Efficient and scalable AFQMC with Dynamical Trial wave function from variational Ansatz

Zixiang Lu
University of Illinois Urbana-Champaign, Champaign, IL, USA

Zhi-Yu Xiao
Institute of Physics, Chinese Academy of Sciences, P.O. Box 603, Beijing 100190, China *

Shiwei Zhang
Center for Computational Quantum Physics, Flatiron Institute, New York, NY 10010, USA †
(Dated: November 24, 2025)

We present a benchmark study of the scalable framework that couples auxiliary-field quantum Monte Carlo (AFQMC) with a variationally optimized auxiliary-field trial wave function (VAFQMC). The ansatz provides a compact but correlated reference whose dynamically sampled determinants act as an effective multi-determinant trial without requiring large static expansions. Benchmarks across W4-11 atoms, HEAT and W4-MR main-group molecules, heavy lanthanide/actinide atoms, and a Cu^+ basis-set series show that AFQMC/VAFQMC consistently attains chemical accuracy and significantly reduces the systematic bias observed with ROHF and truncated CASSCF trials. Analysis of the optimized auxiliary-field operators reveals that the ansatz selectively enhances short-range AO couplings and yields a phaseless constraint that remains robust under AO-space pruning and variations in ansatz capacity. These results highlight variationally optimized auxiliary-field ansätze as a practical and transferable route for constructing compact, accurate trial wave functions for AFQMC.

I. INTRODUCTION

The accurate prediction of electronic structure in interacting many-electron systems is one of the central challenges in modern quantum chemistry and materials science[1, 2]. The electronic correlation that arises from the interplay between kinetic and Coulomb interactions spans a wide range of energy and length scales, giving rise to rich physical and chemical phenomena that are often beyond the reach of mean-field or perturbative approximations. Despite major advances in density functional theory[3, 4] and coupled-cluster approaches[5, 6], no single electronic structure method can simultaneously achieve high accuracy, broad applicability, and favorable computational scaling across weakly and strongly correlated regimes. Continued progress therefore relies on the development of general and systematically improvable many-body methods that can balance efficiency and accuracy in describing correlated electronic systems.

Among ab initio approaches, quantum Monte Carlo (QMC)[7, 8] offers a statistically exact and highly parallel framework for solving the electronic Schrödinger equation. Within QMC, auxiliary-field quantum Monte Carlo (AFQMC)[9–13] projects to the ground state in imaginary time and controls the sign/phase problem via a phaseless constraint, which introduces a bias whose magnitude and sampling efficiency depend sensitively on the trial wave function. Because AFQMC is formally exact when the trial is exact, practical accuracy hinges on the trial's quality. A common strategy is to employ multi-determinant expansions, often built from complete active spaces[14, 15]; chemically motivated orbital selections capture static correlation but introduce user choice and sensitivity to the active-space definition. Automated constructions—e.g., using natural orbitals[16] selected-CI engines[17], enable very long expansions that systematically approach the exact limit. The trade-off is cost: even with efficient screening—or alternatives such as DMRG-based references—this path is not a black-box, polynomial-scaling solution.

Motivated by black-box usability, several polynomial-scaling trial families have been explored that preserve the favorable AFQMC scaling while allowing systematic improvements. Previous work focuses on the design of trial wave function ansatz to maintain a desired balance between low scaling and high efficiency. Symmetry-broken mean-field references (UHF, GHF, unrestricted Kohn–Sham)[18] can capture static correlation at low cost; symmetry projection[19, 20] (e.g., spin or particle-number projection) partially restores good quantum numbers while retaining compactness. Another line of work updates the mean-field trial self-consistently using AFQMC one-body density matrices[21], yielding orbital-optimized single- or few-determinant trials without expanding to huge configuration spaces. There are also "single-reference but correlated" inputs—such as CISD[22] or natural-orbital-truncated expansions—that avoid manual active-space choices yet remain operationally simple.

To further improve the quality of the constraint, a complementary direction is to import variationally optimized

* zxiaoMain@outlook.com
† szhang@flatironinstitute.org

ansätze as trial wave functions. Variational Monte Carlo (VMC)[23] uses parameterized ansatz to variationally optimize the energy expectations, and with ideally expressive variational ansatz and robust optimization strategies the wave functions can always be optimized to near-exact, providing the potential to serve as a great trial wave function. In the first quantization, Diffusion Monte Carlo often take the trial wave function optimized by Variational Monte Carlo and use fixed-node approximations to perform further projection[24], and in our recent paper[25] and a following paper[26], it has been shown that through dynamically sampling determinants as trial wave function, the variational AFQMC[27] and NNQS can serve as trial wave function of AFQMC and this framework works in some small systems.

In this work, we take a further step to study AFQMC with the variationally optimized, dynamically sampled VAFQMC trial. We begin by presenting the computational framework, then validate the approach on small atoms from W4-11[28], comparing static ROHF trials with dynamically sampled single- and multi-determinant trials. We then benchmark the method on small open-shell atoms from the W4-11 dataset [28], where we compare conventional static ROHF trials with dynamically sampled single- and multi-determinant trials. Moving beyond atoms, we evaluate main-group molecules in the HEAT and W4-MR datasets and assess the resulting projector energies against high-level reference values. We further examine heavy atoms across the lanthanide and actinide series (Eu, Tm, Yb; Am, Md, No), demonstrating that dynamically sampled trials offer significantly improved robustness relative to truncated CASSCF references. Finally, we investigate transition-metal systems through a Cu^+ basis-set study, analyze the learned VHS operators using channel-power visualizations, and probe the sensitivity of the approach to ansatz capacity and AO-space pruning. Together, these results clarify how a variationally optimized auxiliary-field ansatz can serve as an accurate, scalable trial for AFQMC and outline practical considerations for designing compact and robust trial wave functions.

II. RESULTS

A. Computational Framework

We first outline the computational framework used in this work. Our procedure has two stages: we first optimize a variational auxiliary-field ansatz and then embed the optimized ansatz as the trial state in a phaseless AFQMC projection. The ansatz we employ is a single-slice variational AFQMC (VAFQMC) wavefunction [27], which has the advantage that it represents explicit electronic correlations directly in the same auxiliary-field language as AFQMC while remaining lightweight enough to be sampled alongside the walkers. Formally, the VAFQMC ansatz is written as an integral over auxiliary fields,

$$|\Psi_\theta\rangle = \int \prod_\gamma \frac{dx_\gamma}{\sqrt{2\pi}}\, e^{-\frac{1}{2}x_\gamma^2}\, \hat{U}_\theta(x)\, |\tilde{\Psi}_{\mathrm{I}}\rangle, \tag{1}$$

$$\hat{U}_\theta(x) = e^{-\tilde{T}(\theta)}\, e^{\sum_\gamma x_\gamma \tilde{L}_\gamma(\theta)}\, e^{-\tilde{T}_0(\theta)}. \tag{2}$$

Here $|\tilde{\Psi}_{\mathrm{I}}\rangle$ is a parameterized Slater determinant and all operator blocks $\tilde{T}_0(\theta), \tilde{T}(\theta), \tilde{L}_\gamma(\theta)$ carry variational parameters. Although the formal ansatz in Eqs. (1)–(2) contains both one- and two-body operator blocks, in Fig. 1(a) we depict only the auxiliary-field channels $\{\tilde{L}_\gamma\}$. These channels play the role analogous to the Jastrow correlation factor in a Slater–Jastrow wavefunction, and constitute the dominant correlation-carrying component of the ansatz. The one-body terms $\tilde{T}_0$ and $\tilde{T}$ are absorbed into the surrounding linear transformations of the determinant in practice and are omitted from the schematic for clarity.

We optimize $|\Psi_\theta\rangle$ by Monte Carlo sampling pairs of field configurations (x, x') and minimizing the stochastic energy with a soft sign constraint. For more details regarding the optimization of VAFQMC, refer to [27]. This stage yields an explicitly correlated, walker-compatible trial bra $\langle\Psi_T|$ that lives in the same auxiliary-field manifold as AFQMC. In the second stage (Fig. 1(b)), this trial is attached to each AFQMC walker and its overlap ratios are re-sampled on the fly, providing the phaseless constraint without reverting to a large deterministic multi-determinant expansion. To further enhance scalability, we extended the original VAFQMC implementation to a distributed multi-GPU framework. Independent Monte Carlo walkers are assigned to different GPUs and communicate only through lightweight reductions of observables, preserving stochastic independence while achieving near-linear strong scaling across devices. Leveraging JAX's functional programming model and automatic parallelization primitives[29], this distributed design can be implemented with minimal overhead, enabling us to treat substantially larger molecular and transition-metal systems than previously feasible.

Once the variational parameters are converged, we freeze them and regard the resulting state as the trial wave function for the projector stage. Denoting the optimized ansatz by $|\Psi_T\rangle$, we emphasize that it lives in the same auxiliary-field representation as the AFQMC propagators, because it was constructed from exponentials of parameterized one- and two-body operators. This removes the usual mismatch between a deterministically generated multi-determinant trial and the stochastic projector. In the AFQMC stage (Fig. 1(b)), each random walker carries its Slater determinant and, when required by the phaseless constraint, evaluates overlap ratios with respect to the VAFQMC trial. Concretely, for a walker state $|\phi_k^i\rangle$ at imaginary-time step i, we form

$$\frac{\langle\Psi_T|\phi_k^{i+1}\rangle}{\langle\Psi_T|\phi_k^i\rangle},$$

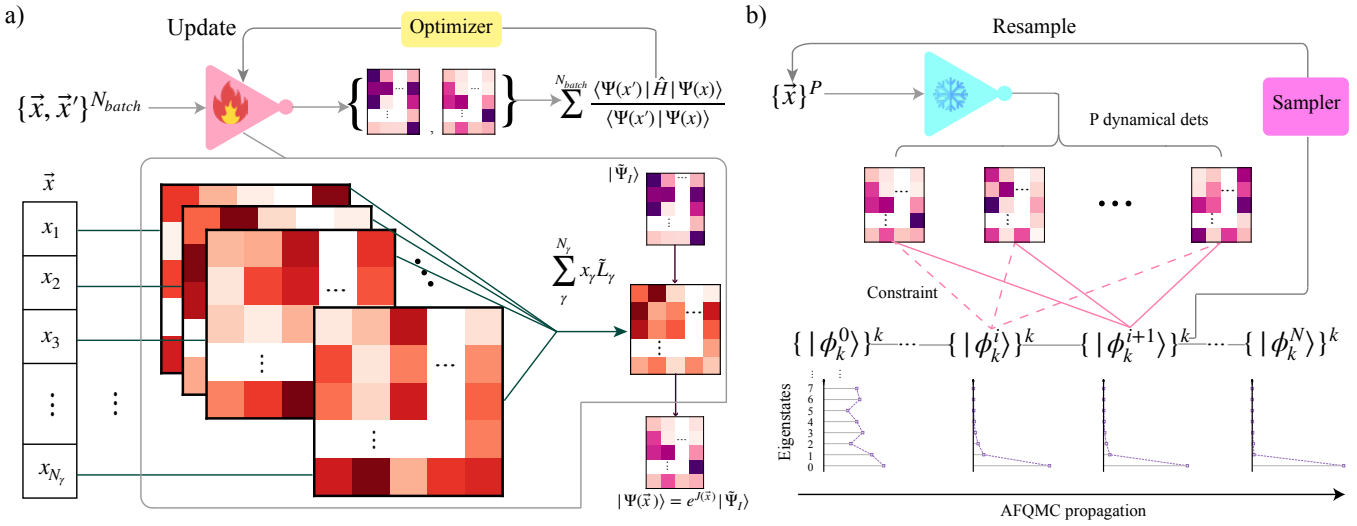

FIG. 1. **Computational framework of this work.**[a] **(a)** Variational optimization of the VAFQMC trial wave function. A batch of sampled auxiliary–field configurations $\{\vec{x}, \vec{x}'\}_{N_{\text{batch}}}$ is used to stochastically evaluate the energy and its gradients, and all variational parameters of the correlated ansatz are updated by the optimizer. Zoom-in: the VAFQMC trial is constructed as a product of exponential one-body propagators, $e^{J}(\vec{x})$, applied successively to an optimizable initial determinant $|\tilde{\Psi}_I\rangle$. The propagator is evaluated through a truncated Taylor expansion $e^{\hat{A}} \approx 1 + \hat{A} + \hat{A}^2/2$, producing the optimized correlated determinant $|\Psi_T(\vec{x})\rangle$. **(b)** Use of the optimized VAFQMC ansatz as a dynamical trial in AFQMC with the phaseless constraint. The optimized ansatz is stochastically resampled to generate P dynamical determinants that constitute the trial manifold enforcing the constraint. During AFQMC propagation, each walker is coupled to a refreshed set of auxiliary fields, yielding updated dynamical determinants that co-evolve with the walker. This leapfrog procedure continuously suppresses excited-state components present in the ensemble-averaged walker state and stabilizes the phaseless constraint, enabling efficient convergence of the projected walkers toward the ground state.

[a] Flame and ice icons in this figure were obtained from Flaticon.com.

and use this ratio to define the importance sampling and the phaseless gauge as in standard phaseless AFQMC. Because $|\Psi_T\rangle$ was itseskelf sampled and optimized in auxiliary-field space, these overlaps can be evaluated with the same determinant machinery as in the variational stage, without expanding $|\Psi_T\rangle$ into an explicit large configuration interaction series. For a detailed implementation, refer to our previous paper[25].

This design allows the projector to benefit from a trial that is already adapted to the system's correlation pattern, while the AFQMC propagation further reduces the residual bias from the variational stage. As a result, the combined VAFQMC+AFQMC framework retains the flexibility of a learned trial state and the robustness of phaseless imaginary-time projection.

B. Application to Small Atoms

(Open shell) To demonstrate the effectiveness of our approach, we first examine the total energies of small atoms from the W4-11 dataset [28]. Previous studies [30] reported that phaseless AFQMC with a static ROHF trial exhibits sizable systematic errors for open-shell atoms, particularly for systems with non-zero angular momentum, and that achieving chemical accuracy

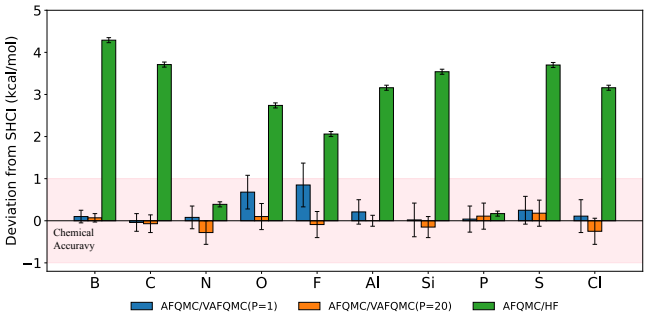

FIG. 2. **Absolute deviations (in kcal/mol) from SHCI reference energies for different trial wave functions.** The green bars represent AFQMC calculations using static ROHF single-determinant trial wave functions, the blue bars correspond to AFQMC with single-determinant dynamical trial wave functions, and the orange bars depict AFQMC with 20 dynamical determinants trial. All AFQMC results employing dynamical trial wave functions fall within chemical accuracy, demonstrating significant improvement over the ROHF trial. The SHCI reference and ROHF-based AFQMC data are taken from Ref. [30].

in such cases typically requires incorporating hundreds of determinants in the trial wave function. In contrast, our framework could attain high accuracy using only a

dynamical single-determinant trial wave function sampled from the variational ansatz. The results, shown in Fig. 2, demonstrate that all computed total energies agree with the reference SHCI values within chemical accuracy (1 kcal/mol), effectively eliminating the bias associated with static single-determinant trials. All variational ansätze were first optimized to convergence. The optimized ansatz was then employed to stochastically sample the single- or multi-determinant configurations used as AFQMC trials. All calculations were performed in the aug-cc-pVTZ basis set. Because the

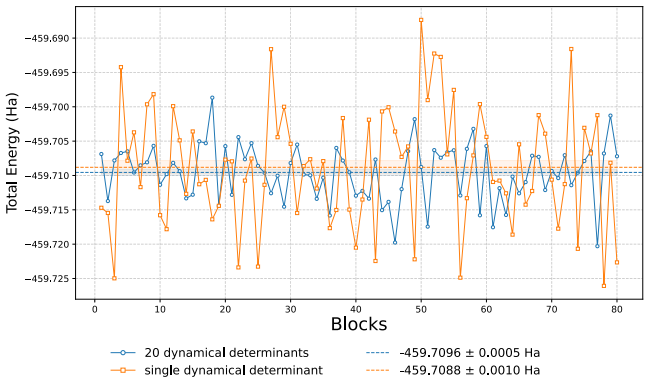

FIG. 3. AFQMC block-averaged total energies for the Cl atom using dynamically sampled trial wave functions: a single determinant (orange) versus 20 determinants (blue). Each marker is one estimation block after equilibration. Dashed lines denote the sample means; the legend reports mean ± one–standard-error. With the same number of blocks (80), the single-determinant trial shows about twice the statistical uncertainty (~ 1 mHa) and visibly larger block-to-block fluctuations than the 20-determinant trial (~ 0.5 mHa).

trial is sampled dynamically, residual stochastic variance is introduced, which amplifies fluctuations in the phaseless constraint and increases the statistical uncertainty when only a single determinant is used. As shown in Fig. 3, the AFQMC energy for Cl obtained with a single dynamical determinant exhibits markedly larger block-to-block fluctuations than with twenty dynamically sampled determinants, so more blocks are required to reach the same precision. Running for 80 blocks in both cases, the single-determinant run yields an error bar of ≈1,mHa, roughly twice that of the 20-determinant run (≈0.5,mHa). Extending the single-determinant simulation to ∼240 blocks reduces its error bar to a comparable level. Population-control effects also become more pronounced under stronger phaseless fluctuations, intermittently removing additional walkers and adding variance. These effects are substantially mitigated by dynamical multi-determinant trials; in practice, using about 20 determinants is sufficient, and we adopt this setting in subsequent larger or more strongly correlated systems. In what follows, we will use **VAFQMC** to denote the variationally optimized ansatz, and **AFQMC/VAFQMC** to denote AFQMC calculations that employ 20 dynamically sampled determinants from this ansatz as the trial wave function.

The variational optimization in VAFQMC requires repeated evaluations of the energy functional and its gradients, each involving Slater determinant updates and contractions over the Cholesky factors and therefore scaling as $O(N^3)$ with the number of basis functions N. The total optimization cost is $O(N^3 N_{\mathrm{conv}})$, where N_{conv} denotes the number of optimization iterations. Consistent with experience in variational Monte Carlo, N_{conv} exhibits only weak dependence on system size and is well approximated by $N_{\mathrm{conv}} \propto N^{\beta}$ with $\beta \lesssim 1$ for chemically relevant systems. Following optimization, dynamical trial configurations are generated through Metropolis–Hastings sampling of the auxiliary-field paths. Each Metropolis proposal modifies only a small subset of fields and requires no recomputation of determinant ratios, Green's functions, or Cholesky contractions, resulting in a subcubic cost $O(N^{\alpha})$ with $1 \leq \alpha \leq 2$. Even with P dynamical trials and M updates per AFQMC time step, this overhead remains negligible relative to the $O(N^3)$ AFQMC propagation. Thus, AFQMC with dynamical trials is effectively equivalent to AFQMC employing a P-determinant trial while preserving the cubic dependence on N. Collecting all contributions, the overall computational complexity of the combined VAFQMC–AFQMC framework may be summarized as

$$T_{\mathrm{total}} = O(N^{3+\beta}) + O(PN^{3+\gamma}),$$

where $N_\tau \propto N^{\gamma}$ with $\gamma \lesssim 1$ denotes the number of AFQMC projection steps. Hence, the introduction of dynamical trials preserves the favorable cubic scaling of AFQMC while introducing only modest prefactors from trial optimization and sampling.

We attribute the superior performance of the dynamical trials to their higher effective fidelity to the exact ground state. In static approaches, hundreds of determinants are often needed to reach a near-unity overlap with the exact wave function. By contrast, a variationally optimized ansatz already provides an accurate correlated reference, so that even a small number of dynamically sampled determinants yields highly accurate trials, with the remaining stochastic variance decreasing as the number of determinants increases.

C. Application to main group molecules

Dataset	AFQMC/HF	AFQMC/CISD	AFQMC/VAFQMC
HEAT	2.84	0.75	0.22
W4-MR	3.33	1.87	0.93

TABLE I. RMSD (mHa) of AFQMC energies using Hartree–Fock, CISD, and VAFQMC trials across the HEAT and W4-MR datasets.

We evaluate the performance of our approach using two established benchmarks, HEAT[33] and the

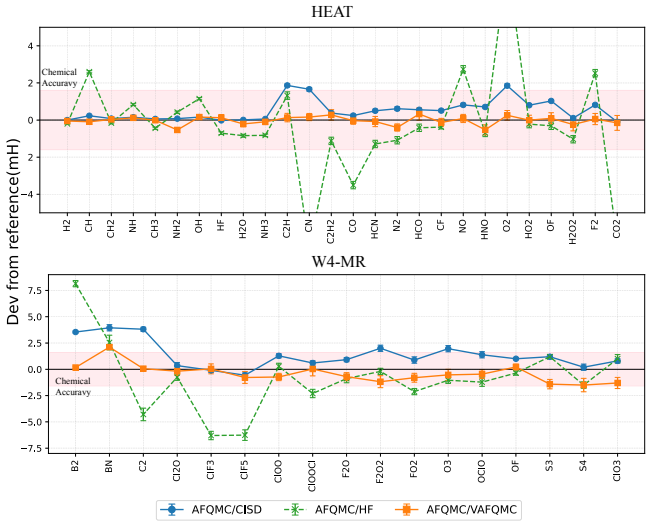

FIG. 4. Deviation of the AFQMC total energies obtained using VAFQMC trials, Hartree–Fock trials, and CISD trials across the HEAT and W4-17-MR benchmark molecules in CC-pVDZ basis. For both datasets, the AFQMC/CISD results are taken from Ref. [22]. For the HEAT dataset, the CCSDTQP reference values are from Ref. [31], and the AFQMC/HF data are from Ref. [32]. For the W4-17-MR dataset, both the reference values and the AFQMC/HF data are taken from Ref. [22].

.

strongly multireference subset W4 (W4-MR)[34], both of which present stringent tests for many-body electronic-structure methods. The HEAT set consists mainly of first-row species with relatively mild multireference character, whereas W4-MR was explicitly designed to contain systems dominated by strong static correlation. Together, these datasets provide a consistent and widely adopted reference for assessing correlated wave function methods. For comparison, we include AFQMC calculations using Hartree–Fock (AFQMC/HF) and CISD-based (AFQMC/CISD) trial wave functions. The HF determinant provides the simplest single-reference baseline, while the CISD expansion supplies a compact yet correlated trial that is commonly used in AFQMC benchmark studies. These baselines establish a clear context for quantifying the additional accuracy afforded by the variationally optimized VAFQMC trial. All calculations employ a uniform set of parameters across all species for both VAFQMC and AFQMC, ensuring that the performance of the method is assessed under a consistent computational protocol rather than system-dependent tuning. Furthermore, to remain consistent with Ref. [22], we freeze the He core for second-row atoms and the Ne core for third-row atoms in all calculations.

Across the HEAT dataset, the VAFQMC–guided AFQMC energies show uniformly small and tightly bounded deviations relative to the reference values. No chemical subgroup displays noticeably larger errors, indicating that the method maintains a stable constraint throughout the set and adapts reliably even in comparatively simple electronic environments. The W4-17-MR dataset presents a more challenging test, with an overall RMSD of 0.93 mHa, yet most of the results still fall within chemical accuracy. A mild trend appears for the larger and more strongly correlated molecules, such as S_3, S_4, and ClO_3, where the deviations increase slightly. This behavior is consistent with the expectation that, under fixed optimization parameters, the variationally optimized trial may not retain the same level of quality as in smaller systems.

Taken together, these results highlight an essential feature of the approach: the adaptivity of the trial manifold significantly reduces constraint variability across chemically diverse systems. The method performs consistently from single-reference to strongly multireference regimes while maintaining a low computational overhead. The HEAT and W4-MR benchmarks therefore illustrate that the central advantage lies not in targeting any specific class of molecules, but in providing a generally robust trial formulation that maintains accuracy across heterogeneous chemical environments.

D. Heavy Atoms

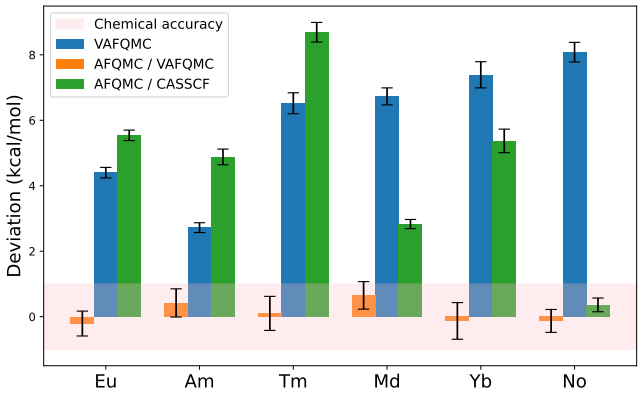

FIG. 5. Deciation of VAFQMC, AFQMC/VAFQMC and AFQMC/CASSCF from high order couple cluster(HOCC) energies.

Heavy atoms—the lanthanides and actinides—are known to be difficult for AFQMC when CASSCF trial wave functions are used. As a compact yet representative subset, we consider six atoms spanning both series and the three canonical f-shell regimes: Eu and Am (half-filled), Tm and Md (nearly filled), and Yb and No (closed shell). This choice captures open- and nominally closed-shell cases and the characteristic challenges—near-degeneracy, core–valence correlation, and trial nodal/phase sensitivity—while keeping the Hamiltonian and basis fixed (X2C/cc-pVDZ-X2C). Localized $4f/5f$ shells and dense $f/d/s$ near-degeneracies make absolute total energies highly sensitive to the bal-

ance between non-dynamical and dynamical correlation. In this setting, CAS-based trials are fragile: the active-space boundary shifts with $f \leftrightarrow d/s$ mixing, and practical CI truncations bias absolute energies; thus, the perceived difficulty is primarily a trial-robustness issue rather than a limitation of AFQMC.

Figure 5 summarizes our results under the same X2C/cc-pVDZ-X2C setup. For the actinides (Am, Md, No) we correlate {6s, 6p, 7s, 5f} and freeze all deeper shells; for the lanthanides (Eu, Tm, Yb) we analogously correlate {5s, 5p, 6s, 4f} and freeze all deeper shells. In our calculations, the variational energies of the VAFQMC ansatz still deviate from the HOCC reference by several kcal/mol. When 20 dynamically sampled determinants from this ansatz are used as the AFQMC trial, the results are consistently within chemical accuracy for all six atoms. For the CASSCF trials, we truncated the expansion to 10,000 determinants; the resulting AFQMC with the CASSCF trial (green) is less accurate and often falls outside the chemical-accuracy band. This shows that, once a stable dynamically sampled multi-determinant trial is used, AFQMC itself is reliable for heavy-atom total energies at this level of theory.

E. Cu^+ basis behavior and trial sensitivity

Transition-metal systems remain prototypical testbeds for strong correlation and continue to challenge standard quantum-chemistry approximations. Prior AFQMC benchmarks have shown that truncated CASSCF multi-determinant trials can still leave non-negligible bias for these molecules [35, 36]. Our recent work demonstrated that AFQMC with a stochastically sampled VAFQMC trial systematically attains chemical accuracy for a panel of transition-metal diatomics in a modest basis [25]. Here we examine how this approach behaves in larger bases and how sensitive it is to the quality of the optimized VAFQMC ansatz.

We take Cu^+ as a representative case. Using the same VAFQMC and AFQMC hyperparameters across all calculations, we first optimize the VAFQMC ansatz to convergence and then employ it as the trial in ph-AFQMC. We compare directly to near-exact SHCI energies from Ref. [35] under identical settings (equilibrium geometry, pseudopotential, and basis sets). Figure 6(a) reports deviations from SHCI for four bases (vdz, vtz, vqz, v5z). The variational VAFQMC energies (green) already track the reference within a few kcal/mol, with a mild basis dependence. When used as trials, AFQMC/VAFQMC (orange) yields residual errors consistently within the chemical-accuracy band for all bases, indicating that the constraint built from the stochastic trial is sufficiently faithful even as the one-particle space grows. By contrast, AFQMC/CASSCF (blue) retains a systematic underestimation of ∼3–5 kcal/mol across bases. Overall, the results show that the variational quality of the VAFQMC ansatz is good but not perfect, while once embedded as a trial, the resulting AFQMC/VAFQMC corrects the remaining bias and maintains accuracy as the basis is enlarged to v5z.

To make the effect of variational optimization directly visible, we proceed in two steps. First, we transform the optimized operators from the canonical molecular–orbital (MO) basis, where both VAFQMC and AFQMC are formulated—back to the atomic–orbital (AO) basis. This eliminates system–dependent MO rotations and allows structural patterns in the auxiliary fields to be compared on a common, chemically meaningful basis. When viewed in the AO basis, each individual VHS channel appears remarkably sparse and localized, indicating that every L_γ captures only a specific fragment of the total correlation structure rather than acting as a fully delocalized operator. Next, we aggregate the full set of auxiliary one–body operators $\{L_\gamma\}_{\gamma=1}^{\Gamma}$ (the "VHS" channels produced by the modified–Cholesky factorization) into a single heatmap. Although the Cholesky procedure initializes different channels with different numerical scales, within the VAFQMC framework all channels are treated as trainable parameters on an equal footing. We therefore combine them uniformly by forming a channel–power map,

$$P_{ij} = \Big(\sum_{\gamma=1}^{\Gamma} [L_{\gamma,ij}]^2 \Big)^{1/2},$$

which measures how strongly each pair of orbitals (i, j) is coupled across all auxiliary channels. Before plotting, for each setting (initial/optimized) we apply a single global Frobenius normalization to the whole stack $\{L_\gamma\}$ so that the heatmaps emphasize structure rather than overall scale. All magnitude panels share a common color scale, and we omit numeric annotations since the visual goal is pattern recognition.

To summarize the change compactly, we compute a stabilized relative change,

$$\Delta_{ij} = \frac{P_{\mathrm{opt},ij} - P_{\mathrm{init},ij}}{P_{\mathrm{init},ij} + \tau},$$

with a small stabilizer $\tau = 10^{-6} \times \mathrm{median}\big(P_{\mathrm{init}}^{(B)}\big)$ to avoid division–by–near–zero artifacts; entries where both P_{init} and P_{opt} fall below the joint 10th percentile are masked. These difference maps (right column) highlight that in the AO representation the adjustments are localized—optimization selectively increases the weight of a few p– and d–shell blocks and slightly suppresses others, without introducing long–range structures. This supports the interpretation that VAFQMC mainly refines short–range couplings while preserving the global organization of the trial.

To probe how the quality/size of the variational trial impacts projector accuracy, we reduce the number of trainable VHS operators in the ansatz to $n_{\mathrm{vhs}} \in \{10, 30, 50, 70, 90, 100\}$. For each n_{vhs} we optimize the VAFQMC energy to convergence on the `vdz` basis, and

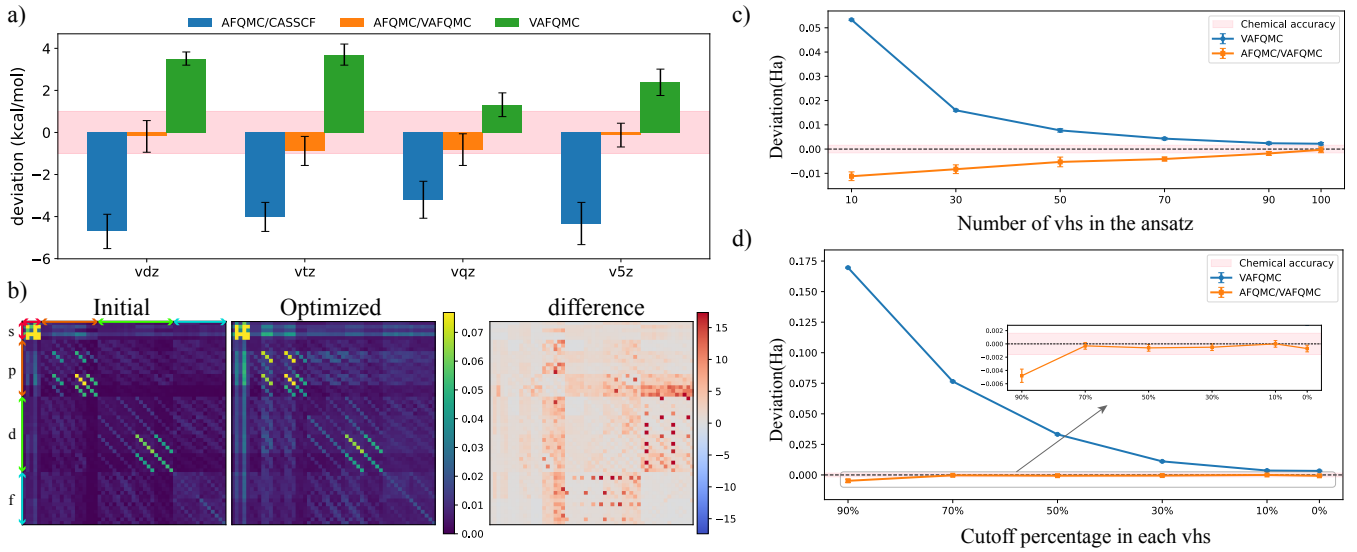

FIG. 6. **Cu^+ basis-set study. (a)** Deviations from SHCI (kcal/mol) for four basis sets. The shaded band marks the chemical-accuracy window (±1 kcal/mol). Blue: AFQMC with truncated CASSCF trials (AFQMC/CASSCF). Orange: AFQMC with stochastically sampled VAFQMC trials (AFQMC/VAFQMC). Green: variational VAFQMC energies (trial energies). SHCI reference energies, basis-set and pseudopotential settings, and the AFQMC/CASSCF data are taken from Ref. [35]. **(b)**Visualization of VAFQMC trial operators in vdz basis. Each panel shows the channel–power map $P_{ij} = \left(\sum_\gamma L^2_{\gamma,ij}\right)^{1/2}$ aggregated from the VHS operators $\{L_\gamma\}$. AO panels are labeled by $s/p/d/f$ blocks. Left and middle columns display the initial (chunked Cholesky) and optimized parameters, respectively, after a single global Frobenius normalization of each stack; magnitude panels share a common color scale. The right column shows the difference of the initial and optimized vhs. **(c)** Deviations from SHCI (Ha) for the variational energy (VAFQMC, blue) and for AFQMC using the optimized VAFQMC trial (AFQMC/VAFQMC, orange) as a function of the number of trainable VHS operators $n_{\rm vhs} \in \{10, 30, 50, 70, 90, 100\}$ on the `vdz` basis. **(d)** Deviations from SHCI (Ha) versus AO–basis elementwise cutoff percentage $p \in \{90, 70, 50, 30, 10, 0\}\%$ applied to each VHS operator: starting from the optimized trial ($n_{\rm vhs} = 100$), the operators are rotated to the AO basis, within each $L^{\rm AO}_\gamma$ the smallest $|\cdot|$ entries at rate p are set to zero, and the operators are rotated back to the canonical MO basis; no further optimization is performed. The inset provides a zoom of the AFQMC/VAFQMC curve near zero.

then use the optimized ansatz as the AFQMC trial. Figure 6(c) reports deviations from SHCI (Ha). As expected for a higher–capacity ansatz, the variational error decreases monotonically with $n_{\rm vhs}$. AFQMC using the same optimized trials follows the same trend but remains within chemical accuracy only when the expresiveness of the ansatz is not overly reduced: as $n_{\rm vhs}$ decreases the AFQMC/VAFQMC deviation grows and, in our `vdz` test, it leaves the ±1 kcal/mol band ($\pm 1.593 \times 10^{-3}$ Ha) once $n_{\rm vhs} < 90$. Therefore, within this ansatz and optimization framework, the VAFQMC energy is a useful indicator not only of how well the wave function has been optimized but also of the trial's suitability for imposing the phaseless constraint; at the same time, the results caution against aggressive pruning of $n_{\rm vhs}$.

Motivated by the AO sparsity observed in Fig. 6(b), we performed a second test in which we start from the well–optimized trial, rotate the VHS to the AO basis, and then within each $L^{(\rm AO)}_\gamma$ zero out a fixed percentage $p \in \{10, 30, 50, 70, 90\}\%$ of the smallest $|\cdot|$ entries (keeping the largest ones). We then rotate back to the canonical MO basis and, without any further optimization, evaluate the variational energy of this pruned ansatz and, use it as the AFQMC trial. As shown in Fig. 6b), the VAFQMC(trial) energy deteriorates rapidly with p because many small–magnitude AO entries are removed; however, the AFQMC projector energy is remarkably robust and stays within the chemical–accuracy band up to about $p = 90\%$. This behavior is consistent with the view that pruning in AO mainly discards weak, long–range couplings while leaving the dominant local structure—and hence the sign/phase information most relevant for the phaseless constraint—largely intact. Two takeaways follow: (i) the variational energy is not always a sufficient golden standard for the utility of a trial—trials with poor variational energy can still furnish a good phaseless constraint if their physical structure is preserved; and (ii) this suggests room for alternative optimization strategies and more scalable ansätze that seek trials which are compact in AO yet retain the essential sign/phase structure, thereby reducing compute while keeping AFQMC accurate.

III. SUMMARY AND DISCUSSION

The present work demonstrates that variational auxiliary–field ansätze provide an efficient and transferable

strategy for constructing compact trial wave functions for phaseless AFQMC. By optimizing the ansatz directly in auxiliary–field space and sampling a small determinant manifold dynamically, the resulting AFQMC/VAFQMC scheme achieves uniform performance across chemically diverse systems—including open-shell atoms, main-group molecules, heavy lanthanides and actinides, and transition-metal, without system-specific active-space choices or tailored hyperparameters. The behaviour of the learned VHS operators in the AO basis further clarifies the mechanism: optimization chiefly reorganizes short-range couplings while preserving the global operator structure, yielding a trial that stabilizes the phaseless constraint even when the variational energy alone is not perfect. Together, these results highlight that the primary requirement for an effective AFQMC trial is not exhaustive amplitude accuracy but a faithful representation of the sign/phase structure relevant to the constraint, which the auxiliary–field formulation naturally captures.

These findings motivate a broader perspective on combining variational optimization with projector methods. The VAFQMC stage provides a correlated, system-adaptive reference of modest dimension, while the AFQMC projection efficiently removes remaining variational bias. This separation of roles preserves the favorable polynomial scaling of AFQMC and enables straightforward deployment on distributed multi-GPU architectures. The observed AO sparsity and robustness to operator pruning suggest opportunities for more structured or sparsity-promoting parameterizations, including symmetry-equivariant forms and neural-network augmentations of the VHS manifold. Overall, the results show that variationally optimized auxiliary–field trials constitute a reliable and scalable approach for AFQMC, yielding consistently improved accuracy across both weakly and strongly correlated systems even when employed in a straightforward, black-box manner. Moreover, as a variationally optimized trial wave function, the VAFQMC form naturally accommodates alternative objective functions, allowing the optimization to target specific electronic states[37, 38] when desired and providing a versatile protocol for future applications.

IV. METHODS

In this section, we briefly introduce the methods we use to perform the AFQMC with VAFQMC trial wave function. For the details regarding the implementation, please refer to [25, 27].

Variational auxiliary-field ansatz

We employ a variational auxiliary-field quantum Monte Carlo (VAFQMC) ansatz to construct a compact yet expressive correlated wave function. The central idea is to treat a short imaginary-time AFQMC propagator as a variational object whose one-body and two-body components are optimized jointly with the initial Slater determinant. The resulting state takes the continuous form

$$|\psi_\theta\rangle = \exp\!\left[-\int_0^\tau \tilde{H}(s)\,\mathrm{d}s\right]\left|\tilde{\psi}_I\right\rangle, \tag{3}$$

where θ denotes all variational parameters. After discretization and a Hubbard–Stratonovich decoupling, the ansatz becomes an integral over auxiliary fields,

$$\begin{aligned}|\psi_\theta\rangle = \int\!\left(\prod_{\gamma,l}\frac{\mathrm{d}x_{\gamma l}}{\sqrt{2\pi}}e^{-x_{\gamma l}^2/2}\right)\!\left(\prod_{l=1}^{N_l} e^{-t_l\tilde{T}_l}\, e^{\sqrt{-s_l}\sum_{\gamma=1}^{N_\gamma} x_{\gamma l}\tilde{L}_\gamma}\right)\\ \times e^{-t_0\tilde{T}_0}\left|\tilde{\psi}_I\right\rangle \equiv \int \mathrm{d}\mathbf{x}\,\tilde{U}(\mathbf{x})\left|\tilde{\psi}_I\right\rangle,\end{aligned} \tag{4}$$

with $\mathbf{x} = \{x_{\gamma l}\}$ and $\{t_l, s_l\}$ denoting optimizable step sizes. Because the full propagator is treated variationally, the ansatz is free of Trotter error and admits relatively large time steps; in practice we adopt a single propagation layer ($l = 1$), which yields a favorable balance between expressiveness and computational efficiency.

The variational parameters are obtained by minimizing the projected energy

$$E_\theta = \frac{\displaystyle\int \mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{x}'\,\langle\tilde{\psi}_I|\tilde{U}^\dagger(\mathbf{x}')\hat{H}\tilde{U}(\mathbf{x})|\tilde{\psi}_I\rangle}{\displaystyle\int \mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{x}'\,\langle\tilde{\psi}_I|\tilde{U}^\dagger(\mathbf{x}')\tilde{U}(\mathbf{x})|\tilde{\psi}_I\rangle}. \tag{5}$$

For Monte Carlo evaluation, each overlap is decomposed into its magnitude $\mathcal{D}$, phase $\mathcal{S}$, and local energy $\mathcal{E}$,

$$\mathcal{D}(\mathbf{x},\mathbf{x}') = \left|\langle\tilde{\psi}_I|\tilde{U}^\dagger(\mathbf{x}')\tilde{U}(\mathbf{x})|\tilde{\psi}_I\rangle\right|, \tag{6}$$

$$\mathcal{S}(\mathbf{x},\mathbf{x}') = \frac{\langle\tilde{\psi}_I|\tilde{U}^\dagger(\mathbf{x}')\tilde{U}(\mathbf{x})|\tilde{\psi}_I\rangle}{\mathcal{D}(\mathbf{x},\mathbf{x}')}, \tag{7}$$

$$\mathcal{E}(\mathbf{x},\mathbf{x}') = \frac{\langle\tilde{\psi}_I|\tilde{U}^\dagger(\mathbf{x}')\hat{H}\tilde{U}(\mathbf{x})|\tilde{\psi}_I\rangle}{\langle\tilde{\psi}_I|\tilde{U}^\dagger(\mathbf{x}')\tilde{U}(\mathbf{x})|\tilde{\psi}_I\rangle}. \tag{8}$$

The energy then becomes

$$E_\theta = \frac{\langle\mathcal{E}\mathcal{S}\rangle_{\mathcal{D}}}{\langle\mathcal{S}\rangle_{\mathcal{D}}}. \tag{9}$$

To stabilize the optimization against sign fluctuations, we penalize reductions of the average sign

$$S_\theta = \mathrm{Re}\big[\langle\mathcal{S}\rangle_{\mathcal{D}}\big], \tag{10}$$

and minimize the augmented objective

$$\min_\theta\big[E_\theta + \lambda\max(B - S_\theta,\,0)^2\big], \tag{11}$$

with fixed $\lambda = 1$ and $B = 0.7$. This soft constraint preserves variational integrity while preventing the optimization from collapsing into low-sign regions of configuration space.

AFQMC with VAFQMC as trial wave function

The optimized VAFQMC state provides an effective trial wave function for phaseless AFQMC. Although expressed compactly as a single variational propagator, the ansatz expands into an auxiliary-field–dependent linear combination of Slater determinants,

$$|\psi_\theta\rangle = \int \mathrm{d}\mathbf{x}\, \tilde{U}(\mathbf{x}) \left|\tilde{\psi}_I\right\rangle, \tag{12}$$

so that every evaluation of overlaps or local energies requires sampling a new set of fields $\mathbf{x}$. In practice, AFQMC therefore interacts not with a fixed trial determinant, but with a dynamical single- or multi-determinant representation whose composition changes at each call. This provides access to correlation patterns that would otherwise require very large static multi-determinant expansions.

For each walker $|\phi\rangle$, the trial enters through the AFQMC overlap ratio

$$O_\theta(\phi) = \int \mathrm{d}\mathbf{x} \langle\tilde{\psi}_I|\tilde{U}^\dagger(\mathbf{x})|\phi\rangle \tag{13}$$

and the corresponding mixed local energy

$$E^\theta_{\mathrm{loc}}(\phi) = \frac{\int \mathrm{d}\mathbf{x} \left\langle\tilde{\psi}_I\right|\tilde{U}^\dagger(\mathbf{x})\hat{H}\left|\phi\right\rangle\right\rangle}{\int \mathrm{d}\mathbf{x} \left\langle\tilde{\psi}_I\right|\tilde{U}^\dagger(\mathbf{x})\left|\phi\right\rangle\right\rangle}. \tag{14}$$

Each term $\tilde{U}(\mathbf{x})\left|\tilde{\psi}_I\right\rangle$ is itself a Slater determinant, so the stochastic sampling over $\mathbf{x}$ produces a sequence of determinants that collectively act as a time-dependent multi-determinant trial.

During AFQMC propagation, the phaseless constraint is applied using the phase of the sampled overlap ratio,

$$w \to w\, \max(0, \cos\Delta\theta), \tag{15}$$

$$\Delta\theta = \arg O_\theta(\phi_{n+1}) - \arg O_\theta(\phi_n). \tag{16}$$

while $\mathrm{Re}[E^\theta_{\mathrm{loc}}]$ contributes to the mixed estimator for the total energy.

Because the effective trial seen by AFQMC consists of a large number of on-the-fly determinants generated through the auxiliary-field sampling, the resulting constraint is substantially more flexible than a conventional single-determinant or static multi-determinant trial. This dynamical representation significantly improves the stability of the phaseless projection and enhances accuracy across challenging correlated systems.

V. ACKNOWLEDGMENTS

VI. COMPETING INTERESTS

The authors declare no competing interests.

[1] J. A. Pople, Nobel lecture: Quantum chemical models, Rev. Mod. Phys. **71**, 1267 (1999).

[2] W. Kohn, Nobel lecture: Electronic structure of matter—wave functions and density functionals, Rev. Mod. Phys. **71**, 1253 (1999).

[3] P. Hohenberg and W. Kohn, Inhomogeneous electron gas, Phys. Rev. **136**, B864 (1964).

[4] W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, Phys. Rev. **140**, A1133 (1965).

[5] R. J. Bartlett and M. Musiał, Coupled-cluster theory in quantum chemistry, Rev. Mod. Phys. **79**, 291 (2007).

[6] T. D. Crawford and H. F. Schaefer III, An introduction to coupled cluster theory for computational chemists, in Reviews in Computational Chemistry (John Wiley Sons, Ltd, 2000) pp. 33–136.

[7] F. Becca and S. Sorella, Quantum Monte Carlo Approaches for Correlated Systems (Cambridge University Press, 2017).

[8] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Quantum monte carlo simulations of solids, Rev. Mod. Phys. **73**, 33 (2001).

[9] S. Zhang, J. Carlson, and J. E. Gubernatis, Constrained path quantum monte carlo method for fermion ground states, Phys. Rev. Lett. **74**, 3652 (1995).

[10] S. Zhang, J. Carlson, and J. E. Gubernatis, Constrained path monte carlo method for fermion ground states, Phys. Rev. B **55**, 7464 (1997).

[11] S. Zhang and H. Krakauer, Quantum monte carlo method using phase-free random walks with slater determinants, Phys. Rev. Lett. **90**, 136401 (2003).

[12] M. Motta and S. Zhang, Ab initio computations of molecular systems by the auxiliary-field quantum monte carlo method, WIREs Computational Molecular Science **8**, e1364 (2018).

[13] H. Shi and S. Zhang, Some recent developments in auxiliary-field quantum monte carlo for real materials, The Journal of Chemical Physics **154**, 024107 (2021).

[14] W. A. Al-Saidi, S. Zhang, and H. Krakauer, Bond breaking with auxiliary-field quantum monte carlo, The Journal of Chemical Physics **127**, 144101 (2007).

[15] W. Purwanto, S. Zhang, and H. Krakauer, An auxiliary-field quantum monte carlo study of the chromium dimer, The Journal of Chemical Physics **142**, 064302 (2015).

[16] J. S. Kurian, H.-Z. Ye, A. Mahajan, T. C. Berkelbach, and S. Sharma, Toward linear scaling auxiliary-field quantum monte carlo with local natural orbitals, Journal of Chemical Theory and Computation **20**, 134 (2024).

[17] A. Mahajan, J. Lee, and S. Sharma, Selected configuration interaction wave functions in phaseless auxiliary field quantum monte carlo, The Journal of Chemical Physics **156**, 174111 (2022).

[18] D. Danilov, B. Ganoe, M. Munyi, and J. Shee, Capturing strong correlation in molecules with phaseless auxiliary-field quantum monte carlo using generalized hartree–fock trial wave functions, Journal of Chemical Theory and Computation **21**, 1136 (2025).

[19] H. Shi and S. Zhang, Symmetry in auxiliary-field quantum monte carlo calculations, Phys. Rev. B **88**, 125132 (2013).

[20] H. Shi, C. A. Jiménez-Hoyos, R. Rodríguez-Guzmán, G. E. Scuseria, and S. Zhang, Symmetry-projected wave functions in quantum monte carlo calculations, Phys. Rev. B **89**, 125129 (2014).

[21] Y.-Y. He, M. Qin, H. Shi, Z.-Y. Lu, and S. Zhang, Finite-temperature auxiliary-field quantum monte carlo: Self-consistent constraint and systematic approach to low temperatures, Phys. Rev. B **99**, 045108 (2019).

[22] A. Mahajan, J. H. Thorpe, J. S. Kurian, D. R. Reichman, D. A. Matthews, and S. Sharma, Beyond ccsd(t) accuracy at lower scaling with auxiliary field quantum monte carlo, Journal of Chemical Theory and Computation **21**, 1626 (2025).

[23] C. J. Umrigar, Observations on variational and projector monte carlo methods, The Journal of Chemical Physics **143**, 164105 (2015).

[24] W. Ren, W. Fu, X. Wu, and J. Chen, Towards the ground state of molecules via diffusion monte carlo on neural networks, Nature Communications **14**, 1860 (2023).

[25] Z.-Y. Xiao, T. Xiang, Z. Lu, Y. Chen, and S. Zhang, Implementing advanced trial wave functions in fermion quantum monte carlo via stochastic sampling, The Journal of Chemical Physics **163**, 164109 (2025).

[26] Z.-Y. Xiao, B. Kan, H. Ma, B. Zhao, and H. Shang, Nnqs-afqmc: Neural network quantum states enhanced fermionic quantum monte carlo, Journal of Chemical Theory and Computation **21**, 9587 (2025).

[27] Y. Chen, L. Zhang, W. E, and R. Car, Hybrid auxiliary field quantum monte carlo for molecular systems, Journal of Chemical Theory and Computation **19**, 4484 (2023).

[28] A. Karton, S. Daon, and J. M. Martin, W4-11: A high-confidence benchmark dataset for computational thermochemistry derived from first-principles w4 data, Chemical Physics Letters **510**, 165 (2011).

[29] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, JAX: composable transformations of Python+NumPy programs (2018).

[30] J. Lee, H. Q. Pham, and D. R. Reichman, Twenty years of auxiliary-field quantum monte carlo in quantum chemistry: An overview and assessment on main group chemistry and bond-breaking, Journal of Chemical Theory and Computation **18**, 7024 (2022).

[31] Y. J. Bomble, J. F. Stanton, M. Kállay, and J. Gauss, Coupled-cluster methods including noniterative corrections for quadruple excitations, The Journal of Chemical Physics **123**, 054101 (2005).

[32] Z. Sukurma, M. Schlilpf, M. Humer, A. Taheridehkhordi, and G. Kresse, Benchmark phaseless auxiliary-field quantum monte carlo method for small molecules, Journal of Chemical Theory and Computation **19**, 4921 (2023).

[33] A. Tajti, P. G. Szalay, A. G. Császár, M. Kállay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vázquez, and J. F. Stanton, Heat: High accuracy extrapolated ab initio thermochemistry, The Journal of Chemical Physics **121**, 11599 (2004).

[34] A. Karton, N. Sylvetsky, and J. M. L. Martin, W4-17: A diverse and high-confidence dataset of atomization energies for benchmarking high-level electronic structure methods, Journal of Computational Chemistry **38**, 2063 (2017).

[35] K. T. Williams, Y. Yao, J. Li, L. Chen, H. Shi, M. Motta, C. Niu, U. Ray, S. Guo, R. J. Anderson, J. Li, L. N. Tran, C.-N. Yeh, B. Mussard, S. Sharma, F. Bruneval, M. van Schilfgaarde, G. H. Booth, G. K.-L. Chan, S. Zhang, E. Gull, D. Zgid, A. Millis, C. J. Umrigar, and L. K. Wagner (Simons Collaboration on the Many-Electron Problem), Direct comparison of many-body methods for realistic electronic hamiltonians, Phys. Rev. X **10**, 011041 (2020).

[36] J. Shee, B. Rudshteyn, E. J. Arthur, S. Zhang, D. R. Reichman, and R. A. Friesner, On achieving high accuracy in quantum chemical calculations of 3 d transition metal-containing systems: A comparison of auxiliary-field quantum monte carlo with coupled cluster, density functional theory, and experiment for diatomic molecules, J. Chem. Theory Comput. **15**, 2346 (2019).

[37] Z. Li, Z. Lu, R. Li, X. Wen, X. Li, L. Wang, J. Chen, and W. Ren, Spin-symmetry-enforced solution of the many-body schrödinger equation with a deep neural network, Nature Computational Science **4**, 910 (2024).

[38] P. B. Szabó, Z. Schätzle, M. T. Entwistle, and F. Noé, An improved penalty-based excited-state variational monte carlo approach with deep-learning ansatzes, Journal of Chemical Theory and Computation **20**, 7922 (2024).

[39] T. Jiang, B. O'Gorman, A. Mahajan, and J. Lee, Unbiasing fermionic auxiliary-field quantum monte carlo with matrix product state trial wavefunctions, Phys. Rev. Res. **7**, 013038 (2025).

[40] R. Levy, M. A. Morales, and S. Zhang, Automatic order detection and restoration through systematically improvable variational wave functions, Phys. Rev. Res. **6**, 013237 (2024).

[41] S. Sorella, Systematically improvable mean-field variational ansatz for strongly correlated systems: Application to the hubbard model, Phys. Rev. B **107**, 115133 (2023).