Performance Analysis of Phaseless Auxiliary Field Quantum Monte Carlo

Namita Shah
University of Michigan
Ann Arbor, Michigan, USA
namishah@umich.edu

Neil Mehta
Lawrence Berkeley National Laboratory
Berkeley, California, USA
NeilMehta@lbl.gov

Pooja Rao
NVIDIA Corporation
Santa Clara, California, USA
porao@nvidia.com

Nickolas Joyner
Lawrence Berkeley National Laboratory
Berkeley, California, USA
najoyner@lbl.gov

Katie Klymko
Lawrence Berkeley National Laboratory
Berkeley, California, USA
kklymko@lbl.gov

Ermal Rrapaj
Lawrence Berkeley National Laboratory
Berkeley, California, USA
ermalrrapaj@lbl.gov

Lisa Claus
Lawrence Berkeley National Laboratory
Berkeley, California, USA
LClaus@lbl.gov

Charles Lively III
Lawrence Berkeley National Laboratory
Berkeley, California, USA
charleslively@lbl.gov

I. Introduction

Drug discovery depends on accurately predicting how candidate molecules behave, which is determined by how their electrons interact. Predicting this requires the molecule's ground-state energy, the lowest energy its electrons can collectively have.

Calculating this ground-state energy exactly is intractable for all but the smallest molecules, since the difficulty of accounting for electron correlation grows exponentially with system size (Table 1). Because an exact solution is difficult, Monte Carlo methods are used instead to perform many random walks that converge toward an approximate ground-state energy. Unguided walks converge slowly and can drift toward incorrect answers, so phaseless auxiliary-field quantum Monte Carlo (ph-AFQMC) [1] guides them with a trial wavefunction, a starting guess for the ground state that accelerates convergence and keeps the walk on track.

The quality of this trial wavefunction largely determines both accuracy and computational cost in this approach, making the choice of how to generate it an important design decision. This experiment explores the cost-accuracy tradeoffs between two different quantum methods for producing the trial wavefunction, the Variational Quantum Eigensolver (VQE) and the Generative Quantum Eigensolver (GQE), on NERSC Perlmutter. We explore the cost-accuracy tradeoffs between two quantum methods for producing the trial wavefunction, the Variational Quantum Eigensolver (VQE) [2] and the Generative Quantum Eigensolver (GQE) [3], on NERSC Perlmutter.

II. Methodology

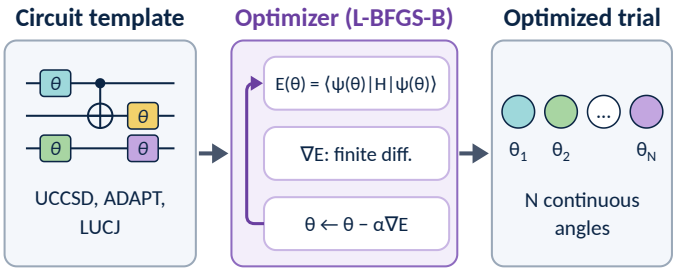
This pipeline has three stages: (1) constructing the molecular Hamiltonian for a given hydrogen chain geometry, (2) preparing a trial wavefunction using either VQE or GQE (Fig. 1), and (3) running ph-AFQMC, whose random walks are guided by the trial wavefunction toward a projected ground-state energy.

Both VQE and GQE build their trial wavefunction from the same pool of excitation operators, each representing one way an electron could move from an occupied orbital to an unoccupied one. By combining these operators, the quantum circuit can approximate a molecule's true electronic state.

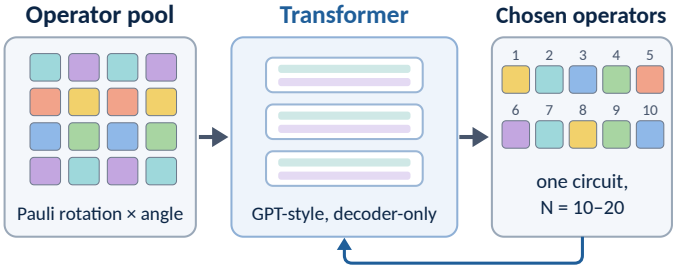
VQE assembles these operators into a fixed-structure circuit with adjustable rotation angles, then uses a classical optimizer adjusts the angles to minimize the circuit's energy. This experiment benchmarks three VQE ansätze: UCCSD [4], which uses a fixed, exhaustive set of single and double excitations; ADAPT-UCCSD [5], which iteratively adds the operator that most reduces the energy; and LUCJ [6], which restricts operators to hardware-local interactions for a shallower circuit. GQE instead trains a classical transformer to determine the sequence of operators that forms the circuit. Each generated sequence is scored by its energy, and the transformer is nudged to favor lower-energy sequences over time, similar to how a reinforcement-learning agent learns from reward. Regardless of method, the final trial wavefunction is passed to the same ph-AFQMC solver.

III. Cost-Accuracy Tradeoffs

We use two metrics: (i) end-to-end application wall time and (ii) the ph-AFQMC projected energy relative to full configuration interaction (FCI), the exact solution used as ground truth. All runs use a single Perlmutter node with four NVIDIA A100 GPUs, on linear hydrogen chains [7] H_4 (8 qubits) and H_8 (16 qubits), to test how each method's cost



(a) VQE: an optimizer tunes the angles of a fixed circuit.



(b) GQE: a transformer, trained on sampled-circuit energies, selects N fixed-angle Pauli rotations.

Fig. 1: Trial-wavefunction preparation methods.

TABLE I: Problem size and circuit complexity

Quantity	H_4	H_8
Electrons	4	8
Qubits (Jordan–Wigner)	8	16
UCCSD excitations (VQE angles)	26	360
UCCSD Pauli rotations	160	2,688
GQE pool (rotations × 12 angles)	1,920	32,256
GQE chosen rotations, N	10, 20	10, 20

scales with problem size. For GQE, the number of operators N is a key hyperparameter: more operators capture more electron correlation but deepen the circuit and raise cost.

Preliminary findings show that at H_4 GQE is the slowest method (∼1,300 s vs 560–810 s mean pipeline time), because transformer training is a fixed cost that dominates at this size. At H_8, however, ADAPT-UCCSD drives the full pipeline to ∼2 hours, while GQE completes end-to-end in ∼41 min on average, about 3× faster (Fig. 2). UCCSD and LUCJ sit between these extremes.

Regarding accuracy, for H_8 at a short bond length (R = 0.8 Å), GQE reaches 0.7 mHa of the FCI energy, within chemical accuracy (error below 1.6 mHa, the threshold for a chemically meaningful result), though above ADAPT-UCCSD (0.03 mHa) and UCCSD (0.015 mHa). Under bond stretching, where electron correlation becomes much stronger, GQE's error grows to ∼19 mHa at R = 1.2 Å and ∼27 mHa at 1.3 Å, while UCCSD and ADAPT-UCCSD remain accurate at 0.48 and 0.61 mHa at 1.2 Å.

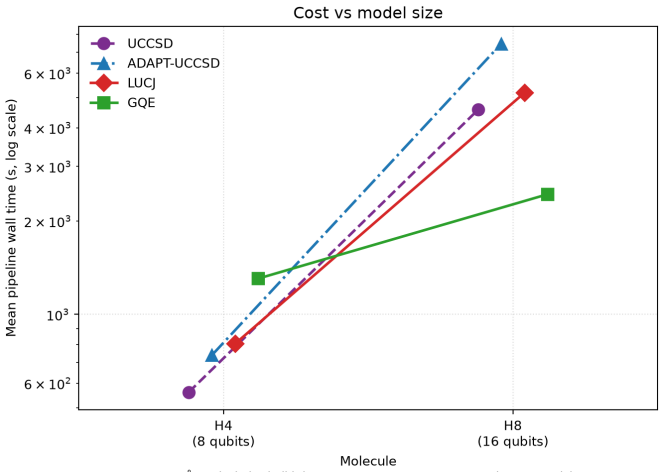


Fig. 2: Mean pipeline wall time over R = 0.8, 1.0, 1.2 Å (one node, 4 A100 GPUs). H_8 VQE times exclude Hamiltonian build.

IV. Next Steps

These preliminary results suggest GQE can substantially reduce trial-generation cost without sacrificing accuracy up to a certain bond length. We plan to extend benchmarking to H_{10} chains (20 qubits) to test whether this cost advantage and accuracy gap persist at larger system sizes and stronger correlation. We will also measure strong scaling of trial generation on 1, 2, and 4 nodes (4, 8, and 16 GPUs) for both the VQE family and GQE; this will help determine which method is more practical on larger HPC allocations. Finally, we are profiling and tuning GPU utilization so that cost comparisons reflect each method's best achievable performance rather than implementation overhead.

References

[1] R. Rofougaran, N. Mehta, K. Klymko, P. Rao, J. W. Mullinax, S. Stein, N. M. Tubman, and E. Rrapaj, "Benchmarking quantum trial wavefunctions for phaseless auxiliary-field quantum Monte Carlo," arXiv:2605.02056, 2026.

[2] A. Peruzzo *et al.*, "A variational eigenvalue solver on a photonic quantum processor," *Nat. Commun.*, vol. 5, p. 4213, 2014.

[3] K. Nakaji *et al.*, "The generative quantum eigensolver (GQE) and its application for ground state search," arXiv:2401.09253, 2024.

[4] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik, "Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz," *Quantum Sci. Technol.*, vol. 4, no. 1, p. 014008, 2019.

[5] H. 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.*, vol. 10, p. 3007, 2019.

[6] M. Motta, K. J. Sung, K. B. Whaley, M. Head-Gordon, and J. Shee, "Bridging physical intuition and hardware efficiency for correlated electronic states: the local unitary cluster Jastrow ansatz for electronic structure," *Chem. Sci.*, vol. 14, pp. 11213–11227, 2023.

[7] M. Motta *et al.*, "Towards the solution of the many-electron problem in real materials: Equation of state of the hydrogen chain with state-of-the-art many-body methods," *Phys. Rev. X*, vol. 7, p. 031059, 2017.