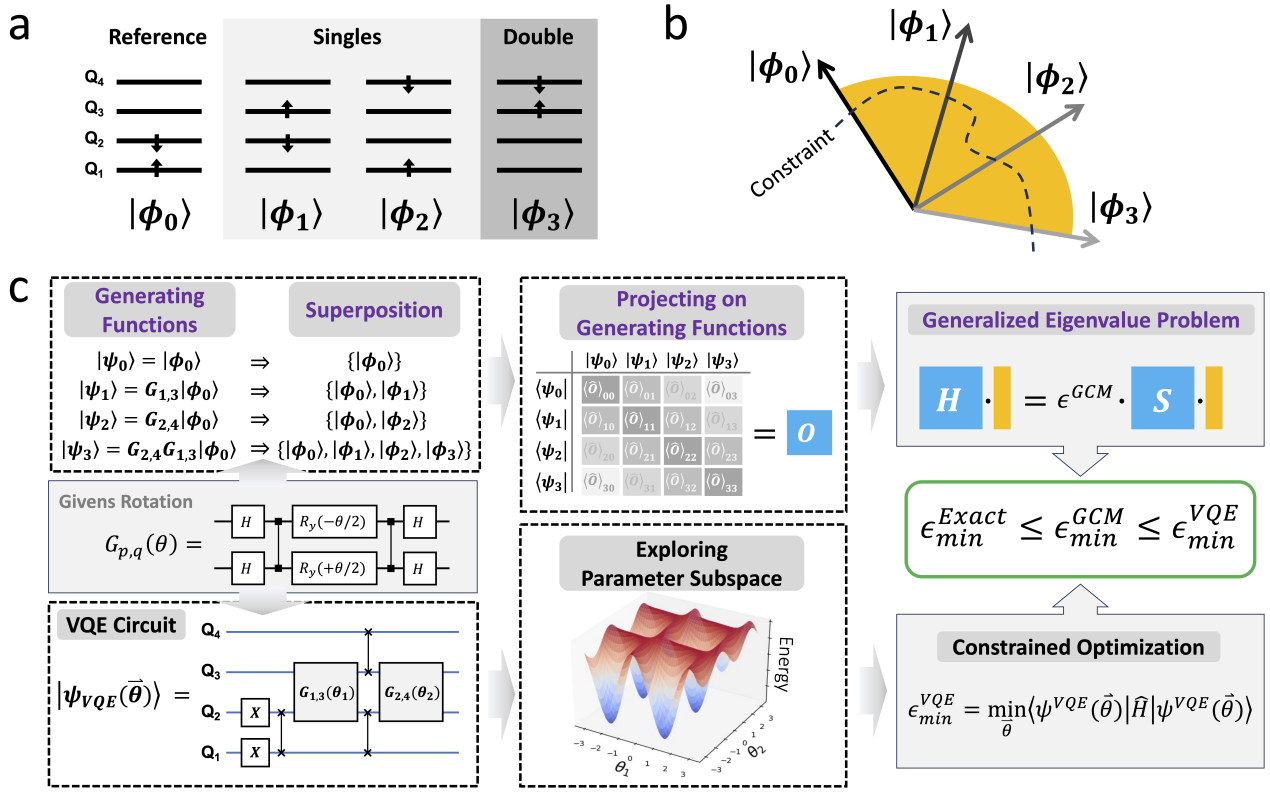

Figure 1. (a) A toy model consists of two-electron (one alpha electron and one beta electron) in four spin-orbitals (two alpha spin-orbitals and two beta spin-orbitals), where the spin-flip transition is assumed forbidden. (b) The projection of the exact wave function on each configuration (yellow shadown). A constrained optimization of the free parameters will put limits on the projection (dashed line). (c) Comparative demonstration between GCM and VQE. The GCM generates a set of non-orthogonal bases, called generating functions. Then the GCM explores the projection of the system on these generating functions and solves a corresponding generalized eigenvalue problem for target state and its energy. The conventional VQE essentially explores the parameter subspace for a given wave function ansätz through numerical optimization that can be usually constrained by many factors ranging from ansätz inexactness to barren plateaus and else. For given Givens rotations that generated excited state configurations, the lowest eigenvalue obtained from GCM guarantees a lower bound of the most optimal solution from the VQE.

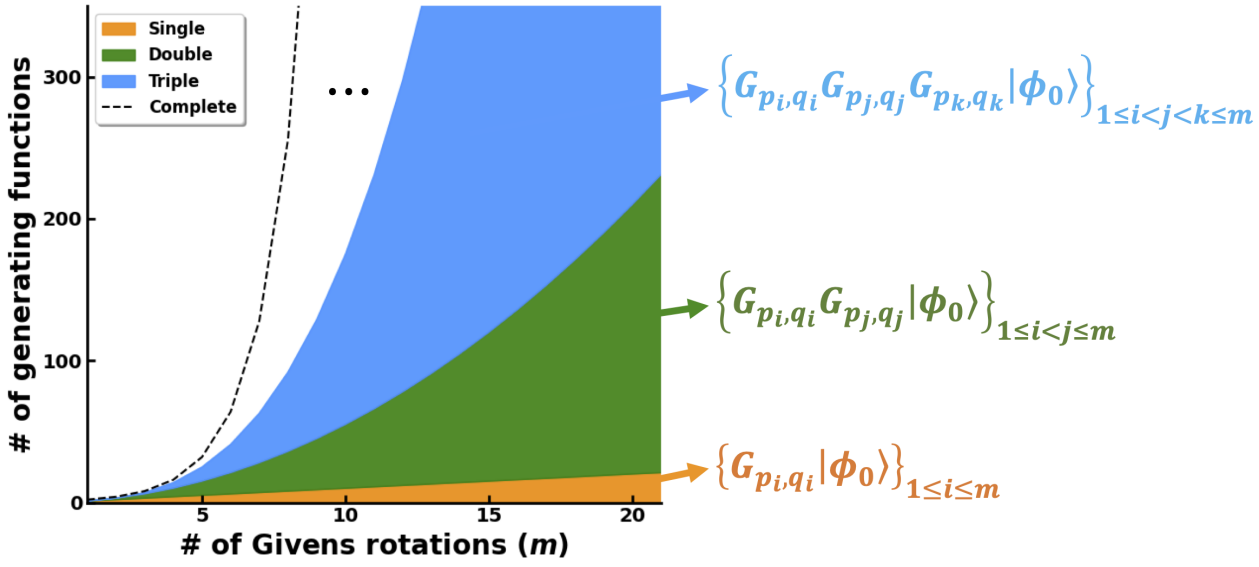

Figure 2. Number of generating functions as functions of the number of Givens rotations (denoted as m) included in the GCM wave function ansätz at different truncation levels.

superposition of no more than two states. The ansätz can then be expressed as:

$$|\psi(\vec{\theta})\rangle = \prod_{i=1}^{m} G_{p_i,q_i}(\theta_i)|\phi_0\rangle, \tag{6}$$

Consequently, the total number of configurations, n_c contained within the superpositioned ansätz (6), where $\vec{\theta} = \{\theta_i | i = 1, \cdots, m\}$ varies, cannot exceed 2^m. This figure represents the maximum number of generating functions utilized to achieve the most optimal solution within the corresponding subspace, exhibiting an exponential scaling relative to the count of Givens rotations applied.

A strategy to potentially generate the full set of 2^m generating functions entails applying $1, 2, \cdots, N$ Givens rotations separately to the reference, as shown below:

- $|\phi_0\rangle$,
- $G_{p_i,q_i}(\theta_i)|\phi_0\rangle$, $1 \leq i \leq m$,
- $G_{p_i,q_i}(\theta_i)G_{p_j,q_j}(\theta_j)|\phi_0\rangle$, $1 \leq i < j \leq m$,
- $G_{p_i,q_i}(\theta_i)G_{p_j,q_j}(\theta_j)G_{p_k,q_k}(\theta_k)|\phi_0\rangle$, $1 \leq i < j < k \leq m$,
- $\vdots$
- $\prod_{i=1}^{m} G_{p_i,q_i}(\theta_i)|\phi_0\rangle$

Here, the number of generating functions that employ k Givens rotations is equivalent to the number of k-combinations drawn from the set of m Givens rotations, aligning with the combinatorial identity:

$$\sum_{k=0}^{m} \binom{m}{k} = 2^m. \tag{7}$$

Importantly, when $m = n_e$, this methodology for constructing generating functions can encapsulate the complete configuration space of the molecular system, thereby guaranteeing exactness in this limit. This characteristic bears notable similarity to full configuration interaction treatments, with the distinctive aspect that this approach constructs a non-orthogonal many-body basis set.

The strategy outlined above can also be employed to establish a hierarchy of approximations to the GCM wave function, aiming towards a complete configuration subspace where the ansätz (6) can be fully projected out. Within this framework, the approximation at level k ($0 \leq k < m$) is defined by a working subspace that incorporates both the ansätze (6) and some auxiliary generating functions produced by acting a product of at most k Givens rotations on the reference. Notably, at level $k = 0$, the approximation narrows to the subspace containing only the ansätz (6) and $|\phi_0\rangle$. In this context, the lowest eigenvalue of the Hamiltonian H can still be less than or equal to its variational expectation value with respect to the ansätz (6) obtained by optimizing the rotations in the ansätz. When $k = m$, independent of the rotations in the generating functions, the configuration subspace has expanded sufficiently for the ansätz (6) to be fully projected out. For other approximation levels where $0 < k < m$, a variational procedure can be undertaken before the configurational subspace expansion, aiming to deduce a lower bound for the optimal expectation value of H with respect to the ansätz (6). In this fashion, the GCM can act as a subsequent step to further refine the energy, especially when the VQE and its ADAPT version, imposing constraints due to the inexactness of the ansätz and local minima.

It's important to recognize that the generating functions, including the ansätz, are not generally orthogonal. This non-orthogonality necessitates extra resources for evaluating the associated overlap matrix. Additionally, if there are higher-order Givens rotations involved, the way of the configuration subspace expansion is influenced by the ordering of the Givens rotations in the ansätz. However, this ordering limitation can be partially circumvented by permitting all permutations of the Givens rotations within the auxiliary generating functions. Such a step increases the measurement pre-factors without affecting the scaling behavior. For instance, complete permutations can be permitted when incorporating the auxiliary generating functions produced by the product of two Givens rotations respectively corresponding to one single excitation and one double excitation.

V. INTEGRATING GCM WITH ADAPT APPROACH

As mentioned in the previous section, the GCM can be implemented as a subsequent step either after each ADAPT-VQE iteration or following a complete ADAPT-VQE calculation. In the ensuing discussion, we designate these approaches as ADAPT-VQE-GCM and ADAPT-VQE-GCM(1), respectively (refer to Appendix C). The numeral '1' in ADAPT-VQE-GCM(1) signifies a 'one-shot' GCM calculation performed at the conclusion of an ADAPT-VQE calculation to enhance the results.

To facilitate a more robust selection of the auxiliary generating functions, we consider the fluctuation of the GCM energy when a new generating function is included in the basis set. In the VQE context, similar considerations have led to its ADAPT version. In the GCM framework, several routines can be proposed for a gradient-based ADAPT-GCM approach. One straightforward method is to directly compute the GCM energy gradient with respect to the scalar rotation (see Appendix B). The energy gradient computed in this manner depends on the corresponding eigenvector, which in turn requires solving the GCM general eigenvalue problem for every new basis. However, when solving a generalized eigenvalue problem–the eigenvalue (i.e. the energy) is more sensitive to the working subspace than to the scalar rotation in the employed generating functions. This is exemplified in the **Lemma 1** of **Theorem 1** in Appendix A, where, for a 2×2 generalized eigenvalue problem, the choice of the scalar rotations in the employed generating functions is relatively flexible.

Our proposed gradient-based ADAPT-GCM algorithm in this paper is summarized in Fig. 3a, in which we consider a surrogate product state $|\psi\rangle_s$ that, irrespective of its detailed non-singular scalar rotations, primarily serves as a metric to (1) abstract the change in the GCM working subspace and (2) approximate the GCM energy gradient calculation. Through this metric, we can then compute the

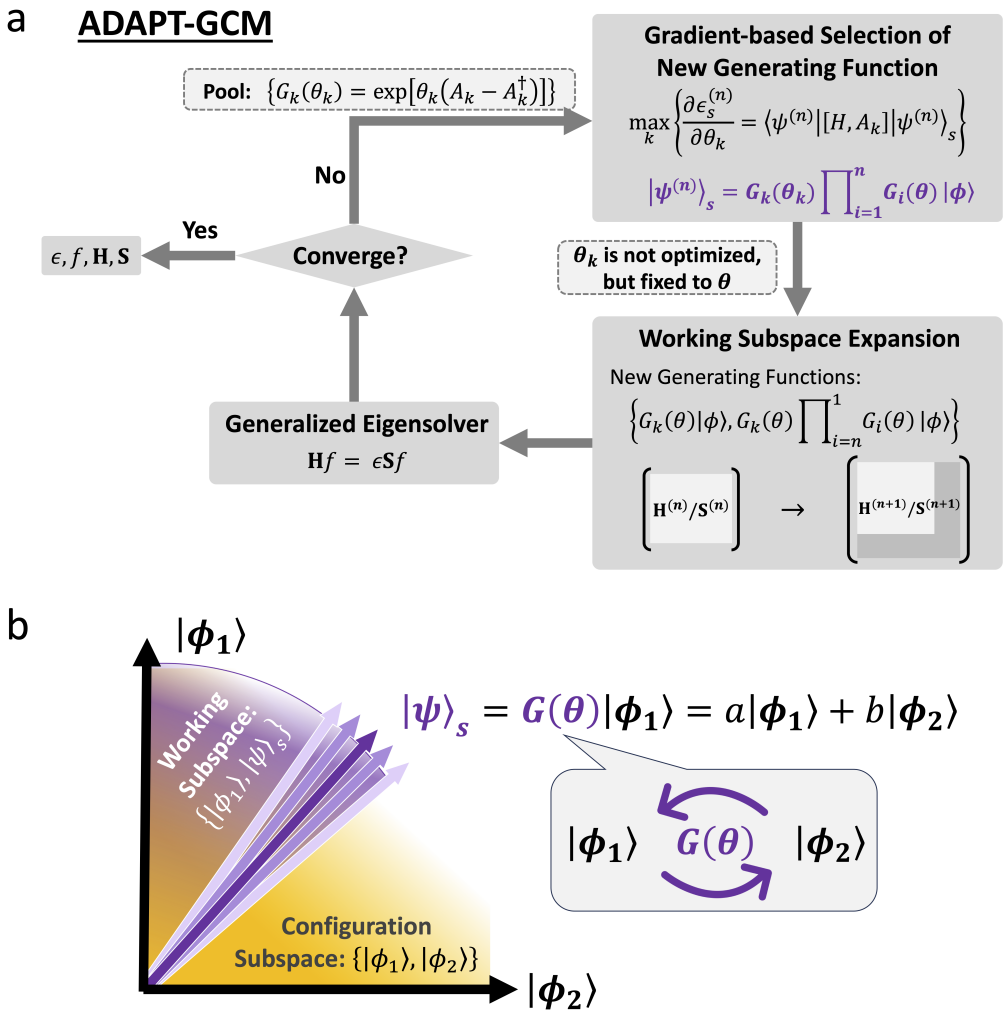

Figure 3. (a) Schematic depiction of the ADAPT-GCM algorithm. The working subspace and its corresponding configuration subspace are characterized by a surrogate state $|\psi\rangle_s$. (b) In a two-configuration subspace, a Givens rotation can couple two configurations, effectively generating a surrogate state, a superposition of two configurations, that can be employed as a metric to characterize this subspace. The quality of the working subspace, comprising one configuration and the surrogate state, is not sensitive to the specific choice of scalar rotation.

gradient in a straightforward way similar to that in the conventional ADAPT-VQE. But significantly different from ADAPT-VQE, we have to emphasize that this ADAPT-GCM approach skips the optimization of the phases and its energy is not from the measurement of the expectation value with respect to the surrogate state, but obtained from solving a generalized eigenvalue problem in the working subspace.

In the proposed ADAPT-VQE-GCM(1), the number of bases in the working subspace equals to the number of Givens rotations in the ansätz generated at the last ADAPT-VQE iteration, incremented by one. This increment accounts for the product of these Givens rotations corresponding to the ansätz itself. Conversely, in the ADAPT-GCM and ADAPT-VQE-GCM approaches, the number of bases is twice the number of iterations. To illustrate this, consider the n-th iteration: when the Givens rotation G_k with the greatest gradient is selected, two associated basis vectors,

- $G_k(\theta_k)|\phi_{\rm HF}\rangle$ and
- $G_k(\theta_k)\prod_{i=n-1}^{1} G_i(\theta_i)|\phi_{\rm HF}\rangle$

are added to the working subspace. So the number of bases is increased by two in every iteration. The key difference is that in ADAPT-GCM, the phase is set to a constant scalar (e.g., $\pi/4$) for any Givens rotation in the ansätz, while in ADAPT-VQE-GCM, the phases θ_1 to θ_{n-1} are optimized during the $(n-1)$-th iteration. In the following experiments of ADAPT-GCM in the next section, θ_k is set to $\pi/4$ for all iterations in all molecules.

VI. NUMERICAL EXPERIMENTS

We have performed numerical experiments on four molecular systems in six geometries for testing convergence performance of the proposed adapt versions of GCM approaches under different electronic configurations. For example, for H_4 and H_6 molecular systems, we consider two geometries for each, i.e. linear and square geometries

for H_4, and two H_6 geometries with different $H-H$ length. The more pronounced quasidegeneracy in the square H_4 and the stretched H_6 exhibit strong static correlations that are challenging for traditional single-reference methods[55]. The operator pool employed in the ADAPT approaches consists of all possible singly and doubly spin-adapted excitation operators. In all the ADAPT-VQE simulations, the convergence is achieved when the norm of the gradient vector is less than 10^{-4}. Here, we use the more strict criteria than in Ref. 56 to show the full picture of the performance of ADAPT-VQE. Other criteria, such as the magnitude of the change of the lowest eigenvalue, can be used for earlier convergence with the assistance of ADAPT-VQE-GCM. More details are listed in Appendix E.

The performance results are shown in Fig. 4, where the conventional ADAPT-VQE results are also included for comparison. Generally speaking, the GCM-related approaches provide several recipes for locating a lower bound to the ADAPT-VQE energy. As can be seen, in all the test cases, the ADAPT-VQE-GCM and ADAPT-GCM approaches exhibit a steady monotonic converging curve until close to the machine precision. The "one-shot" GCM, ADAPT-VQE-GCM(1), also slightly improves the ADAPT-VQE energy when the latter converges. Notably, the ADAPT-VQE-GCM gives a lower energy than the ADAPT-VQE *at every single iteration*, and thus converges faster. As a trade-off of skipping the phase optimization, the ADAPT-GCM generally requires more iterations to converge than ADAPT-VQE and ADAPT-VQE-GCM, but both ADAPT-GCM and ADAPT-VQE-GCM are capable to achieve exact numerical solutions for all the testing cases as indicated by the corresponding energy difference from the exact are $\leq 10^{-13}$ a.u. at the convergence.

One issue in the GCM simulations is the numerical instability of the generalized eigenvalue problem. It is possible that the overlap matrix $\mathbf{S}$ becomes numerically "singular " at certain iterations. In classical computing, the issue of singularity can be perfectly avoided by orthogonalizing the generating function basis set. However, in quantum computing, the quantum version of the orthogonalization algorithm like the quantum Gram-Schmidt process usually requires extra quantum resources, such as QRAM[57]. Assuming the orthogonalization is enabled, the number of iterations in the employed ADAPT procedures can be greatly reduced (see Appendix E). In the present study, when a singular overlap matrix is encountered, we choose to reconstruct an approximate but more stable overlap matrix by imposing a small disturbance ($\leq 10^{-14}$) on all diagonal entries of the overlap matrix.

VII. QUANTUM RESOURCE COMPARISON

In comparison to ADAPT-VQE, the ADAPT-VQE-GCM and ADAPT-VQE-GCM(1) require additional quantum resource to improve the ADAPT-VQE energy. However, the resource comparison between ADAPT-VQE and ADAPT-GCM is not straightforward. Therefore, in this section, we focus on the differences in quantum resources, specifically the number of CNOT gates and measurements required between ADAPT-VQE and ADAPT-GCM.

The number of CNOT gates in these two ADAPT approaches depends on how the basis sets are prepared at each iteration and the number of iterations needed to achieve a certain accuracy. In Fig. 5, we compare the CNOT counts for these two ADAPT approaches at each iteration, both before and after quantum circuit transpilation. For relatively small test cases, the CNOT counts in ADAPT-VQE state preparation are typically smaller than those in ADAPT-GCM, which includes only the single and product bases. For larger systems with stronger correlation, higher accuracy in ADAPT-VQE may require more CNOT gates than in ADAPT-GCM. It is important to note that this comparison focuses only on state or basis preparation in the two ADAPT approaches and does not account for the number of optimizations governed by the minimizer in ADAPT-VQE, which could potentially increase the CNOT counts. The actual numbers of classical optimizations used in each ADAPT iteration in numerical experiments are recorded in Fig. 6. Additionally, the circuit used in obtaining the $\mathbf{H}$ and $\mathbf{S}$ matrices in ADAPT-GCM differs from the Hadamard circuit used in ADAPT-VQE. Note that obtaining an off-diagonal matrix element requires more CNOT gates than the diagonal ones since two bases need to be prepared simultaneously in the former.

Regarding the number of measurements in the fault tolerant quantum computation, ADAPT-VQE requires measurements for two parts, VQE and gradient, at each iteration. The number of the measurements in VQE is not deterministic and depends on the average number of optimizations in the minimizer, $\tilde{N}_{\text{opt}}$, at each iteration. Assuming there are N_{iter} iterations, the number of measurements in this part is $\tilde{N}_{\text{opt}} \times N_{\text{iter}}$. Meanwhile, the number of gradient measurements roughly equals the size of the pool, which is the same as N_{iter}, multiplied by the number of terms in the Hamiltonian, N_{term}. The latter number can be further reduced by applying grouping techniques extensively discussed in the previous report[50]. Therefore, the total number of measurements in ADAPT-VQE scales as $\mathcal{O}(\tilde{N}_{\text{opt}} \times N_{\text{iter}} + N_{\text{term}} \times N_{\text{iter}})$.

In ADAPT-GCM, there are also two parts of measurements: the gradient measurements and the measurements for constructing the projected Hamiltonian $\mathbf{H}$ and overlap matrices $\mathbf{S}$ at each iteration. The former is the same as in ADAPT-VQE. The number of measurements for constructing the entire $\mathbf{H}$ and $\mathbf{S}$ matrices scales as $\mathcal{O}(N_{\text{GF}}^2)$, with N_{GF} being the number of generating functions. However, at each iteration, since only new matrix elements associated with new generating functions are required, the actual number of measurements only scales as $\mathcal{O}(N_{\text{GF}})$ if the number of the generating functions grows linearly with the number of iterations, as studied in this paper. Since $N_{\text{GF}} \sim \mathcal{O}(N_{\text{iter}})$, the measurement comparison between the two ADAPT approaches then boils down to $\mathcal{O}(\tilde{N}_{\text{opt}} \times N_{\text{iter(VQE)}})$ vs. $\mathcal{O}(N_{\text{iter(GCM)}}^2)$.

The above discussion is within the fault-tolerance context. In real quantum measurements, noise from various sources can be significant. For solving the generalized eigenvalue problem, the uncertainty in the finite shots of measurements is often a concern. In our previous study[50], assuming the measurement data to be normally distributed,

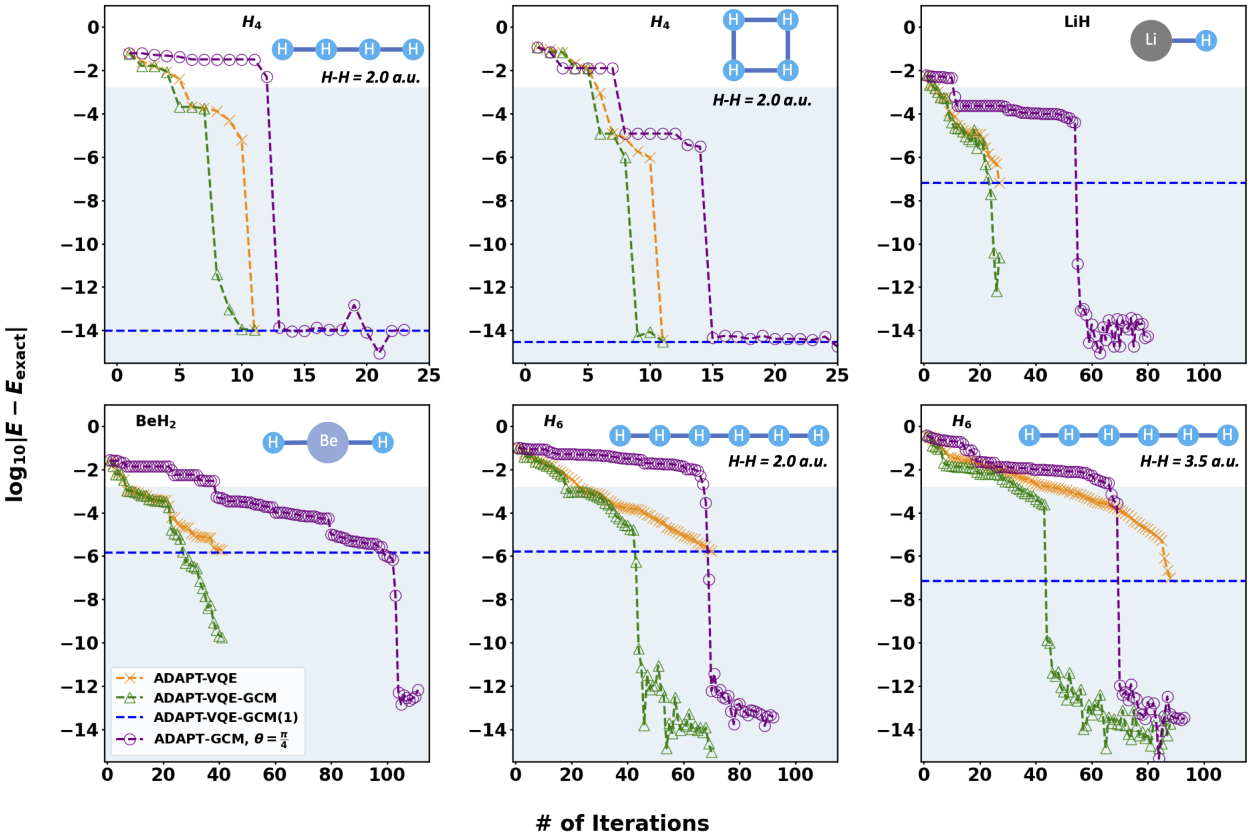

Figure 4. Convergence performance of ADAPT-VQE-GCM, ADAPT-VQE-GCM(1), and ADAPT-GCM in computing the ground states of six molecules. For comparison, results from ADAPT-VQE are also included. The shaded region is when error is smaller than the chemical accuracy. ADAPT-VQE and ADAPT-VQE-GCM are terminated when the sum of the magnitudes of all gradients falls below 10^{-4}. ADAPT-GCM is set to converge if the energy differences are under 10^{-6} a.u. for a certain number of consecutive iterations, where this number depends on the size of the operator pool.

we've shown that stronger correlation requires orders of magnitude more measurements than relatively weak correlation. Additionally, it has been shown that finite sampling error in the overlap **S** matrix seems more influential than in the **H** matrix. A more detailed discussion on this topic in the ADAPT-GCM context is given in Appendix F.

VIII. CONCLUSION AND OUTLOOK

In this study, we have demonstrated a new hybrid quantum chemical approach for accurate and efficient quantum simulations. This approach represents another subspace expansion method, employing non-orthogonal Givens rotations as bases for subspace expansion. This differs from typical constrained optimization methods such as VQE and its ADAPT variant. Compared with other subspace expansion approaches, our approach automates the basis selection. Compared with the ADAPT-VQE, our Givens rotation-based method uses the same ansätze pool but solves a generalized eigenvalue problem within a working subspace. Here, the global minimum ensures a lower bound to the ADAPT-VQE solutions. Our results indicate that the numerical performance of the new ADAPT-GCM surpasses that of the ADAPT-VQE method across several benchmark systems, particularly in instances of strong electron correlation.

Unlike other subspace expansion methods and those that feature solving generalized eigenvalue problems, the GCM-based approach is size-extensive. This is due to the fact that the bases are essentially UCC ansätze. However, since the bases are generally non-orthogonal, sampling in the overlap matrix may have a greater impact than in the projected Hamiltonian matrix. Despite this, the GCM-based approach can be seamlessly applied directly to qubit space. This is achieved by using the exponential of antihermitian Pauli strings as generating functions, thereby improving scalability for larger systems. A similar approach has been reported in the development of qubit-ADAPT-VQE as a scalable version of ADAPT-VQE[59].

The simplicity of the GCM Ansatz renders it an appealing solver in excited state computation and Hamiltonian downfolding that are the byproducts of the ADAPT-GCM procedure as indicated in Fig. 3a, laying the groundwork for sophisticated quantum simulations in chemistry. Besides, it will also be useful in designing constant-circuit-depth quantum algorithms, such as the QFlow algorithm[60]. However, utilizing the GCM solver in QFlow necessitates robust

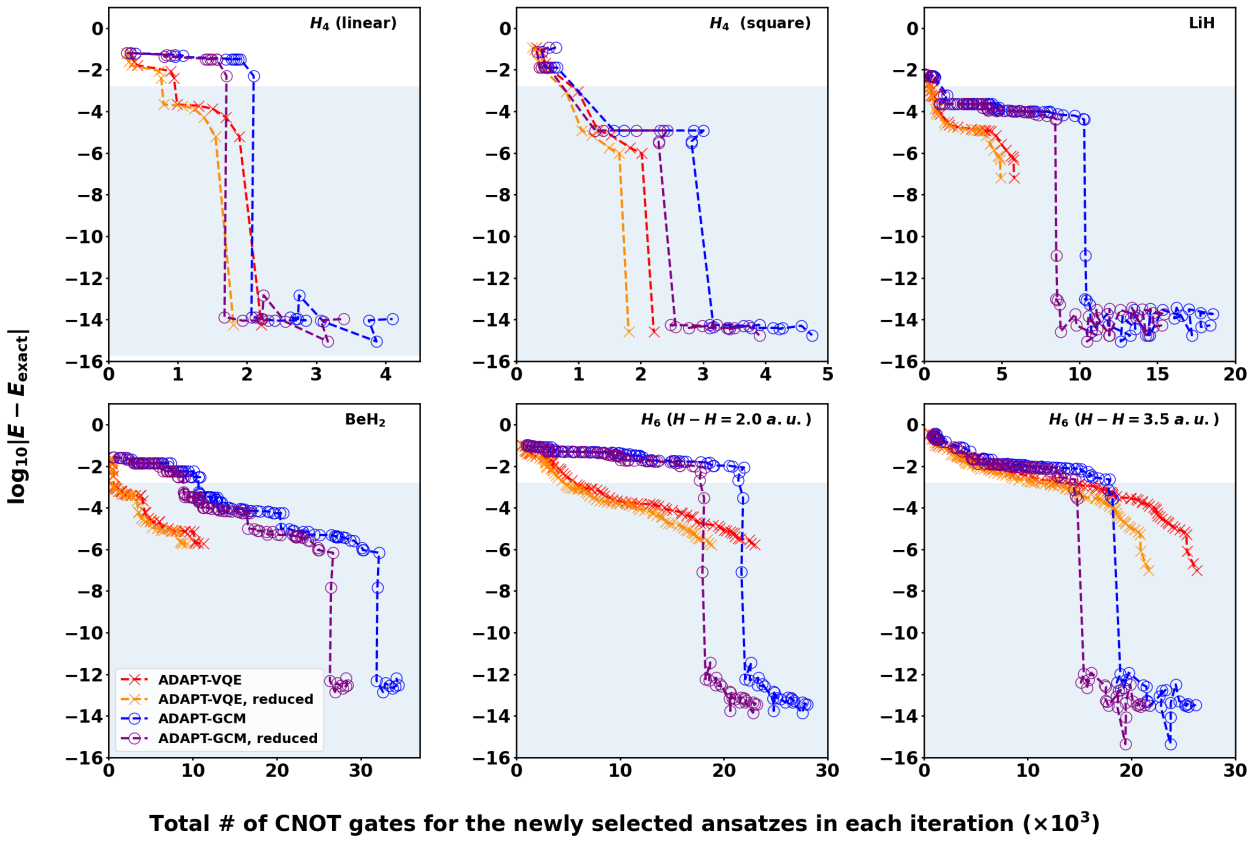

Figure 5. Energy difference as a function of the number of CNOTs in the corresponding circuits of newly selected ansätze. Note that the number of CNOTs for ADAPT-GCM at each iteration is the sum of the selected single basis and a product basis. All circuits are generated from the order-1 trottered Fermionic operators by Qiskit[58]. The reduced counts are obtained using *qiskit.transpile* with optimization level 3, which includes canceling back-to-back CNOTs, "commutative cancellation," and unitary synthesis. The implementation of Givens rotation is discussed in Appendix D.

procedures for approximating GCM by UCC-type ansätz, which can be achieved through the maximization of the overlap $|\langle\psi_{\text{GCM}}|\psi_{\text{UCC}}\rangle|$.

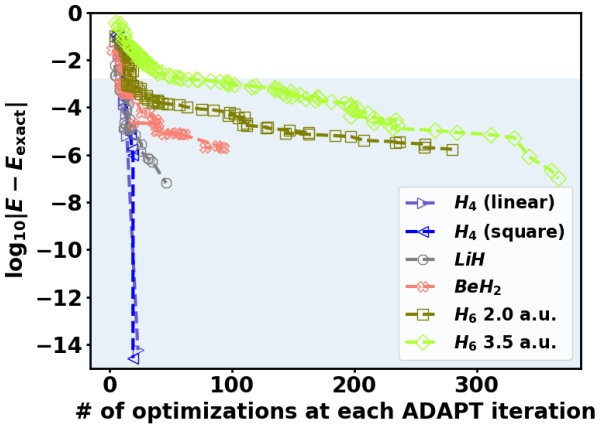

Figure 6. The performance of classical minimizer (in terms of number of optimization) at each ADAPT-VQE iteration.

IX. DATA AND CODE AVAILABILITY

The data and the code for the related numerical simulation is available in GitHub repository[61].

X. ACKNOWLEDGEMENTS

M. Z., B. P. and K. K. were supported by the "Embedding QC into Many-body Frameworks for Strongly Correlated Molecular and Materials Systems" project, which is funded by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences (BES), the Division of Chemical Sciences, Geosciences, and Biosciences under FWP 72689. A. L. was supported by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, Quantum Science Center. X. Y. was supported by National Science Foundation CAREER DMS-2143915. M. Z. and X. Y. both also were supported by Defense Advanced Research Projects Agency as part of the project W911NF2010022: *The Quantum Computing Revolution and Optimization: Challenges and Opportunities*. This research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725.

The Pacific Northwest National Laboratory is operated by Battelle for the U.S. Department of Energy under Contract DE-AC05-76RL01830.

Appendix A: Mapping optimization to a configuration subspace

Theorem 1. For a Hamiltonian H and a parametrized normalized ansätz $|\Phi(\boldsymbol{\theta})\rangle = U(\boldsymbol{\theta})|\phi\rangle$, where $|\phi\rangle$ is a fixed reference state and $U(\boldsymbol{\theta})$ is a unitary operator dependent on the parameter $\boldsymbol{\theta} = \{\theta_1, \theta_2, \cdots \theta_P\}$, the minimal expectation value

$$\varepsilon = \min_{\boldsymbol{\theta}} \langle\phi|U^\dagger(\boldsymbol{\theta})|H|U(\boldsymbol{\theta})|\phi\rangle \tag{A1}$$

is the upper bound of the minimal eigenvalue of H in an arbitrary linearly independent M-configuration subspace $\{\psi_1, \psi_2, \cdots, \psi_M\}$ ($M \geq P$) where $U(\boldsymbol{\theta})|\phi\rangle$ can be completely projected out, i.e.,

$$U(\boldsymbol{\theta})|\phi\rangle = \sum_{i=1}^{M} c_i(\boldsymbol{\theta})|\psi_i\rangle., \tag{A2}$$

where

$$c_i(\boldsymbol{\theta}) = \sum_{j=1}^{M} (\mathbf{S}^{-1})_{ij} \langle\psi_j|U(\boldsymbol{\theta})|\phi\rangle \tag{A3}$$

and $\mathbf{S}$ is the the overlap matrix with

$$\mathbf{S}_{ij} = \langle\psi_i|\psi_j\rangle. \tag{A4}$$

Proof : Given (A2), $U(\boldsymbol{\theta})|\phi\rangle$ is one state in the linearly independent M-configuration subspace, thus the minmal expectation value from minimizing (A1) is constrained by the form of $U(\boldsymbol{\theta})$, and does not necessarily give the minimal expectation value of H in the subspace. Instead, searching the minimal expectation value of H in the subspace can be re-formulated to minimizing a generalized Rayleigh quotient for the projected Hamiltonian matrix $\mathbf{H}$ and the overlap matrix $\mathbf{S}$

$$\min_f \; R(f) = \frac{f^*\mathbf{H}f}{f^*\mathbf{S}f}, \quad \text{subjec to} \;\; f^*\mathbf{S}f = 1, \tag{A5}$$

with matrix elements defined as

$$\mathbf{H}_{ij} = \langle\psi_i|H|\psi_j\rangle, \tag{A6}$$
$$\mathbf{S}_{ij} = \langle\psi_i|\psi_j\rangle. \tag{A7}$$

By introducing a Lagrange multiplier and forming the Lagrangian,

$$L(f, \varepsilon) = f^*\mathbf{H}f - \varepsilon(f^*\mathbf{S}f - 1), \tag{A8}$$

the minimization corresponds to taking the derivative of $L(f, \varepsilon)$ with respect to f and setting it to zero, leading to a generalized eigenvalue equation:

$$\mathbf{H}f = \varepsilon\mathbf{S}f, \tag{A9}$$

where the lowest eigenvalue corresponds to the global minimum in the M-configuration subspace.

Lemma 1 Consider a unitary operation U parametrized by a single scalar θ in a variational ansätze. If, for any value of θ, the action of U on $|\phi\rangle$ produces at most two distinct configurations, then the task of minimizing $\langle\phi|U^\dagger(\theta)|H|U(\theta)|\phi\rangle$ can be equivalently approached by solving the following 2×2 generalized eigenvalue problem:

$$\begin{pmatrix} \langle\phi|U^\dagger(\theta_1)HU(\theta_1)|\phi\rangle & \langle\phi|U^\dagger(\theta_1)HU(\theta_2)|\phi\rangle \\ \langle\phi|U^\dagger(\theta_2)HU(\theta_1)|\phi\rangle & \langle\phi|U^\dagger(\theta_2)HU(\theta_2)|\phi\rangle \end{pmatrix} f = \varepsilon \begin{pmatrix} 1 & \langle\phi|U^\dagger(\theta_1)U(\theta_2)|\phi\rangle \\ \langle\phi|U^\dagger(\theta_2)U(\theta_1)|\phi\rangle & 1 \end{pmatrix} f. \tag{A10}$$

Here θ_1 and θ_2 are chosen such that $\|\langle\phi|U^\dagger(\theta_1)U(\theta_2)|\phi\rangle\|^2 \neq 1$. The lowest eigenvalue of the 2×2 problem yields the minimal energy within the defined subspace.

Proof : For any θ, let's express

$$U(\theta)|\phi\rangle = a(\theta)|\phi_1\rangle + b(\theta)|\phi_2\rangle \tag{A11}$$

where $|\phi_1\rangle$ and $|\phi_2\rangle$ are orthogonal, i.e., $\langle\phi_1|\phi_2\rangle = 0$, and $a(\theta)$ and $b(\theta)$ are scalar functions of θ with $|a(\theta)|^2 + |b(\theta)|^2 = 1$

According to **Theorem 1**, the minimization of $\langle\phi|U^\dagger(\theta)|H|U(\theta)|\phi\rangle$ over all θ can be recast into a 2×2 eigenvalue problem with the Hamiltonian matrix

$$\mathbf{H} = \begin{pmatrix} \langle\phi_1|H|\phi_1\rangle & \langle\phi_1|H|\phi_2\rangle \\ \langle\phi_2|H|\phi_1\rangle & \langle\phi_2|H|\phi_2\rangle \end{pmatrix}. \tag{A12}$$

Alternatively, the expansion of $U(\theta)|\phi\rangle$ can also be expanded by non-orthogonal basis, e.g.

$$U(\theta)|\phi\rangle = c_1 U(\theta_1)|\phi\rangle + c_2 U(\theta_2)|\phi\rangle, \tag{A13}$$

where θ_1 and θ_2 are two scalars satisfying $\|\langle\phi|U^\dagger(\theta_1)U(\theta_2)|\phi\rangle\|^2 \neq 1$, and c_1 and c_2 depends on the choice of $\theta, \theta_1, \theta_2$ and can be obtained through

$$(a(\theta_1)b(\theta_2) - b(\theta_1)a(\theta_2))c_1 = a(\theta)b(\theta_2) - b(\theta)a(\theta_2), \tag{A14}$$
$$(a(\theta_2)b(\theta_1) - b(\theta_2)a(\theta_1))c_2 = a(\theta)b(\theta_1) - b(\theta)a(\theta_1). \tag{A15}$$

If the expansion is done in such a non-orthogonal subspace, the minimization of $\langle\phi|U^\dagger(\theta)|H|U(\theta)|\phi\rangle$ transforms into a generalized eigenvalue problem (A9) with matrices defined as

$$\mathbf{H} = \begin{pmatrix} \langle\phi|U^\dagger(\theta_1)HU(\theta_1)|\phi\rangle & \langle\phi|U^\dagger(\theta_1)HU(\theta_2)|\phi\rangle \\ \langle\phi|U^\dagger(\theta_2)HU(\theta_1)|\phi\rangle & \langle\phi|U^\dagger(\theta_2)HU(\theta_2)|\phi\rangle \end{pmatrix}. \tag{A16}$$

and

$$\mathbf{S} = \begin{pmatrix} 1 & \langle\phi|U^\dagger(\theta_1)U(\theta_2)|\phi\rangle \\ \langle\phi|U^\dagger(\theta_2)U(\theta_1)|\phi\rangle & 1 \end{pmatrix}. \tag{A17}$$

Solving this generalized eigenvalue problem allows us to determine the combination of $U(\theta_1)|\phi\rangle$ and $U(\theta_2)|\phi\rangle$ that corresponds to the lowest expectation value for H in this two-configuration subspace.

Appendix B: Energy gradient in ADAPT-GCM framework

In the GCM framework, the k-th eigenvalue of the projected H, ε_k, can be represented as

$$\varepsilon_k = \frac{f_k^T \mathbf{H} f_k}{f_k^T \mathbf{S} f_k}, \tag{B1}$$

where f_k is the corresponding normalized eigenvectors, $f_k^T f_k = 1$. Assume the working subspace comprises M generating functions, $\{|\psi_i\rangle\}_{1\le i\le M}$, in which $M-1$ are generated by acting one Givens rotation on the reference, and only one is generated by acting the product form of Givens rotation on the reference, i.e.

$$|\psi_i\rangle = \begin{cases} G_{p_i,q_i}(\theta_i)|\phi_0\rangle, & 1 \le i \le M-1 \\ \prod_{j=1}^{M-1} G_{p_j,q_j}(\theta_j)|\phi_0\rangle, & i = M \end{cases}. \tag{B2}$$

Then $\mathbf{H}$ and $\mathbf{S}$ are $M\times M$ matrices projected in this working subspace with their matrix elements being defined as

$$\mathbf{H}_{ij} = \langle\psi_i|H|\psi_j\rangle, \tag{B3}$$
$$\mathbf{S}_{ij} = \langle\psi_i|\psi_j\rangle, \tag{B4}$$

and the derivative of the eigenvalue ε_k with respect to the rotation θ_s can be computed through

$$\frac{\partial\varepsilon_k}{\partial\theta_s} = \frac{\langle\frac{\partial\mathbf{H}}{\partial\theta_s}\rangle_k \cdot \langle\mathbf{S}\rangle_k - \langle\mathbf{H}\rangle_k \cdot \langle\frac{\partial\mathbf{S}}{\partial\theta_s}\rangle_k}{\langle\mathbf{S}\rangle_k^2}, \tag{B5}$$

with $\langle\mathbf{O}\rangle_k = f_k^T\mathbf{O}f_k$ for matrix $\mathbf{O}$. The matrix elements of $\frac{\partial\mathbf{S}}{\partial\theta_s}$ and $\frac{\partial\mathbf{H}}{\partial\theta_s}$ are the following,

- if neither $|\psi_i\rangle$ nor $|\psi_j\rangle$ is parametrized by θ_s

$$\frac{\partial(\mathbf{H}_{ij})}{\partial\theta_s} = \frac{\partial(\mathbf{S}_{ij})}{\partial\theta_s} = 0, \tag{B6}$$

- if only $|\psi_i\rangle$ is parametrized by θ_s and $i\neq j$

$$\frac{\partial(\mathbf{H}_{ij})}{\partial\theta_s} = \langle\psi_i|(a_{q_s}^\dagger a_{p_s} - a_{p_s}^\dagger a_{q_s})H|\psi_j\rangle, \tag{B7}$$
$$\frac{\partial(\mathbf{S}_{ij})}{\partial\theta_s} = \langle\psi_i|(a_{q_s}^\dagger a_{p_s} - a_{p_s}^\dagger a_{q_s})|\psi_j\rangle, \tag{B8}$$

- if only $|\psi_j\rangle$ is parametrized by θ_s and $i\neq j$

$$\frac{\partial(\mathbf{H}_{ij})}{\partial\theta_s} = \langle\psi_i|H(a_{p_s}^\dagger a_{q_s} - a_{q_s}^\dagger a_{p_s})|\psi_j\rangle, \tag{B9}$$
$$\frac{\partial(\mathbf{S}_{ij})}{\partial\theta_s} = \langle\psi_i|(a_{p_s}^\dagger a_{q_s} - a_{q_s}^\dagger a_{p_s})|\psi_j\rangle, \tag{B10}$$

- if both $|\psi_j\rangle$ are parametrized by θ_s

$$\frac{\partial(\mathbf{H}_{ij})}{\partial\theta_s} = \langle\psi_i|[H,(a_{p_s}^\dagger a_{q_s} - a_{q_s}^\dagger a_{p_s})]|\psi_j\rangle, \tag{B11}$$
$$\frac{\partial(\mathbf{S}_{ij})}{\partial\theta_s} = 0. \tag{B12}$$

As can be seen matrices $\frac{\partial\mathbf{S}}{\partial\theta_s}$ and $\frac{\partial\mathbf{H}}{\partial\theta_s}$ are very sparse with the non-zero elements only residing in the column and row whose index is equal to p, which can greatly reduce the computational complexity. The gradient employed in the proposed ADAPT-GCM in Fig. 3 can be considered as an approximation to the above gradient calculations, where only (B11) and (B12) are taken into consideration.

Appendix C: ADAPT-VQE-GCM and ADAPT-VQE-GCM(1) algorithms

The main structure of ADAPT-VQE-GCM in Alg. 1 is built upon the a complete ADAPT-VQE process from Ref. 56. In fact, the essential use of ADAPT-VQE is to provide the unitary operators and phase values (except the phase for the latest selected ansätze) to form the bases for GCM. ADAPT-VQE-GCM(1) only includes the selected single Givens rotations from all previous ADAPT iterations and the final product of the Givens rotations when ADAPT-VQE converges in the basis set.

Algorithm 1 ADAPT-VQE-GCM

Require: Operator pool $\{G_l\}_{(l)\in L}$ with L the orbital index pool, initial state $|\psi_{(0)}^{VQE}\rangle := |\phi_{HF}\rangle$, VQE optimal parameter set $\Theta_{(0)}^* = \emptyset$, GCM basis set $\mathbf{B} = \emptyset$, and Hamiltonian H

$n \leftarrow 0$
while convergence criteria is not met **do**
  $n \leftarrow n+1$
  Select operator:
    $G_{(n)} \leftarrow \max_{l\in L}\left\{\frac{\partial\varepsilon}{\partial\theta_l}\big|_{\theta_l=0} = \langle\psi_{(n-1)}^{VQE}|[H,A_l]|\psi_{(n-1)}^{VQE}\rangle\right\}$;
  Prepare VQE ansätz and optimize free parameters:
    $|\tilde{\psi}_{(n)}^{VQE}(\Theta_{(n)})\rangle = \prod_{i=n}^{1} G_{(i)}(\theta_i)|\phi_{HF}\rangle$;
    $\Theta_{(n)}^* = \mathrm{argmin}_{\Theta_{(n)}}\langle\tilde{\psi}_{(n)}^{VQE}(\Theta_{(n)})|H|\tilde{\psi}_{(n)}^{VQE}(\Theta_{(n)})\rangle$;
    $|\psi_{(n)}^{VQE}\rangle \leftarrow |\tilde{\psi}_{(n)}^{VQE}(\Theta_{(n)}^*)\rangle$;
  Expand the basis set B:
    $\mathbf{B} \leftarrow G_{(n)}(\theta_n^*)|\phi_{HF}\rangle$;
    $\mathbf{B} \leftarrow |\psi_{(n)}^{VQE}\rangle$;
  Compute matrices and solve for eigenvalues:
    $\mathbf{H} = \mathbf{B}^T H\mathbf{B}$;
    $\mathbf{S} = \mathbf{B}^T\mathbf{B}$;
    $\mathbf{H}f = \varepsilon\mathbf{S}f$;
end while
return eigenvalues ε after convergence

Appendix D: Implementing the Givens rotations

For skew-Hermitian operators

$$A_{p_i,q_i} = a^\dagger_{p_i} a_{q_i} - a^\dagger_{q_i} a_{p_i}, \quad p_i < q_i \tag{D1}$$

$$A_{p_i,r_i,q_i,s_i} = a^\dagger_{p_i} a^\dagger_{r_i} a_{q_i} a_{s_i} - a^\dagger_{s_i} a^\dagger_{q_i} a_{r_i} a_{p_i}, \quad q_i < s_i < p_i < r_i \tag{D2}$$

the number of CNOT gates in their corresponding Givens rotation circuits can be estimated as follows. For single excitations, Eq. (D1), its corresponding Givens rotation circuit for two adjacent indices is given in Fig. 1, i.e., two CNOT gates. If the two indices are not adjacent, $q_i - p_i - 1$ FSWAP gates will be employed. If each FSWAP contains three CNOT gates,[62] the total number of CNOT gates for implementing a single excitation Givens rotation would be $3(q_i - p_i) - 1$. For double excitations, Eq. (D2), the typical implementation involves the Trotterization of the exponential of the Pauli representation of the second quantized expression. For example, through Jordan-Wigner transformation,

$$a_{p_i} = \frac{1}{2}(X_{p_i} + \mathrm{i}Y_{p_i}) \prod_{k=0}^{p_i-1} Z_k, \quad a^\dagger_{p_i} = \frac{1}{2}(X_{p_i} - \mathrm{i}Y_{p_i}) \prod_{k=0}^{p_i-1} Z_k, \tag{D3}$$

the exponential of Eq. (D2) can be expressed as

$$\begin{aligned} \exp(\theta A_{p_i,r_i,q_i,s_i}) = \exp\Bigg(-\mathrm{i}\frac{\theta}{8} (&X_{q_i} Y_{s_i} X_{p_i} X_{r_i} + Y_{q_i} X_{s_i} X_{p_i} X_{r_i} \\ &+ Y_{q_i} Y_{s_i} Y_{p_i} X_{r_i} + Y_{q_i} Y_{s_i} X_{p_i} Y_{r_i} - X_{q_i} X_{s_i} Y_{p_i} X_{r_i} - X_{q_i} X_{s_i} X_{p_i} Y_{r_i} \\ &- Y_{q_i} X_{s_i} Y_{p_i} Y_{r_i} - X_{q_i} Y_{s_i} Y_{p_i} Y_{r_i}) \prod_{k=p_i+1}^{r_i-1} Z_k \prod_{l=q_i+1}^{s_i-1} Z_l \Bigg), \end{aligned} \tag{D4}$$

whose first Trotter approximation can be implemented using 48 CNOT gate if p_i, r_i, q_i, s_i are adjacent to each other.[63] If not, longer CNOT ladders will be employed to connect mapped qubits. For all the singlet operators in the spin-adapted pool in Section VI, under the Jordan-Wigner transformation, the average number of CNOTs gates of an individual operator is

$$\bar{N}_{Pauli\ per\ string} \times \bar{N}_{string} \times 2 \tag{D5}$$

where $\bar{N}_{Pauli\ per\ string}$ is the average number of non-identity Pauli gates in all strings, $\bar{N}_{Paulistring}$ is the average number of Pauli strings in an operator after mapping, and the multiplier 2 comes from the symmetries of the CNOT ladders.

Appendix E: Other computational details

The implementation of ADAPT-VQE is the original code in Ref. 56 from the corresponding repository[64], and its computation is based on the SciPy linear algebra package[65]. To have the same foundation for comparison, calculations in Section VI for three GCM related algorithms follow the same manner. However, the quantum resources estimation, shown in Section VII, utilizes Qiskit[58]. All the numerical experiments are carried on a MacBook Pro with 2.6 GHz 6-Core Intel Core i7 CPU and Intel UHD Graphics 630 GPU.

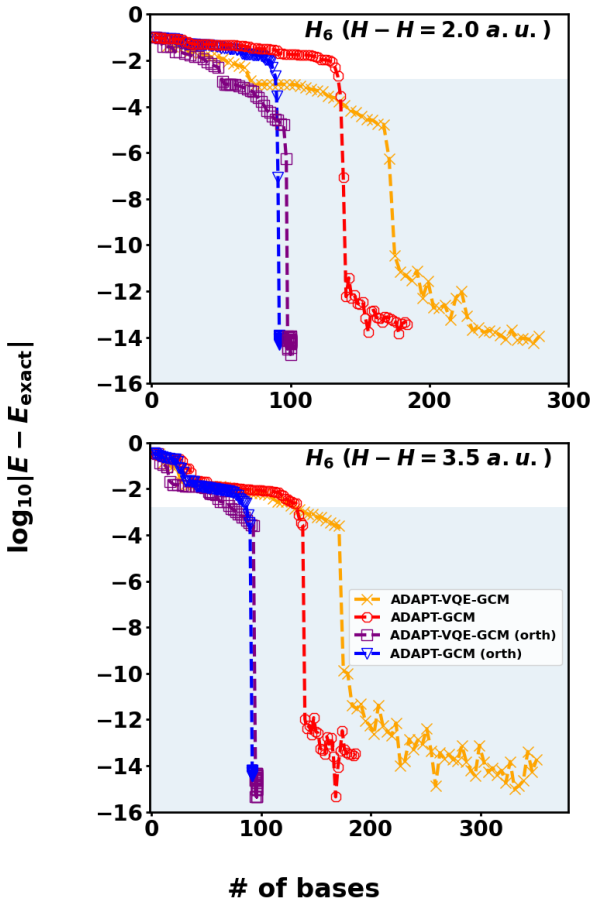

Figure 7. Number of bases used in each iteration before and after orthgonalization in ADAPT-VQE-GCM and ADAPT-VQE in both H_6 examples. The "orth" in the labels mean the calculation was done with the "orthogonalized bases."

As mentioned in Section VI, some small disturbances are used to mitigate numerical instability in the generalized eigenvalue problems of ADAPT-GCM. This problem can be resolved with orthgonalization of basis vectors, while the current quantum implementation requires special hardware like QRAM[57], which is beyond the scope of this paper. To illustrate the effects, numerical simulations with classical orthgonalization on basis vectors in both H_6 examples is conducted and demonstrated in Fig. 7. For ADAPT-VQE-GCM, we increase four basis vectors in each iteration by setting $\theta_k := \pm\pi/4$ for the vectors mentioned in Section V, and for ADAPT-GCM, $\theta_k := \pi/4$ is sufficient. So orthogonalized ADAPT-VQ-GCM converges slightly faster at the beginning in both plots in Figure 7. However, at the final convergence, the number of orthogonal bases are both at around 100.

Appendix F: Randomness due to finite number of shots

Recall that in both ADAPT-GCM and GCM part of the ADAPT-VQE-GCM, we compute project Hamiltonian, $\mathbf{H}$, and overlap matrix, $\mathbf{S}$, entry by entry

$$\mathbf{H}_{ij} = \langle \psi_i | H | \psi_j \rangle = \langle \psi_{HF} | B_i^\dagger H B_j | \psi_{HF} \rangle \tag{F1}$$

$$\mathbf{S}_{ij} = \langle \psi_i | \psi_j \rangle = \langle \psi_{HF} | B_i^\dagger B_j | \psi_{HF} \rangle, \tag{F2}$$

where B_i and B_j are short-hand notations for some products of Givens rotation matrices for simplicity. These expectations can be computed by converting $B_i^\dagger H B_j$ and $B_i^\dagger B_j$ to the sums of Pauli strings, respectively, then grouping the Pauli strings to do simultaneous measurements. The quantum and classical complexities and effects of shot noise have been explicitly discussed in Ref. 50.

Instead, in this section, we illustrate the shot noise calculation when employing Hadamard test for the expectation computation. Due to the randomness of the collapse of qubits, we will transfer $\mathbf{H}_{ij}$ to a random variable where the number of shots controls its variance.

Assume the Pauli decomposition of Hamiltonian $H = \sum_{k=1}^{N_{term}} c_k P_k$, $c_k \in \mathbb{R}$ for all k, has been provided, we have

$$\mathbf{H}_{ij} = \sum_{k=1}^{N_{term}} c_k \langle \psi_{HF} | B_i^\dagger P_k B_j | \psi_{HF} \rangle. \tag{F3}$$

Without the loss of the generality, we only focus on $\mathbf{H}_{ij}$ since the calculation of $\mathbf{S}_{ij}$ can be regarded as a special case for (F3) when $N_{term} = 1$, $c_k = 1$, and P_k is an identity matrix. It is known that, Hadamard test gives

$$\Pr(\text{ancillary is } |0\rangle) = \frac{1}{2} + \frac{1}{2}\mathrm{Re}\langle \psi_{HF} | B_i^\dagger P_k B_j | \psi_{HF} \rangle \tag{F4}$$

$$\Pr(\text{ancillary is } |1\rangle) = \frac{1}{2} - \frac{1}{2}\mathrm{Re}\langle \psi_{HF} | B_i^\dagger P_k B_j | \psi_{HF} \rangle, \tag{F5}$$

after measuring the ancillary qubit in the circuit. Thus, a binary random variable can be defined by

$$\Gamma_l^{(k)} = \begin{cases} 1 & \text{if ancillary is } |0\rangle \text{ in } l^{th} \text{ shot when compute } \langle B_i^\dagger P_k B_j \rangle \\ -1 & \text{if ancillary is } |1\rangle \text{ in } l^{th} \text{ shot when compute } \langle B_i^\dagger P_k B_j \rangle \end{cases} \tag{F6}$$

with

$$p_k := \mathbb{E}\left[\Gamma_l^{(k)}\right] = \mathrm{Re}\langle \psi_{HF} | B_i^\dagger P_k B_j | \psi_{HF} \rangle, \tag{F7}$$

$$\mathrm{Var}\left[\Gamma_l^{(k)}\right] = \mathbb{E}\left[\left(\Gamma_l^{(k)}\right)^2\right] - \left(\mathbb{E}\left[\Gamma_l^{(k)}\right]\right)^2 = 1 - p_k^2, \tag{F8}$$

for all l by assuming $\Gamma_l^{(k)}$'s are mutually independent and identically distributed. For simplicity, we omit the indices i and j for now. In other words, each measurement on the ancillary qubit in the Hadamard test circuit is modelled by sampling the random variable $\Gamma_l^{(k)}$.

We further construct another random variable, omitting indices i and j again,

$$\Lambda^{(k)} = \sum_{i=l}^{N_{shot}^{(k)}} \Gamma_l^{(k)} \tag{F9}$$

where $N_{shot}^{(k)}$ is the number of shots for the Hadamard test circuit of $\langle B_i^\dagger P_k B_j \rangle$, and it is considered as a user-defined constant. Then, it follows that

$$\frac{1}{2}\left(\Lambda^{(k)} + N_{shot}^{(k)}\right) \sim \mathrm{Bin}\left(N_{shot}^{(k)}, \frac{1}{2} + \frac{1}{2}p_k\right), \tag{F10}$$

by realizing $\frac{1}{2} + \frac{1}{2}\Gamma_i^{(k)}$ is a standard Bernoulli random variable. Further combining (F3) and (F10), another discrete random variable

$$\Xi^{(ij)} = \sum_{k=1}^{N_{term}} \frac{c_k}{N_{shot}^{(k)}} \Lambda^{(k)} = \sum_{k=1}^{N_{term}} \frac{c_k}{N_{shot}^{(k)}} \sum_{l=1}^{N_{shot}^{(k)}} \Gamma_l^{(k)} \tag{F11}$$

has

$$\mathbb{E}\left[\Xi^{(ij)}\right] = \sum_{k=1}^{N_{term}} \frac{c_k}{N_{shot}^{(k)}} N_{shot}^{(k)} p_k = \mathbf{H}_{ij} \tag{F12}$$

$$\begin{aligned} \mathrm{Var}\left[\Xi^{(ij)}\right] &= \sum_{k=1}^{N_{term}} \left(\frac{c_k}{N_{shot}^{(k)}}\right)^2 N_{shot}^{(k)} \mathrm{Var}\left[\Gamma_l^{(k)}\right] \\ &= \sum_{k=1}^{N_{term}} \frac{c_k^2}{N_{shot}^{(k)}} \left(1 - p_k^2\right). \end{aligned} \tag{F13}$$

In this case, $\mathbf{H}_{ij}$ has a corresponding random variable $\Xi^{(ij)}$ where the randomness is caused by the uncertainty from the collapse of the qubits during the measurement. The role of the number of shots, as (F13) implies, is to govern the scale of the variance. The variance (F13) goes to 0 when every $N_{shot}^{(k)}$ goes to infinity.

Chebyshev's inequality shows that, for every $a > 0$,

$$\Pr\left(\left|\Xi^{(ij)} - \mathbf{H}_{ij}\right| \geq a\right) \leq \frac{\mathrm{Var}\left[\Xi^{(ij)}\right]}{a^2} =: \eta_{ij}. \tag{F14}$$

Setting $N_{shot}^{(k)}$ to N_{shot} for all i, j, k, Table I provides the relation between N_{shot} and $\max_{i,j} \eta_{ij}$ when a is fixed at 10^{-4}. However, the tightness of the bound is still unclear for this discrete variable, and this is not sufficient to illustrate the effects of the disturbance of the matrix entries on the eigenvalues. This leads to numerical simulations.

Since it is very common to have huge $N_{shot}^{(k)}$ in practice, we can assume $\Lambda^{(k)}$ is closely approximated by

$$\Lambda_{gauss}^{(k)} \sim \mathrm{N}\left(N_{shot}^{(k)} p_k,\, N_{shot}^{(k)}\left(1 - p_k^2\right)\right) \tag{F15}$$

and mutually independent for different k. Thus, $\Xi^{(ij)}$ is also approximated by a Gaussian random variable $\Xi_{gauss}^{(ij)}$

Table I. N_{shot} for the various choices of $\max_{i,j} \eta_{ij}$ when $a = 10^{-4}$.

Molecule	$\max \eta_{ij} = 5\%$	$\max \eta_{ij} = 1\%$
$H4$ (square)	6.32×10^{10}	3.16×10^{11}
LiH	1.34×10^{11}	6.68×10^{11}
H_6 (2.0 *a.u.*)	1.26×10^{11}	6.29×10^{11}
H_6 (3.5 *a.u.*)	6.17×10^{10}	3.08×10^{11}

Table II. Reduction by IS on the total number of shots for **H** when $N_{shot}^{(k)} = \tau = \tau_{simu}$ regardless the change of accuracy.

Molecule	N_{term}	$\sum_k \lvert c_k \rvert$	% reduction by IS
$H4$ (square)	529	43.4	91.8%
LiH	1861	65.8	96.5%
H_6 (2.0 $a.u.$)	2629	115.0	95.6%
H_6 (3.5 $a.u.$)	2629	99.2	96.2%

with the same mean and variance as in (F12) and (F13), respectively.

The idea of importance sampling (IS) can be applied here to significantly reduce the number of shots. By setting

$$N_{shot}^{(k)} := |c_k|\tau \tag{F16}$$

for some large constant τ, the random variable

$$\Xi_{IS}^{(ij)} = \frac{1}{\tau}\sum_{k=1}^{N_{term}} \mathrm{sign}(c_k)\Lambda_{gauss}^{(k)} \tag{F17}$$

is also Gaussian and has the mean $\mathbf{H}_{ij}$ and the variance

$$\mathrm{Var}\left[\Xi_{IS}^{(ij)}\right] = \frac{1}{\tau}\sum_{k=1}^{N_{term}} |c_k|\left(1-p_k^2\right). \tag{F18}$$

It is suggested to use larger N_{shot}^{k} and τ for the overlap matrix **S** than **H** to improve the numerical stability. In the following simulations, **S** will always have one hundred times larger N_{shot}^{k} and τ than **H**.

The Monte Carlo simulations for squared H_4, LiH, and both H_6 are conducted by treating each matrix entry as a normal random variable follows (F12) and (F13). We select a pair of **H** and **S** from a specific iteration of ADAPT-GCM of each molecule such that the error of ground-energy is around 10^{-11} to 10^{-14} a.u. without the finite-sampling noise.

However, the initial simulations show that the shot noise severely aggravates the numerical instability in the eigenvalue problems. This makes us impose another strategy when solve the eigenvalue problems. Besides adding 10^{-10} on the diagonal entries of every overlap matrix, we also discard eigenvalues that are smaller than an lower bound based on the HF energy, $\lfloor E_{\mathrm{HF}} \rfloor - 1$. In practice, this bound can be modified from the energy in the previous iteration.

As summarized in Fig. 8, without IS, H_4 and LiH require $N_{shot}^{(k)} = \tau_{simu}$ to be about 10^7 to reach the chemical accuracy stably while both H_6 instances need the number to be around $10^{11.5}$ and $10^{12.5}$, respectively. When IS is applied, setting $\tau = \tau_{simu}$, H_6 (3.5 $a.u.$) gives the largest accuracy drop, but as Table II shows, the reduction by IS on the total number of shots for **H** when $N_{shot}^{k} = \tau = \tau_{simu}$ is around 96.2% for H_6 (3.5 $a.u.$). So importance sampling is still effectively reduce the number of shots in all four molecules.

REFERENCES

[1] M. Imada, A. Fujimori, and Y. Tokura, Rev. Mod. Phys. **70**, 1039 (1998).

[2] R. J. Witzke, D. Hait, K. Chakarawet, M. Head-Gordon, and T. D. Tilley, ACS Catalysis **10**, 7800 (2020).

[3] C. A. Gould, K. R. McClain, D. Reta, J. G. C. Kragskow, D. A. Marchiori, E. Lachman, E.-S. Choi, J. G. Analytis, R. D. Britt, N. F. Chilton, B. G. Harvey, and J. R. Long, Science **375**, 198 (2022).

[4] M. Askerka, G. W. Brudvig, and V. S. Batista, Accounts of Chemical Research **50**, 41 (2017).

[5] J. Raymond and R. E. Blankenship, Coordination Chemistry Reviews **252**, 377 (2008), the Role of Manganese in Photosystem II.

[6] I. Shavitt and R. Bartlett, *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*, Cambridge Molecular Science (Cambridge University Press, 2009).

[7] A. Peruzzo, J. R. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'brien, Nat. Commun. **5**, 4213 (2014).

[8] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, New J. Phys. **18**, 023023 (2016).

[9] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik, Quantum Sci. Technol. **4**, 014008 (2018).

[10] Y. Shen, X. Zhang, S. Zhang, J.-N. Zhang, M.-H. Yung, and K. Kim, Phys. Rev. A **95**, 020501 (2017).

[11] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Nature **549**, 242 (2017).

[12] A. Kandala, K. Temme, A. D. Corcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, Nature **567**, 491 (2019).

[13] J. I. Colless, V. V. Ramasesh, D. Dahlen, M. S. Blok, M. E. Kimchi-Schwartz, J. R. McClean, J. Carter, W. A. de Jong, and I. Siddiqi, Phys. Rev. X **8**, 011021 (2018).

[14] W. J. Huggins, J. Lee, U. Baek, B. O'Gorman, and K. B. Whaley, New J. Phys. **22**, 073009 (2020).

[15] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. Sawaya, *et al.*, Chemical reviews **119**, 10856 (2019).

[16] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, Nature communications **10**, 1 (2019).

[17] H. R. Grimsley, D. Claudino, S. E. Economou, E. Barnes, and N. J. Mayhall, Journal of chemical theory and computation **16**, 1 (2019).

[18] V. Verteletskyi, T.-C. Yen, and A. F. Izmaylov, The Journal of chemical physics **152**, 124114 (2020).

[19] S. McArdle, T. Jones, S. Endo, Y. Li, S. C. Benjamin, and X. Yuan, npj Quantum Inf. **5**, 1 (2019).

[20] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Reviews of Modern Physics **92**, 015003 (2020).

[21] J. Tilly, H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. H. Booth, *et al.*, arXiv preprint arXiv:2111.05176 (2021).

[22] E. Farhi, J. Goldstone, and S. Gutmann, arXiv preprint arXiv:1411.4028 (2014).

[23] S. Stein, N. Wiebe, Y. Ding, P. Bo, K. Kowalski, N. Baker, J. Ang, and A. Li, in *Proceedings of the 49th Annual International Symposium on Computer Architecture* (2022) pp. 59–71.

[24] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S.Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik, arXiv preprint arXiv:2101.08448 (2021).

[25] T. Albash and D. A. Lidar, Rev. Mod. Phys. **90**, 015002 (2018).

[26] S. Aaronson and A. Arkhipov, in *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing*, STOC '11 (Association for Computing Machinery, New York, NY, USA, 2011) p. 333–342.

[27] A. Trabesinger, Nat. Phys. **8**, 263 (2012).

[28] I. M. Georgescu, S. Ashhab, and F. Nori, Rev. Mod. Phys. **86**, 153 (2014).

[29] M. Motta, C. Sun, A. T. K. Tan, M. J. O'Rourke, E. Ye, A. J. Minnich, F. G. S. L. Brandão, and G. K.-L. Chan, Nat. Phys. **16**, 205 (2020).

[30] R. M. Parrish and P. L. McMahon, arXiv preprint arXiv:1909.08925 (2019).

[31] O. Kyriienko, npj Quantum Inf. **6**, 1 (2020).

[32] J. R. McClean, M. E. Kimchi-Schwartz, J. Carter, and W. A. de Jong, Phys. Rev. A **95**, 042308 (2017).

[33] T. Takeshita, N. C. Rubin, Z. Jiang, E. Lee, R. Babbush, and J. R. McClean, Phys. Rev. X **10**, 011004 (2020).

[34] J. McClean, Z. Jiang, N. Rubin, R. Babbush, and H. Neven, Nat. Commun. **11**, 636 (2020).

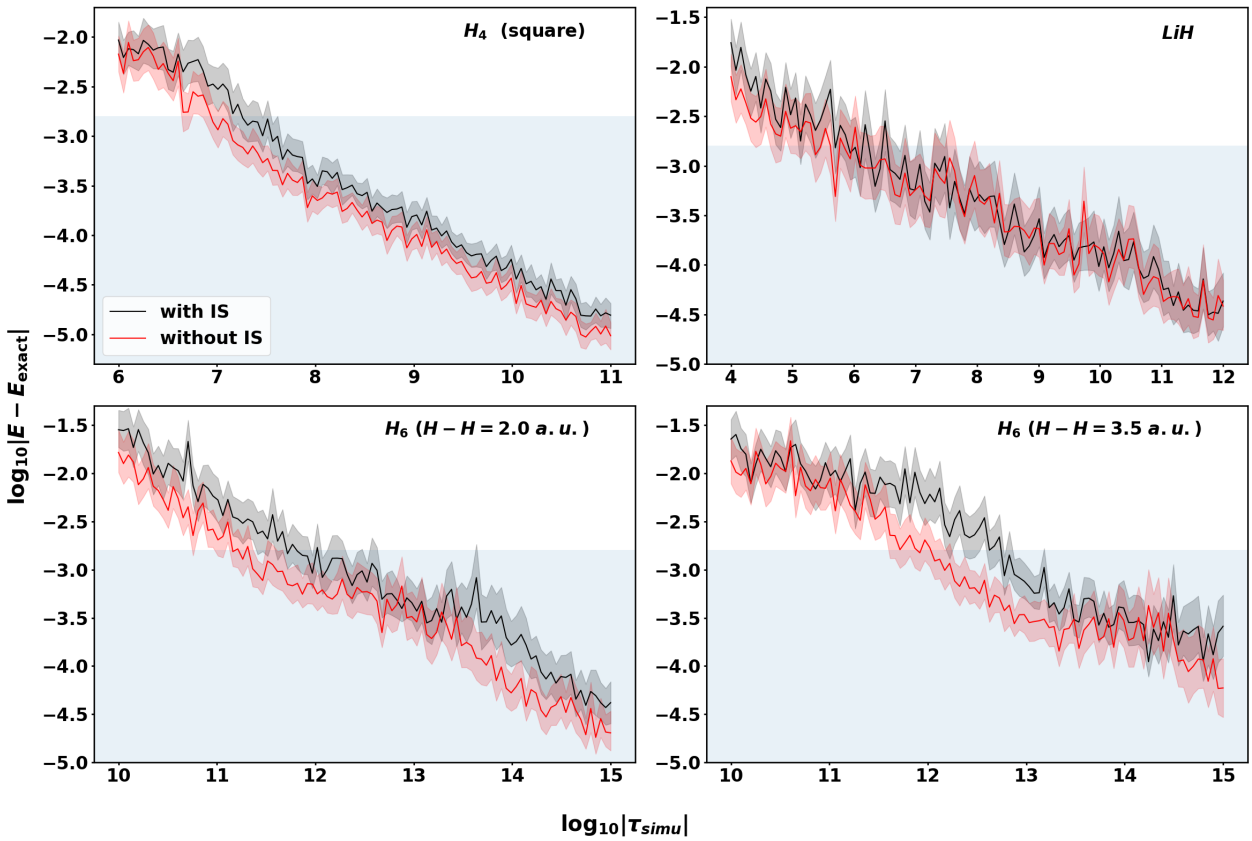

Figure 8. Monte Carlo simulations for the effects of the shot noise on the accuracy of the selected iterations in ADAPT-GCM on different molecules. The error under infinite number of shots are in 10^{-11} to 10^{-14}. The black and red-shaded area is the 95% confidence interval estimated from 100 runs for each given τ_{simu}. When not using importance sampling, $N^k_{shot} = \tau_{simu}$ for a projected Hamiltonian $\mathbf{H}$ and $N^k_{shot} = 100\tau_{simu}$ for an overlap matrix $\mathbf{S}$. With importance sampling, $N^k_{shot} = |c_k|\tau_{simu}$ for $\mathbf{H}$ and $N^k_{shot} = 100|c_k|\tau_{simu}$ for $\mathbf{S}$. Note that the total number of shots for estimating a matrix entry is different when $N^{(k)}_{shot} = \tau = \tau_{simu}$.

[35] P. Suchsland, F. Tacchino, M. H. Fischer, T. Neupert, P. K. Barkoutsos, and I. Tavernelli, Quantum **5**, 492 (2021).
[36] N. V. Tkachenko, L. Cincio, A. I. Boldyrev, S. Tretiak, P. A. Dub, and Y. Zhang, arXiv , 2204.10741 (2023).
[37] K. Seki and S. Yunoki, PRX Quantum **2**, 010333 (2021).
[38] K. Kowalski and B. Peng, J. Chem. Phys. **153**, 201102 (2020).
[39] H. J. Vallury, M. A. Jones, C. D. Hill, and L. C. L. Hollenberg, Quantum **4**, 373 (2020).
[40] J. C. Aulicino, T. Keen, and B. Peng, International Journal of Quantum Chemistry **122**, e26853 (2022).
[41] P. J. Ollitrault, A. Kandala, C.-F. Chen, P. K. Barkoutsos, A. Mezzacapo, M. Pistoia, S. Sheldon, S. Woerner, J. M. Gambetta, and I. Tavernelli, Phys. Rev. Res. **2**, 043140 (2020).
[42] D. L. Hill and J. A. Wheeler, Physical Review **89**, 1102 (1953).
[43] J. J. Griffin and J. A. Wheeler, Phys. Rev. **108**, 311 (1957).
[44] R. Rodríguez-Guzmán, J. Egido, and L. Robledo, Nuclear Physics A **709**, 201 (2002).
[45] M. Bender, P.-H. Heenen, and P.-G. Reinhard, Reviews of Modern Physics **75**, 121 (2003).
[46] P. Ring and P. Schuck, *The nuclear many-body problem* (Springer Science & Business Media, 2004).
[47] J. Yao, J. Meng, P. Ring, and D. Vretenar, Physical Review C **81**, 044311 (2010).
[48] J. L. Egido, Physica Scripta **91**, 073003 (2016).
[49] N. Hizawa, K. Hagino, and K. Yoshida, Physical Review C **103**, 034313 (2021).
[50] M. Zheng, B. Peng, N. Wiebe, A. Li, X. Yang, and K. Kowalski, Phys. Rev. Res. **5**, 023200 (2023).
[51] H. Fukutome, Progress of Theoretical Physics **65**, 809 (1981).
[52] F. A. Evangelista, G. K.-L. Chan, and G. E. Scuseria, J. Chem. Phys. **151**, 244112 (2019).
[53] J. Lee, W. J. Huggins, M. Head-Gordon, and K. B. Whaley, Journal of Chemical Theory and Computation **15**, 311 (2019), https://doi.org/10.1021/acs.jctc.8b01004.
[54] U. Baek, D. Hait, J. Shee, O. Leimkuhler, W. J. Huggins, T. F. Stetina, M. Head-Gordon, and K. B. Whaley, PRX Quantum **4**, 030307 (2023).
[55] K. Jankowski and J. Paldus, International Journal of Quantum Chemistry **18**, 1243 (1980).
[56] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, Nature Communications **10** (2019), 10.1038/s41467-019-10988-2.
[57] K. Zhang, M.-H. Hsieh, L. Liu, and D. Tao, Physical Review Research **3** (2021), 10.1103/physrevresearch.3.043095.
[58] Qiskit contributors, "Qiskit: An open-source framework for quantum computing," (2023), qiskit version 0.45.0.
[59] H. L. Tang, V. Shkolnikov, G. S. Barron, H. R. Grimsley, N. J. Mayhall, E. Barnes, and S. E. Economou, PRX Quantum **2**, 020310 (2021).
[60] K. Kowalski and N. P. Bauman, Phys. Rev. Lett. **131**, 200601 (2023).
[61] M. Zheng, B. Peng, A. Li, X. Yang, and K. Kowalski, "QuGCM," `https://github.com/pnnl/QuGCM` (2022).
[62] A. Hashim, R. Rines, V. Omole, R. K. Naik, J. M. Kreikebaum, D. I. Santiago, F. T. Chong, I. Siddiqi, and P. Gokhale, arXiv preprint (2021).
[63] Y. S. Yordanov, D. R. M. Arvidsson-Shukur, and C. H. W. Barnes, Phys. Rev. A **102**, 062612 (2020).
[64] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, "adapt-vqe," `https://github.com/mayhallgroup/adapt-vqe` (2022).
[65] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman,

I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, Nature Methods **17**, 261 (2020), scipy version 1.11.3.