JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

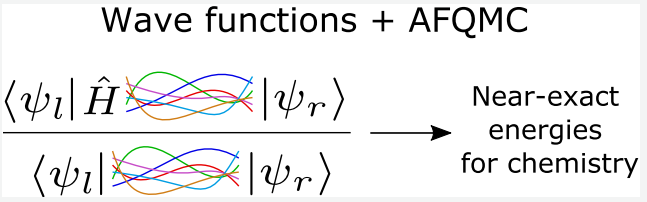
Taming the Sign Problem in Auxiliary-Field Quantum Monte Carlo Using Accurate Wave Functions

Ankit Mahajan* and Sandeep Sharma

Cite This: J. Chem. Theory Comput. 2021, 17, 4786–4798 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: We explore different ways of incorporating accurate trial wave functions into free projection auxiliary-field quantum Monte Carlo (fp-AFQMC). States employed include coupled-cluster singles and doubles, multi-Slater, and symmetry-projected mean-field wave functions. We adapt a recently proposed fast multi-Slater local energy evaluation algorithm for fp-AFQMC, making the use of long expansions from selected configuration interaction methods feasible. We demonstrate how these wave functions serve to mitigate the sign problem and accelerate convergence in quantum chemical problems, allowing the application of fp-AFQMC to systems of substantial sizes. Our calculations on the widely studied model $Cu_2O_2^{2+}$ system show that many previously reported isomerization energies differ substantially from the near-exact fp-AFQMC value.



1. INTRODUCTION

Projector Monte Carlo (PMC) is among the most powerful and versatile methods for calculating the properties of many-fermion systems.[1−4] In this method, one numerically integrates the imaginary-time (τ) Schrödinger equation

$$\frac{\partial|\psi(\tau)\rangle}{\partial\tau} = -\hat{H}|\psi(\tau)\rangle \tag{1}$$

whose solution can formally be written as $|\psi(\tau)\rangle = \exp(-\tau\hat{H})|\psi_r\rangle$, where $|\psi_r\rangle$ is an initial state. Writing the initial state as a linear combination of the Hamiltonian eigenstates ($|\psi_i\rangle$), $|\psi_r\rangle = \sum_i c_i|\psi_i\rangle$, we get

$$|\psi(\tau)\rangle = \sum_i c_i \exp(-E_i\tau)|\psi_i\rangle \tag{2}$$

where E_i are energy eigenvalues. Consequently, as long as the initial state $|\psi_r\rangle$ does not have a vanishing overlap with the ground state ($c_0 \neq 0$), $|\psi(\tau)\rangle$ approaches the ground state exponentially fast in the long τ limit. One can use a trial wave function $|\psi_t\rangle$ to measure various observables in the ground state. In particular, the ground-state energy can be obtained using the expression

$$E(\tau) = \frac{\langle\psi_t|\hat{H}e^{-\tau\hat{H}}|\psi_r\rangle}{\langle\psi_t|e^{-\tau\hat{H}}|\psi_r\rangle} \tag{3}$$

which converges to the exact ground-state energy in the long τ limit. Although PMC is exact in principle, in practice (in the absence of special symmetries in the system), any attempt at stochastic numerical integration is plagued with a severe numerical instability, termed the sign problem, stemming from the antisymmetry of fermionic wave functions.[5] As a result, the signal-to-noise of the PMC simulation decreases exponentially, and after a relatively short imaginary time τ, obtaining useful information from the simulation becomes intractable.

Various flavors of PMC differ in the space in which the simulation is performed, the method used to approximately apply the projector onto a state, and how the sign problem is controlled. Characteristics of four commonly used PMC methods, namely diffusion Monte Carlo (DMC),[3,6] Green's function Monte Carlo (GFMC),[7,8] auxiliary-field QMC (AFQMC),[9−16] and full configuration interaction QMC (FCIQMC),[17,18] are listed in Table 1. These approaches have been developed with different goals in mind based on their respective advantages and pitfalls. The sign problem is

Table 1. Characteristics of Various Commonly Used PMC Approaches[a]

method	space	projector form	approximation
DMC	real	$\exp(-\Delta\tau H)$	fixed-node
GFMC	real	$(1-\Delta\tau H)^{-1}$	fixed-node
GFMC	orbital	$1-\Delta\tau H$	fixed-node
AFQMC	orbital	$\exp(-\Delta\tau H)$	phaseless
FCIQMC	orbital	$1-\Delta\tau H$	initiator

[a]The last column lists the approximation that is commonly used to overcome the sign problem in the respective methods.

Received: April 14, 2021
Published: July 7, 2021

https://doi.org/10.1021/acs.jctc.1c00371
J. Chem. Theory Comput. 2021, 17, 4786–4798

less severe in orbital space methods compared to real-space methods due to the antisymmetry naturally enforced in the Hilbert space. The linear projectors used in orbital space GFMC and FCIQMC have the advantage that they do not have time-step errors that arise in the exponential projectors. The cost of applying the linear projector is also much smaller comparatively. Alternatively, the exponential projector used in AFQMC allows noise-free application of the one-body part of the Hamiltonian, leading to a milder sign problem in AFQMC compared to FCIQMC and GFMC. Regardless of the variation in the severity of the sign problem in different methods, it becomes exponentially worse with system size in all cases, making calculations of large systems very difficult. One of the common ways to deal with it is to use a guiding wave function to enforce a constraint that stabilizes the simulation at the expense of a systematic bias. The phaseless constraint[19] often used in AFQMC (ph-AFQMC), and the fixed-node approximation used in DMC and GFMC lead to polynomially scaling methods, albeit with more or less controlled errors. A different approach is used in FCIQMC, whereby the constraint is applied without an external guiding state using the initiator approximation.[20] Another way of dealing with the sign problem is using transcorrelation, wherein trial wave function information gets folded into the Hamiltonian, making the projection more manageable.[21] A recently proposed approach of combining aspects of variational and projection methods through time-step optimization also seems promising.[22]

In the current work, we will take a different approach. Instead of introducing a bias to stabilize the simulation, we use high-quality states $|\psi_l\rangle$ and $|\psi_r\rangle$ such that the simulation converges quickly enough before the sign problem becomes prohibitively expensive. We will use the exponential projector $\exp(-\Delta\tau H)$ and work in orbital space without enforcing a constraint. Thus, this approach is almost identical to free projection AFQMC (fp-AFQMC). The term free projection was first introduced in ref 19 to distinguish it from the phaseless approach. Due to the very high cost of fp-AFQMC when used with simple wave functions, to date, only a handful of calculations have been reported on systems of substantial sizes, the chromium dimer calculation by Purwanto et al.[23] being notable among them. Looking at eq 3, it is evident that if the states $|\psi_l\rangle$ and $|\psi_r\rangle$ have substantial weight on the ground state and small weights on low-lying excited states, the energy estimator $E(\tau)$ will converge rapidly to the ground-state energy, and one can obtain an accurate estimate of the ground-state energy with a relatively short τ simulation. This approach is analogous to finite-temperature AFQMC simulations with τ playing the role of inverse temperature. It is also closely related to the method presented in ref 24 by Hlubina et al., where a Gutzwiller trial wave function is used to obtain accurate energies in the Hubbard model due to a zero-variance principle. While past efforts to incorporate high-quality wave functions into AFQMC, including those in ph-AFQMC,[25,26] have focused on model lattice systems, in this paper, we consider ab initio systems and wave functions commonly used in quantum chemistry. We study the use of multi-Slater (obtained from selected configuration interaction methods),[27−31] coupled-cluster singles and doubles (CCSD),[32,33] and symmetry-projected mean-field states[34,35] in fp-AFQMC. Our method presents a way to systematically improve upon these wave functions through imaginary-time propagation. In particular, we find that multi-Slater states, in combination with CCSD, give remarkably accurate results for several systems, and one can perform near-exact simulations on systems as large as 52 electrons in 118 orbitals.

The rest of this paper is organized as follows: In Section 2, we begin with the details of the procedure used for sampling the short-time propagator. Then, we present efficient ways of using various wave functions in the estimator. In Section 3, we analyze the efficacy of these states in fp-AFQMC with some illustrative examples. We also gauge the accuracy of focal point corrections to stretch the scope of the method. We present results for larger systems, namely benzene, cyclobutadiene, and $Cu_2O_2^{2+}$. Finally, we conclude with a discussion of the prospects for our approach.

2. THEORY

2.1. Propagator Sampling. Here, we review how the propagator is sampled in AFQMC following the detailed exposition by ref 16. Consider the ab initio electronic Hamiltonian given by

$$\hat{H} = \sum_{ij} h_i^j a_i^\dagger a_j + \frac{1}{2}\sum_{ikjl} v_{ij}^{kl} a_i^\dagger a_k a_j^\dagger a_l = \hat{K} + \hat{V} \tag{4}$$

where i, j, k, and l are spin-orbital indices and h_i^j and v_{ij}^{kl} are one and two electron integrals, respectively. We denote the number of orbitals by M and the number of electrons by N. To sample the imaginary-time propagator, one first divides the propagation into small intervals and makes a Trotter approximation for each interval, given as

$$\begin{aligned} e^{-\tau\hat{H}} &= (e^{-\Delta\tau\hat{H}})^N \\ e^{-\Delta\tau\hat{H}} &= e^{-\Delta\tau\hat{K}/2}e^{-\Delta\tau\hat{V}}e^{-\Delta\tau\hat{K}/2} + O(\Delta\tau^3) \end{aligned} \tag{5}$$

To sample the exponential of the two-body part, we express it as a sum of squares using a modified Cholesky decomposition[36,37]

$$\hat{V} = \frac{1}{2}\sum_\gamma \left(\sum_{jl} [L^\gamma]_j^l a_j^\dagger a_l\right)^2 = -\frac{1}{2}\sum_\gamma \hat{v}_\gamma^2 \tag{6}$$

where $\hat{v}_\gamma = i\sum_{jl}[L^\gamma]_j^l a_j^\dagger a_l$ are one-body operators. We will denote the number of Cholesky matrices, L^γ, by X. Empirically, X is known to scale linearly with the number of orbitals with $X \sim 5-10M$. Using the Hubbard−Stratanovic transform, we get

$$e^{-\Delta\tau\hat{V}} = \prod_\gamma \int \frac{dx_\gamma}{\sqrt{2\pi}} e^{-x_\gamma^2/2} e^{\sqrt{\Delta\tau}x_\gamma\hat{v}_\gamma} + O(\Delta\tau^2) \tag{7}$$

where x_γ are the auxiliary fields. The quadratic Trotter error arises due to the fact that $\hat{v}_\gamma$ do not commute with each other. The propagator on the short interval can be expressed in a compact form as

$$e^{-\Delta\tau\hat{H}} = \int \mathbf{dx}\, p(\mathbf{x})\hat{B}(\mathbf{x}) + O(\Delta\tau^2) \tag{8}$$

where $\mathbf{x}$ is the vector of auxiliary fields, $p(\mathbf{x})$ is the standard Gaussian distribution, and $\hat{B}(\mathbf{x})$ is a one-body operator given by

$$\hat{B}(\mathbf{x}) = e^{-\Delta\tau K/2} e^{\sqrt{\Delta\tau}\sum_\gamma x_\gamma\hat{v}_\gamma} e^{-\Delta\tau K/2} \tag{9}$$

The full propagator can be written as

https://doi.org/10.1021/acs.jctc.1c00371
J. Chem. Theory Comput. 2021, 17, 4786−4798

$$e^{-\tau\hat{H}} = \mathcal{T}\prod_i \int d\mathbf{x}_i\, p(\mathbf{x}_i)\hat{B}(\mathbf{x}_i) = \int d\mathcal{X}\, p(\mathcal{X})\hat{\mathcal{B}}(\mathcal{X}) \tag{10}$$

where $\mathbf{x}_i$ are auxiliary fields at the ith time slice, $\mathcal{T}$ denotes time ordering, $\mathcal{X} = \{\mathbf{x}_i\}$ is the set of auxiliary fields at all time slices, and $\hat{\mathcal{B}}(\mathcal{X}) = \mathcal{T}\prod_i \hat{B}(\mathbf{x}_i)$.

Substituting the propagator integral into the energy expression in eq 3, we get

$$E(\tau) = \frac{\int d\mathcal{X}\, p(\mathcal{X})\langle\psi_l|\hat{H}\hat{\mathcal{B}}(\mathcal{X})|\psi_r\rangle}{\int d\mathcal{X}\, p(\mathcal{X})\langle\psi_l|\hat{\mathcal{B}}(\mathcal{X})|\psi_r\rangle} \tag{11}$$

According to the Thouless theorem, applying the exponential of a one-body operator onto a Slater determinant results in another Slater determinant with rotated orbitals. This allows one to operate the sampled propagator onto a determinant as

$$\hat{\mathcal{B}}(\mathcal{X})|\phi\rangle = |\phi(\mathcal{X})\rangle \tag{12}$$

where $|\phi\rangle$ and $|\phi(\mathcal{X})\rangle$ are Slater determinants. If the initial state is a single determinant, the projector can thus be applied directly to it. However, even when the state $|\psi_r\rangle$ is not a single determinant (e.g., CCSD or Jastrow−Slater wave function), it can be written as

$$|\psi_r\rangle = \int d\phi\, c(\phi)|\phi\rangle \tag{13}$$

where the integral is over an overcomplete set of Slater determinants $|\phi\rangle$, and consequently, the expansion coefficients $c(\phi)$ are not unique. Thus, we have

$$E(\tau) = \frac{\int d\mathcal{X}d\phi\, p(\mathcal{X})c(\phi)\langle\psi_l|\hat{H}\hat{\mathcal{B}}(\mathcal{X})|\phi\rangle}{\int d\mathcal{X}d\phi\, p(\mathcal{X})c(\phi)\langle\psi_l|\hat{\mathcal{B}}(\mathcal{X})|\phi\rangle} \tag{14}$$

Now we can sample multidimensional integrals in the numerator and the denominator using Monte Carlo. The process for directly sampling determinants $|\phi\rangle$ from a CCSD initial state $|\psi_r\rangle$ will be described in Section 2.2. We sample $\mathcal{X}$ directly from the standard Gaussian distribution $p(\mathcal{X})$. So we get the estimator

$$E(\tau) \approx \frac{\sum_i \langle\psi_l|\hat{H}\hat{\mathcal{B}}(\mathcal{X}_i)|\phi_i\rangle}{\sum_i \langle\psi_l|\hat{\mathcal{B}}(\mathcal{X}_i)|\phi_i\rangle} = \frac{\sum_i \langle\psi_l|\hat{H}|\phi_i(\mathcal{X}_i)\rangle}{\sum_i \langle\psi_l|\phi_i(\mathcal{X}_i)\rangle} \tag{15}$$

where $\mathcal{X}_i$ and $|\phi_i\rangle$ are auxiliary field and initial wave function samples, respectively. We note that this estimator of the ratio of two random variables has a bias that goes down as $1/N_{\text{samples}}$, while the statistical noise goes down more slowly as $1/\sqrt{N_{\text{samples}}}$. Therefore, we ignore the more rapidly decaying bias, which we have confirmed to be unimportant in our numerical calculations. In the limit that $|\psi_l\rangle$ is the exact ground state, the estimator in eq 15 becomes noiseless since $\langle\psi_0|\hat{H}|\phi_i(\mathcal{X}_i)\rangle = E_0\langle\psi_0|\phi_i(\mathcal{X}_i)\rangle$. This zero-variance property ensures that more accurate $|\psi_l\rangle$ states lead to lower statistical noise in the estimator. Subtracting the mean-field background from the Hamiltonian is known to substantially reduce statistical fluctuations,[14] and we employ this strategy in our calculations. We also periodically orthogonalize orbitals in the walker, $|\phi_i(\mathcal{X}_i)\rangle$, at fixed intervals during the propagation for numerical stability.

The most expensive operation in propagation is the formation of the one-body operator matrix ($\sum_\gamma x_\gamma \hat{v}_\gamma$ in eq 9), which has a computational cost scaling of $O(XM^2)$. The exponentials of these matrices are not explicitly calculated, instead only their action on the walker matrices, $|\phi_i(\mathcal{X}_i)\rangle$ with dimension $M \times N$, is calculated using a truncated Taylor series, which has a cost scaling of $O(NM^2)$. We retain the first 10 terms in the Taylor series of the exponential.

2.2. Trial and Initial Wave Functions. In this section, we will outline methods for using various accurate states as $|\psi_l\rangle$ and $|\psi_r\rangle$ in eq 3. For using a wave function as $|\psi_l\rangle$, one needs an efficient way to calculate its local energy given by

$$E_L(\phi) = \frac{\langle\psi_l|\hat{H}|\phi\rangle}{\langle\psi_l|\phi\rangle} \tag{16}$$

where $|\phi\rangle$ is a Slater determinant. In terms of local energy, eq 15 is written as

$$E(\tau) \approx \frac{\sum_i \langle\psi_l|\phi_i(\mathcal{X}_i)\rangle\, E_L(\phi_i(\mathcal{X}_i))}{\sum_i \langle\psi_l|\phi_i(\mathcal{X}_i)\rangle} \tag{17}$$

For accurate trial states and larger systems, local energy evaluation can become more expensive than propagation. Many ways of reducing this cost in AFQMC have been proposed for single determinant and multi-Slater trials.[38−42] Below we will describe economic algorithms for evaluating local energies for multi-Slater and symmetry-projected mean-fields trial states, which can be used as $|\psi_l\rangle$. To use a wave function as the initial state $|\psi_r\rangle$ in the sampling method outlined in Section 2.1, one needs to be able to sample a determinant from it (eq 13). We will present a way to do this for CCSD wave functions.

2.2.1. Multi-Slater. A multi-Slater wave function is obtained by particle−hole excitations from a reference configuration. We can write it as

$$|\psi\rangle = \sum_n^{N_c} c_n \prod_{\mu_n}^{k_n} a^\dagger_{t_{\mu_n}} a_{p_{\mu_n}} |\psi_0\rangle \tag{18}$$

where $|\psi_0\rangle$ is the reference configuration, c_n are real expansion coefficients, N_c is the number of configurations in the expansion, and k_n are the excitation ranks. Note that the product of particle−hole excitations needs not be ordered since the excitations commute with each other.

Multi-Slater wave functions can be employed as $|\psi_l\rangle$ trial states in fp-AFQMC. We recently presented an efficient local energy evaluation algorithm for multi-Slater wave functions in variational Monte Carlo[43] that has a favorable scaling with the number of configurations, analogous to developments in real-space QMC algorithms.[44] Here, this algorithm is adapted for AFQMC with some important changes. For convenience, we write the local energy as

$$E_L[\phi] = \frac{\langle\psi|\hat{H}|\phi\rangle}{\langle\psi_0|\phi\rangle} \Big/ \frac{\langle\psi|\phi\rangle}{\langle\psi_0|\phi\rangle} \tag{19}$$

The overlap ratio in the denominator is given by

$$\frac{\langle\psi|\phi\rangle}{\langle\psi_0|\phi\rangle} = \sum_n c_n \frac{\langle\psi_0|\prod_{\mu_n} a^\dagger_{p_{\mu_n}} a_{t_{\mu_n}}|\phi\rangle}{\langle\psi_0|\phi\rangle} \tag{20}$$

Since $|\psi_0\rangle$ and $|\phi\rangle$ are determinants, we can use the generalized Wick's theorem to compute the terms in this sum. The Green's function is given as[16]

https://doi.org/10.1021/acs.jctc.1c00371
J. Chem. Theory Comput. 2021, 17, 4786−4798

$$G_j^i = \frac{\langle \psi_0 | a_i^\dagger a_j | \phi \rangle}{\langle \psi_0 | \phi \rangle} = [\phi(\psi_0^\dagger \phi)^{-1} \psi_0^\dagger]_i^j \tag{21}$$

where ϕ and ψ_0 are the orbital coefficient matrices of the corresponding determinants. Using the generalized Wick's theorem, we get

$$\frac{\langle \psi_0 | \prod_\mu a_{p_\mu}^\dagger a_{t_\mu} | \phi \rangle}{\langle \psi_0 | \phi \rangle} = \det(G_{\{t_\mu\}}^{\{p_\mu\}}) \tag{22}$$

where the sets of indices $\{p_\mu\}$ and $\{t_\mu\}$ denote the $k \times k$ slice of the G matrix, k being the rank of the excitation. Therefore, once the Green's function is calculated at cost $O(N^2M)$ the overlap ratio in eq 20 can be calculated at cost $O(N_c)$. We will ignore the scaling factors with respect to the excitation rank k as it is usually small, typically $(k \leq 6)$.

To calculate the numerator in the local energy expression of eq 19, we first rewrite the Hamiltonian as

$$\hat{H} = \sum_{ij} [h']_j^i a_i^\dagger a_j + \frac{1}{2} \sum_{ikjl} \sum_\gamma [L^\gamma]_i^k [L^\gamma]_j^l a_i^\dagger a_j^\dagger a_k a_l \tag{23}$$

where we have normal-ordered the two-body term in eq 4 and absorbed the resulting one-body terms into h'. Note that we have used the same orbitals as those used in the multi-Slater trial wave function to express the Hamiltonian. The one-body part of the local energy numerator is given by

$$\frac{\langle \psi | \hat{H}_1 | \phi \rangle}{\langle \psi_0 | \phi \rangle} = \sum_n c_n \left(\sum_{ij} [h']_i^j \frac{\langle \psi_0 | \left(\prod_{\mu_n} a_{p_{\mu_n}}^\dagger a_{t_{\mu_n}} \right) a_i^\dagger a_j | \phi \rangle}{\langle \psi_0 | \phi \rangle} \right) \tag{24}$$

Using the generalized Wick's theorem, we get

$$\frac{\langle \psi_0 | \left(\prod_\mu a_{p_\mu}^\dagger a_{t_\mu} \right) a_i^\dagger a_j | \phi \rangle}{\langle \psi_0 | \phi \rangle} = \det \begin{pmatrix} G_j^i & \mathcal{G}_{\{t_\mu\}}^i \\ G_j^{\{p_\mu\}} & G_{\{t_\mu\}}^{\{p_\mu\}} \end{pmatrix} \tag{25}$$

where the matrix on the RHS is written as a block matrix with the blocks given by the specified slices. We have also defined a modified Green's function $\mathcal{G}$ as

$$\mathcal{G}_j^i = G_j^i - \delta_j^i \tag{26}$$

where δ is the Kronecker delta. A naive evaluation of the sums in eq 24 entails contracting over the one-body Hamiltonian separately for each configuration in the multi-Slater expansion. This explicit evaluation of the double sum can be avoided using the following trick. Substituting the matrix element in eq 25 into eq 24 and dropping the configuration index n for convenience, we get

$$\begin{aligned} \sum_{ij} [h']_i^j \det \begin{pmatrix} G_j^i & \mathcal{G}_{\{t_\mu\}}^i \\ G_j^{\{p_\mu\}} & G_{\{t_\mu\}}^{\{p_\mu\}} \end{pmatrix} &= \left(\sum_{ij} [h']_i^j G_j^i \right) \det(G_{\{t_\mu\}}^{\{p_\mu\}}) \\ &+ \sum_{\nu=1}^{k} (-1)^\nu \det \begin{pmatrix} \sum_{ij} [h']_i^j G_j^{p_\nu} \mathcal{G}_{\{t_\mu\}}^i \\ G_{\{t_\mu\}}^{\{p_\mu\} \setminus p_\nu} \end{pmatrix} \end{aligned} \tag{27}$$

where we have Laplace expanded the determinant on the LHS along the first column. $\{p_\mu\} \setminus p_\nu$ denotes the set of indices $\{p_\mu\}$ excluding p_ν. This equation suggests a way of separating the sums over Hamiltonian and configuration indices by precomputing the intermediates given by

$$\begin{aligned} E_0^1 &= \sum_{ij} [h']_i^j G_j^i \\ S_t^p &= \sum_{ij} [h']_i^j G_j^p \mathcal{G}_t^i \end{aligned} \tag{28}$$

Using these intermediates in eq 27, we get

$$\begin{aligned} \sum_{ij} [h']_i^j \det \begin{pmatrix} G_j^i & \mathcal{G}_{\{t_\mu\}}^i \\ G_j^{\{p_\mu\}} & G_{\{t_\mu\}}^{\{p_\mu\}} \end{pmatrix} &= E_0^1 \det(G_{\{t_\mu\}}^{\{p_\mu\}}) + \sum_{\nu=1}^{k} (-1)^\nu \det \begin{pmatrix} S_{\{t_\mu\}}^{p_\nu} \\ G_{\{t_\mu\}}^{\{p_\mu\} \setminus p_\nu} \end{pmatrix} \end{aligned} \tag{29}$$

where the RHS now does not involve any contractions over the one-body Hamiltonian indices.

The two-body contribution to the local energy numerator is given by

$$\frac{\langle \psi | \hat{H}_2 | \phi \rangle}{\langle \psi_0 | \phi \rangle} = \sum_n c_n E_n \tag{30}$$

where

$$E_n = \sum_\gamma \sum_{ikjl} [L^\gamma]_i^k [L^\gamma]_j^l \frac{\langle \psi_0 | \left(\prod_{\mu_n} a_{p_{\mu_n}}^\dagger a_{t_{\mu_n}} \right) a_i^\dagger a_j^\dagger a_k a_l | \phi \rangle}{\langle \psi_0 | \phi \rangle} \tag{31}$$

Again using Wick's theorem, we find

$$\frac{\langle \psi_0 | \left(\prod_{\mu_n} a_{p_{\mu_n}}^\dagger a_{t_{\mu_n}} \right) a_i^\dagger a_j^\dagger a_k a_l | \phi \rangle}{\langle \psi_0 | \phi \rangle} = \det \begin{pmatrix} G_{\{k,l\}}^{\{i,j\}} & \mathcal{G}_{\{t_\mu\}}^{\{i,j\}} \\ G_{\{k,l\}}^{\{p_\mu\}} & G_{\{t_\mu\}}^{\{p_\mu\}} \end{pmatrix} \tag{32}$$

By following a procedure analogous to the one-body case, it is possible to avoid the expensive explicit evaluation of the sums in eq 31, by precomputing the following intermediates

https://doi.org/10.1021/acs.jctc.1c00371
J. Chem. Theory Comput. 2021, 17, 4786−4798

$$
\begin{aligned}
E_0^2 &= \sum_{\gamma} \sum_{ikjl} [L^{\gamma}]_i^k [L^{\gamma}]_j^l \det(G_{\{k,l\}}^{\{i,j\}}) \\
[D_1]_t^p &= \sum_{\gamma} \sum_{ikjl} [L^{\gamma}]_i^k [L^{\gamma}]_j^l \det(G_{\{k,l\}}^{\{i,p\}}) \mathcal{G}_t^j, \\
[D_2^{\gamma}]_t^p &= \sum_{ik} [L^{\gamma}]_i^k G_k^p \mathcal{G}_t^i
\end{aligned} \tag{33}
$$

Using these intermediates along with eqs 31 and 32, dropping the configuration index n for convenience, we get

$$
\begin{aligned}
&\sum_{\gamma ikjl} [L^{\gamma}]_i^k [L^{\gamma}]_j^l \det\begin{pmatrix} G_{\{k,l\}}^{\{i,j\}} & \mathcal{G}_{\{t_\mu\}}^{\{i,j\}} \\ G_{\{k,l\}}^{\{p_\mu\}} & G_{\{t_\mu\}}^{\{p_\mu\}} \end{pmatrix} \\
&= E_0^2 \det(G_{\{t_\mu\}}^{\{p_\mu\}}) + \sum_{\nu=1}^{k} (-1)^{\nu} \det\begin{pmatrix} [D_1]_{\{t_\mu\}}^{p_\nu} \\ G_{\{t_\mu\}}^{\{p_\mu\}\ p_\nu} \end{pmatrix} \\
&\quad + \sum_{\gamma} \sum_{\nu_1,\nu_2,\lambda_1,\lambda_2}{}' (-1)^{\nu_1+\nu_2+\lambda_1+\lambda_2} \times \det([D_2^{\gamma}]_{\{t_{\lambda_1},t_{\lambda_2}\}}^{\{p_{\nu_1},p_{\nu_2}\}}) \\
&\quad \det(G_{\{t_\mu\}\ \{t_{\lambda_1},t_{\lambda_2}\}}^{\{p_\mu\}\ \{p_{\nu_1},p_{\nu_2}\}})
\end{aligned} \tag{34}
$$

where we have Laplace expanded the determinant on the LHS along the first two columns. The primed sum indicates the conditions $\nu_1 \neq \nu_2$ and $\lambda_1 \neq \lambda_2$, and all of these indices can take values from 1 to k. Note that the RHS only involves a sum over the Hamiltonian index γ, while all other Hamiltonian indices are contracted in the separate calculation of the intermediates. Thus, we have partially separated the sums over configuration and Hamiltonian indices.

Now, we consider the cost scaling of this algorithm. We will only look at the more expensive two-body contribution. The cost of calculating G scales as $O(N^2M)$. The intermediate D_1 in eq 33 can be calculated at cost $O(XNM^2)$. The cost of calculating each D_2^{γ} is $O(NM^2)$; thus, the cost of calculating it for all γ is also $O(XNM^2)$. Given these intermediates, the contribution of each determinant in the expansion to the local energy, as given by eq 34, has a cost scaling of $O(X)$ arising from the last term in the equation. Therefore, the total cost scaling is given by $O(XNM^2 + XN_c)$. We note that this scaling is different from that of the algorithm we reported previously in ref 43. It is possible to use our previous algorithm in fp-AFQMC, which amounts to combining D_2^{γ} intermediates to build a four-index tensor intermediate, resulting in a total cost scaling of $O(XN^2M^2 + N_c)$. While this algorithm has a better scaling with N_c, for the multi-Slater expansions employed in this study with $N_c \sim 10^4$, we found it to be slower than the algorithm detailed above in most cases because of the large cost of building the four-index intermediate. Finally, if the multi-Slater expansion is restricted to an active orbital space of size A, the algorithm cost scales as $O(XNAM + XN_c)$.

2.2.2. Symmetry-Projected Mean-Field States. Symmetry-projected mean-field states are given by

$$
|\psi\rangle = \prod_i \hat{P}_i |\psi_0\rangle \tag{35}
$$

where $\hat{P}_i$ denotes the projector onto an eigenstate of the operator for symmetry i and $|\psi_0\rangle$ is a broken symmetry mean-field state. When such states are variationally optimized in the presence of the projectors, they yield more accurate approximations of the desired energy eigenstate at a mean-field level compared to bare symmetry broken states without projectors. The state $|\psi_0\rangle$ can be allowed to break different symmetries of the molecular Hamiltonian like spin, complex conjugation, point group, and number symmetry simultaneously. Here, we consider the breaking and restoration of spin (S) and complex conjugation (K) symmetries resulting in a KSGHF state. Details of optimizing such states can be found in ref 34.

To use KSGHF as the $|\psi_l\rangle$ trial state in fp-AFQMC calculations, we consider the local energy given by

$$
E_L[\phi] = \frac{\langle \psi_0 | \hat{P}_K \hat{P}_S \hat{H} | \phi \rangle}{\langle \psi_0 | \hat{P}_K \hat{P}_S | \phi \rangle} \tag{36}
$$

Note that $\hat{P}_S$ commutes with the Hamiltonian. If determinants sampled from $|\psi_r\rangle$ are closed shell (RHF-like) or high-spin open shell (ROHF-like), the propagator sampling described in Section 2.1 preserves their good spin quantum numbers. So, in these cases, we have

$$
\hat{P}_S |\phi\rangle = |\phi\rangle \tag{37}
$$

This allows us to ignore the spin projector completely because of its trivial action on the walker. If a sampling of $|\psi_r\rangle$ were used that produced UHF-like determinants, one would have to perform the spin projection explicitly, but we will not consider this case here. The complex conjugation symmetry projection is performed as

$$
E_L[\phi] = \frac{\langle \psi_0 | \hat{H} | \phi \rangle + \langle \psi_0^* | \hat{H} | \phi \rangle}{\langle \psi_0 | \phi \rangle + \langle \psi_0^* | \phi \rangle} \tag{38}
$$

where $|\psi_0^*\rangle$ is obtained by complex conjugating the orbitals of $|\psi_0\rangle$. We use this same projector when optimizing the KSGHF state. Evaluation of local energy now proceeds similarly to the RHF case with a cost scaling of $O(XN^2M)$.

2.2.3. CCSD. The coupled-cluster wave function ansatz is given by

$$
\begin{aligned}
|\psi\rangle &= \exp(\hat{T}_1 + \hat{T}_2)|\psi_0\rangle \\
&= \exp\left(\sum_{\mu} t_{\mu} \hat{E}_{\mu} + \sum_{\mu,\nu} t_{\mu\nu} \hat{E}_{\mu} \hat{E}_{\nu}\right)|\psi_0\rangle
\end{aligned} \tag{39}
$$

where $|\psi_0\rangle$ is a reference Slater determinant and $\hat{E}_{\mu}$ are particle−hole excitation operators given by

$$
\hat{E}_{\mu} = a_t^{\dagger} a_p \tag{40}
$$

$\mu = \{t, p\}$ is a composite index for a pair of empty and occupied orbitals in the reference determinant. t_{μ} and $t_{\mu\nu}$ are singles and doubles amplitudes, respectively. To use a CCSD wave function as $|\psi_r\rangle$, we sample determinants from it using a Hubbard−Stratanovic transform in a manner analogous to the imaginary-time propagator. Since the particle−hole excitation operators commute with each other, we have

$$
\exp(\hat{T}_1 + \hat{T}_2)|\psi_0\rangle = \exp(\hat{T}_2)|\tilde{\psi}_0\rangle \tag{41}
$$

https://doi.org/10.1021/acs.jctc.1c00371
J. Chem. Theory Comput. 2021, 17, 4786−4798

where, due to Thouless' theorem, $|\tilde{\psi}_0\rangle = \exp(\hat{T}_1)\,|\psi_0\rangle$ is a Slater determinant. Note that unlike the imaginary-time propagator, here we can split the exponential exactly without incurring a Trotter error. Now, we write $\hat{T}_2$ as a sum of squares by diagonalizing the double amplitudes

$$\hat{T}_2 = \sum_{\mu,\nu} t_{\mu\nu}\hat{E}_\mu\hat{E}_\nu = \sum_i \hat{\mathcal{E}}_i^2 \tag{42}$$

where the one-body operators $\hat{\mathcal{E}}_i$ are defined as

$$\hat{\mathcal{E}}_i = \sqrt{\lambda_i}\sum_\mu V_i^\mu \hat{E}_\mu \tag{43}$$

Here, λ_i and V are the eigenvalues and eigenvectors of the doubles amplitudes $t_{\mu\nu}$, respectively. Using the Hubbard–Stratanovic transform, we get

$$\exp\left(\sum_i \hat{E}_i^2\right) = \prod_i \int \frac{dx_i}{\sqrt{2\pi}} e^{-x_i^2/2} e^{\sqrt{2}x_i\hat{E}_i} \tag{44}$$

Having written the exponential of a two-body operator as an integral over exponentials of one-body operators, we can sample its action upon $|\tilde{\psi}_0\rangle$ using Monte Carlo. We again employ direct sampling of the multidimensional Gaussian to obtain determinant samples (see Section 2.1). In this sense, the CCSD operator can be thought of as performing a long-time propagation of the reference determinant in one step without a Trotter error. Computational cost considerations for sampling the CCSD wave function this way are similar to those for sampling the propagator. We note that this way of sampling CCSD wave functions is similar to the treatment of Jastrow factors in ref 45.

3. RESULTS

In this section, we present and analyze the results of fp-AFQMC calculations on various systems. We used PySCF[46] to obtain molecular electronic integrals and optimized CCSD wave functions. The modified Cholesky decompositions were performed using a script in the AFQMC code PAUXY.[47] HCI calculations were performed with the code Dice.[48,49] All fp-AFQMC calculations were performed with a code that we have made publicly available on GitHub..[50] Input and output files from all large calculations can also be found in a public repository.[51] Additional details of the calculations have been provided in the Supporting information.

Before presenting the results, we make a note of some practical details related to fp-AFQMC calculations. The amount of time required to achieve convergence depends on the gap to low-lying excited states and the convergence threshold. We estimated convergence by performing energy calculations at multiple points during the projection, and energies within error bars of each other at 2−3 consecutive substantially separated times were taken to be converged. All fp-AFQMC calculations have a Trotter error that can be made arbitrarily small by reducing the time step. There are various ways of handling this error common to almost all projection QMC methods. We assessed the amount of Trotter error by performing pilot calculations with different time steps for a fraction of the propagation time required for convergence. Based on these small calculations, conservative time steps were chosen. We employed time steps ranging from 0.01 to 0.1 H^{-1} in this study for different problems.

3.1. Illustrative Examples.

3.1.1. Energy Convergence for Different Trial and Initial States. In this section, we analyze the convergence of the energy estimator in eq 3 for different combinations of the trial and initial wave functions discussed above. While all such estimators should converge to the ground-state energy eventually regardless of the choice of trial states, they differ considerably in the amount of computational effort expended. To understand the convergence behavior of the function $E(\tau)$ in eq 3 with τ, let us denote the spectral decomposition of the trial and initial states in the basis of energy eigenstates as

$$|\psi_l\rangle = \sum_i l_i|\psi_i\rangle,\quad |\psi_r\rangle = \sum_i r_i|\psi_i\rangle \tag{45}$$

where $H|\psi_i\rangle = E_i|\psi_i\rangle$. Substituting into eq 3, we get

$$E(\tau) = \frac{\sum_i E_i e^{-\tau E_i} l_i^* r_i}{\sum_i e^{-\tau E_i} l_i^* r_i} \tag{46}$$

From this expression, it is evident that the convergence is dictated by the energy gap to the low-lying excited states in the symmetry sector of the ground state and the coefficients of these states in the spectral decompositions. Since imaginary-time propagation is efficient at filtering out higher-lying excited states, it is best to project out lower-lying states in the initial and trial wave functions as much as possible. If the problem admits an active space overlapping substantially with low-lying states, compact wave functions restricted to such a subspace represent a natural choice to accomplish this.

Figure 1 shows the convergence behavior in a stretched N_2 molecule with a bond length of d = 3 Bohr using the cc-pVDZ

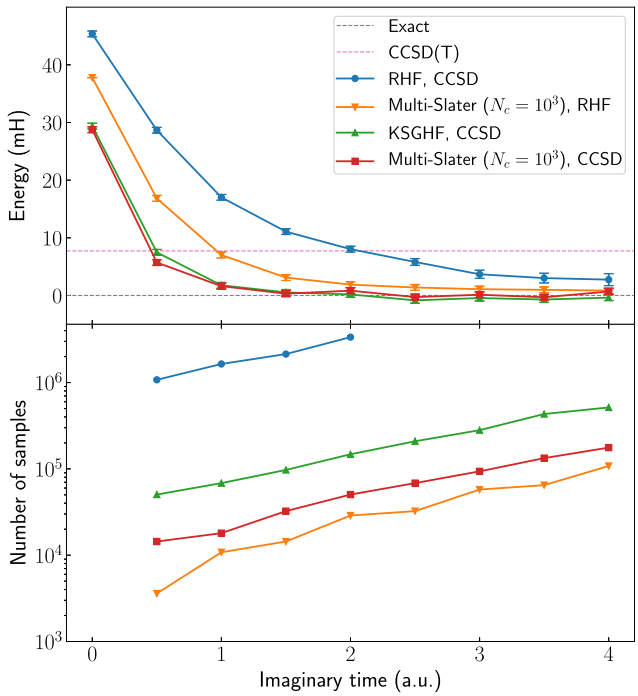


Figure 1. Convergence of the energy estimator for N_2 (d = 3 Bohr) in cc-pVDZ basis for different trial and initial state combinations denoted as $wave_1$ and $wave_2$, respectively. For example, "RHF, CCSD" indicates the estimator using an RHF state as $|\psi_l\rangle$ and a CCSD state as $|\psi_r\rangle$. The bottom plot shows the number of samples required to get a stochastic error less than 0.5 mH as a function of imaginary time.

https://doi.org/10.1021/acs.jctc.1c00371
J. Chem. Theory Comput. 2021, 17, 4786−4798

basis set. We correlated all 14 electrons in the 28 orbitals of this basis set for all calculations. At this geometry, the electronic structure is fairly multireference in character, as evidenced by the substantial error in the CCSD(T) energy. Despite the shortcomings of the CCSD wave function, in this case, it still improves the convergence of energy significantly compared to RHF. Similarly, both multi-Slater and KSGHF wave functions used as $|\psi_l\rangle$ trials lead to faster energy convergence than RHF. For all combinations shown in the plot, the estimators are variational throughout the propagation. Figure 2 shows that when a UHF state is used along with an

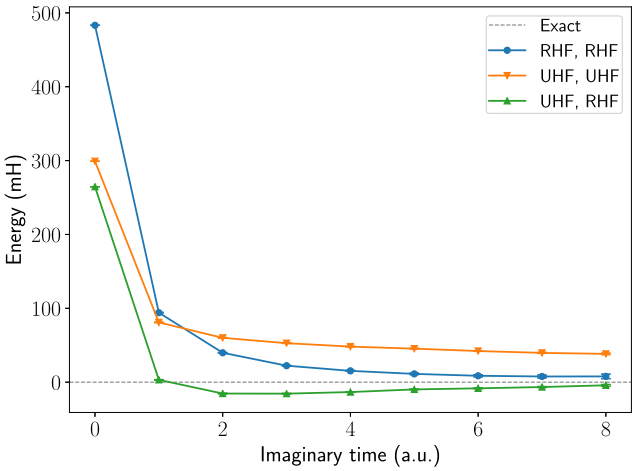


Figure 2. Slow or nonvariational convergence of the energy estimator for different trial and initial state combinations. System and other details are the same as in Figure 1.

RHF state in a mixed estimator, the resulting energy becomes nonvariational after some propagation and eventually converges to the ground-state energy slowly. Curiously, the variational estimator using UHF for both $|\psi_l\rangle$ and $|\psi_r\rangle$ converges rather slowly compared to the RHF variational estimator, even though the UHF energy is lower at $\tau = 0$. Both these observations can be explained by noting that low-lying excited states contribute significantly in the spectral decomposition of the UHF state due to spin contamination, as has been pointed out previously.[52] Empirically, we often found the estimator to be nonvariational when using a symmetry broken wave function as one of the trial or initial states. While such nonvariational estimators could still be employed in fp-AFQMC calculations, determining convergence could become tricky for larger systems. Thus, we do not use symmetry broken states like UHF or UCCSD with a UHF reference in this study.

Another important consideration is the statistical efficiency afforded by various trial states. The bottom panel of Figure 1 shows the number of samples required to get a stochastic error of less than 0.5 mH at different imaginary times. In all cases, the number of samples increases roughly exponentially with imaginary time because of the sign problem, but the estimators differ considerably in the number required to get a fixed stochastic error. As discussed in Section 2.1, based on the zero-variance principle (eq 15), statistical efficiency is dictated by the quality of the $|\psi_l\rangle$ state. This principle is evident from the plot, with the estimator using the RHF state as $|\psi_l\rangle$ requiring a large number of samples. The multi-Slater state, with the leading 10^3 configurations taken from an HCI wave function, is the most accurate and, thus, yields the most statistically efficient estimators. Note that the "multi-Slater, RHF" estimator requires fewer samples to get a fixed statistical error compared to the "multi-Slater, CCSD" estimator for the *same* imaginary time. However, since the CCSD estimator converges much faster than the RHF one, the correct comparison to make is between the number of samples at times when both estimators have the same energy. When this is taken into account, the CCSD estimator requires fewer samples. In addition, since the cost of sampling the CCSD wave function is equivalent to just one imaginary-time propagation step, it also has a lesser overall computational cost compared to the RHF estimator, which requires many propagation steps to reach the same energy.

3.1.2. Multi-Slater Trials with Different Number of Configurations. Multi-Slater trial states can be systematically made more accurate by adding more configurations. In this section, we analyze the cost and benefits of using longer expansions. Figure 3 shows energy convergence with imaginary

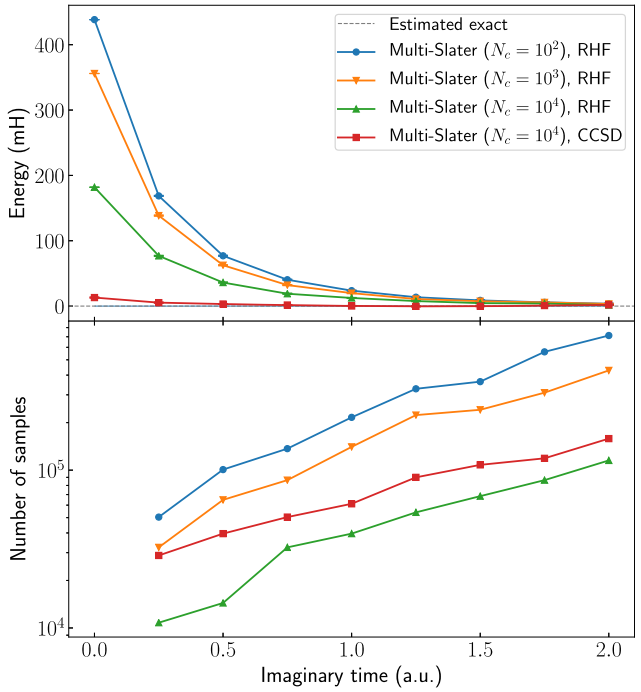


Figure 3. Convergence of the energy estimator for two water molecules at equilibrium geometries in aug-cc-pVDZ basis for different number of configurations. Other details are the same as in Figure 1.

time for two water molecules at equilibrium geometries (provided in the Supporting Information) in the aug-cc-pVDZ basis set. The 1s electrons on oxygen atoms were frozen in all calculations correlating the remaining 16 electrons in 80 orbitals. In contrast to the N_2 example in Section 3.1.1, correlation in this system is primarily dynamic. The main advantage of adding more configurations to the trial state $|\psi_l\rangle$ is to increase the sampling efficiency, as seen from the bottom panel of the figure. It also slightly reduces the projection time required to reach convergence. We have included the curve for the estimator using a CCSD state as $|\psi_r\rangle$ to demonstrate the efficacy of this state in accelerating convergence particularly in problems dominated by dynamic correlation.

https://doi.org/10.1021/acs.jctc.1c00371
J. Chem. Theory Comput. 2021, 17, 4786−4798

In Figure 4, we compare the computational cost scaling of the naive local energy algorithm with that presented in Section 2.2.1 for the same two water molecule system. The naive algorithm effort is estimated by simply multiplying the cost of local energy calculation for a single RHF determinant by the number of configurations. While the cost of calculating local energy for a single determinant includes the cost of evaluating the Green's function, which need not be evaluated for every configuration from scratch, it is a small fraction of the much more expensive operation of contracting the Green's function with Hamiltonian integrals. For less than 10^3 configurations, the cost of the efficient algorithm is seen to increase sublinearly with the number of configurations. In this regime, the cost is dominated by the construction of intermediates scaling as $O(XNM^2)$. Beyond this point, the iteration over configurations begins to dominate, scaling linearly as $O(XN_c)$. To estimate the marginal cost of adding configurations to the trial state, one needs to compare the increase in local energy cost with a reduction in the required number of samples with the associated reduction in propagation costs depending on the system.

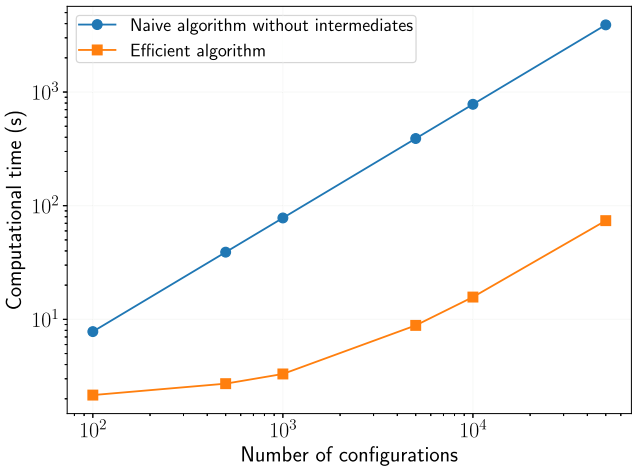


Figure 4. Computational cost scaling of the local energy calculation with the number of configurations (same system as in Figure 3). Computational time is the wall time for calculating 100 local energy samples. "Efficient algorithm" refers to the algorithm described in Section 2.2.1.

3.1.3. Basis Set and Semicore Correlation Focal Point Corrections. Gaussian basis sets used in quantum chemistry are designed to provide a fast convergence for various molecular properties, including energy differences with increasing basis set size. While much of the important correlation is captured in smaller basis sets, larger basis sets are still required to obtain quantitatively accurate properties in some cases. Focal point approaches[53−55] have been developed to yield accurate results close to the continuum limit while avoiding prohibitively expensive calculations with steep-scaling methods in large basis sets. Here, we consider its use to provide basis set and semicore corrections to fp-AFQMC calculations. The basis set correction is given as

$$E(\text{fp} - \text{AFQMC, LB}) \approx E(\text{fp} - \text{AFQMC, SB}) + E(\text{M, LB}) - E(\text{M, SB}) \quad (47)$$

where LB and SB refer to large and small bases, respectively, and M is another method that can be used in the large basis set. We will refer to the resulting method as fp-AFQMC/M. M could be CCSD, CCSD(T), MP2, or any other technique suitable for the problem at hand. The crucial point is that it does not need to be accurate on its own; it should be good enough to capture the effects of increasing the basis size.

Table 2 shows the efficacy of the basis set correction for the N_2 bond-breaking problem. We performed all electron fp-AFQMC ground-state energy calculations of the N_2 molecule for a set of bond lengths in both cc-pVDZ and cc-pVQZ basis sets. A CCSD state was used as the $|\psi_T\rangle$ state, while a multi-Slater wave function was used as $|\psi_I\rangle$. Despite the poor performance of UCCSD(T) in this system, with absolute energy errors of ∼20 mH for some bond lengths, it provides an excellent basis set correction for fp-AFQMC.

Similar to basis set corrections, a semicore correlation correction can be used as

$$E(\text{fp} - \text{AFQMC}, N_v + N_{sc}) \approx E(\text{fp} - \text{AFQMC}, N_v) + E(\text{M}, N_v + N_{sc}) - E(\text{M}, N_v) \quad (48)$$

where N_v is the number of valence electrons and N_{sc} is the number of semicore electrons. This correction is particularly relevant for 3d transition-metal compounds where in addition to valence electrons, the semicore 3s and 3p electrons can play an important role in the chemistry. We note that a CCSD(T) semicore correction has been used along with CASPT2 recently.[56] Table 3 shows the performance of this correction for fp-AFQMC in three transition-metal oxide molecules at equilibrium geometries used in ref 57. We used the scalar relativistic X2C Hamiltonian in a triple-ζ ANO-RCC basis set with a 21s15p10d6f4g/6s5p3d2f1g contraction on metal atoms and a 14s9p4d3f/4s3p2d1f contraction on oxygen. Multi-Slater wave functions were used as $|\psi_I\rangle$ trials. Since these systems are

Table 2. Basis Set Corrections for Ground-State Energies (in H) of N_2 for a Set of Bond Lengths (in Bohr)[a]

	cc-pVDZ		cc-pVQZ		
d	UCCSD(T)	fp-AFQMC	UCCSD(T)	fp-AFQMC	fp-AFQMC/UCCSD(T)
2.0	−109.2694	−109.2706	−109.4592	−109.4610	−109.4603
2.4	−109.2356	−109.2419	−109.4078	−109.4142	−109.4141
2.7	−109.1506	−109.1642	−109.3172	−109.3320	−109.3307
3.0	−109.0689	−109.0897	−109.2311	−109.2543	−109.2519
3.6	−108.9828	−108.9978	−109.1372	−109.1537	−109.1521
4.2	−108.9630	−108.9702	−109.1117	−109.1192	−109.1189

[a]In the last column, we use the notation fp-AFQMC/UCCSD(T) to denote energies obtained by adding the UCCSD(T) cc-pVQZ basis correction to the cc-pVDZ fp-AFQMC energies. Stochastic errors in the QMC calculations are less than 0.8 mH.

https://doi.org/10.1021/acs.jctc.1c00371
J. Chem. Theory Comput. 2021, 17, 4786−4798

Table 3. Semicore Corrections for Ground-State Energies (in H) in Transition-Metal Oxide Molecules[a]

species	method	Ar core	Ne core
CrO	UCCSD(T)	−1115.7928	−1116.1485
	fp-AFQMC	−1115.7953	−1116.1505
	fp-AFQMC/UCCSD(T)		−1116.1510
MnO	UCCSD(T)	−1232.5244	−1232.8777
	fp-AFQMC	−1232.5287	−1232.8815
	fp-AFQMC/UCCSD(T)		−1232.8820
FeO	UCCSD(T)	−1346.7165	−1347.0381
	fp-AFQMC	−1346.7188	−1347.0415
	fp-AFQMC/UCCSD(T)		−1347.0404

[a]Ne and Ar core denote energies with and without semicore correlation, respectively. All calculations were performed with the X2C Hamiltonian in a triple-ζ quality ANO basis set. Stochastic error in QMC calculations is less than 0.7 mH in all cases.

open shell, we used unrestricted CCSD wave functions on top of ROHF references as the $|\psi_r\rangle$ initial states. In all three cases, UCCSD(T) energies have some absolute energy errors, but they still provide reasonable semicore corrections.

Besides the examples presented here, we will check the accuracy of these focal point corrections in larger systems in the following sections. In all of the cases presented in this paper, we have found these corrections to be exceptionally accurate.

3.2. Organic Molecules. *3.2.1. Ground-State Energy of Benzene (C_6H_6).* A recent benchmark study[58] reported the ground-state energy of benzene on the cc-pVDZ basis using various accurate quantum chemistry methods. Here, we calculate fp-AFQMC energies in this basis as well as the cc-pVTZ basis. We used the same geometry as in the benchmark paper and froze the 1s electrons of all carbon atoms at the HF level. This yields correlation spaces of (30e, 108o) for the cc-pVDZ basis and (30e, 258o) for the cc-pVTZ basis. For fp-AFQMC calculations in both basis sets, we used a multi-Slater wave function as $|\psi_l\rangle$ and the CCSD wave function as $|\psi_r\rangle$. In the cc-pVDZ basis calculation, we used canonical HF orbitals to obtain an HCI wave function consisting of about 2.6×10^5 configurations, of which the leading 10^4 were used in the trial state. For the cc-pVTZ basis set, we first performed an HCISCF calculation with a (30e, 100o) active space with a loose $\epsilon_1 = 0.001$ allowing internal rotations. An HCI wave function consisting of about 1.3×10^6 configurations was then obtained in just the active space, and the leading 10^4 configurations were used in the trial state for fp-AFQMC. Table 4 shows the ground-state energies from various methods.

Table 4. Ground-State Energy (in H) of Benzene from Different Methods[a]

method	cc-pVDZ	cc-pVTZ
CCSD(T)	−231.5813	−231.8058
MBE-FCI[58]	−231.5848	
DMRG[58]	−231.5846(7)	
SHCI[58]	−231.586(2)	
ph-AFQMC (RHF)[59]	−231.5879(4)	−231.8122(4)
ph-AFQMC (multi-Slater)[59]	−231.5861(4)	
fp-AFQMC/CCSD(T)		−231.8096(7)
fp-AFQMC	−231.5851(7)	−231.809(1)

[a]Correlation spaces: (30e, 108o) for cc-pVDZ and (30e, 258o) for cc-pVTZ.

The correlation in this system is primarily dynamic in nature and, thus, CCSD(T) recovers most of the correlation energy. All of the high accuracy methods are more or less in agreement within chemical accuracy for the cc-pVDZ basis. In the cc-pVTZ basis, the phaseless approximation error with the RHF trial is about 3 mH. The CCSD(T) basis set correction seems to work well in this case as well. We note that unlike several methods in Table 4, fp-AFQMC energies were obtained without performing any extrapolations.

3.2.2. Automerization of Cyclobutadiene (C_4H_4). Calculating the barrier height for automerization of cyclobutadiene is a challenging problem because of the multireference (biradical) character of the transition state. This system has been used to gauge the accuracy of many multireference methods. We performed fp-AFQMC calculations using geometries provided in ref 60. The minima correspond to a D_{2h} rectangular geometry, while the transition state has a square D_{4h} geometry (Figure 5). Some geometry optimization studies find slightly

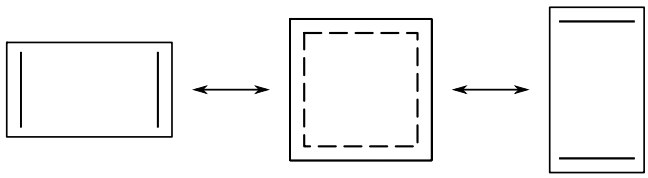


Figure 5. D_{4h} symmetric transition state of cyclobutadiene (middle) between two D_{2h} symmetric degenerate minima (left and right).

bent transition-state structures,[61] but due to the flatness of the energy surface near the transition state, energy differences between different geometries used for transition states in various studies are tiny (less than a mH in most cases). Carbon 1s orbitals were frozen, leading to correlation spaces of size (20e, 72o) and (20e, 172o) in the cc-pVDZ and cc-pVTZ basis sets, respectively. We used the CCSD wave function as the $|\psi_r\rangle$ state in both bases. For $|\psi_l\rangle$, we used multi-Slater wave functions consisting of the order of 10^4 configurations obtained from relatively crude HCI calculations using a procedure similar to that used for benzene calculations. Full details of the wave functions can be found in the Supporting Information.

Table 5 shows barrier heights obtained from different methods. We also provide absolute energies in the Supporting

Table 5. Automerization Barrier Height (kcal/mol) for Cyclobutadiene[a]

method	cc-pVDZ	cc-pVTZ
CCSD(T)	15.8	18.2
CC(t;3)[62]	7.8	10.0
CCSDT[62]	7.6	10.6
TCCSD (2, 2)[60]	9.4	12.9
TCCSD(T) (2, 2)[60]	4.6	7.0
TCCSD (12, 12)[63]		9.2
MR-MkCCSD(T)[64]	7.8	8.9
MRCI+Q[63]		9.2
fp-AFQMC/CCSD(T)		10.9(4)
fp-AFQMC	8.4(4)	10.2(4)
iCAS-CI (6-31+G**basis)[61]	11	
experiment[65]	1.6−10	

[a]Note that different methods used slightly different geometries. Correlation spaces: (20e, 72o) for cc-pVDZ and (20e, 172o) for cc-pVTZ.

https://doi.org/10.1021/acs.jctc.1c00371
J. Chem. Theory Comput. 2021, 17, 4786−4798

Information for reference. The experimental estimate for the barrier spans a large range of 1.6−10 kcal/mol (includes zero-point vibrational energy). CCSD(T) overestimates the gap substantially due to its poor description of the transition state. An iterative treatment of triple excitations in CCSDT or its cheaper approximation CC(t;3)[61] almost entirely corrects this error, signaling the failure of perturbative treatment of triples to describe the biradical transition state. Tailored CCSD (TCCSD)[62] with a small (2e, 2o) active space reference overestimates the barrier height, while the perturbative triples correction overcorrects this error. Recent TCCSD calculations with a (12e, 12o) CAS reference[63] improve upon the previous result. The genuinely multireference Mukherjee's CCSD (MR-MkCCSD(T))[64] yields a barrier height within about 1 kcal/mol of the fp-AFQMC result. Multireference configuration interaction with the Q correction (MRCI+Q) and incremental complete active space configuration interaction (iCAS-CI)[61] also give energy gaps close to the fp-AFQMC value. Finally, despite the poor performance of CCSD(T), it provides a remarkably accurate basis set correction for fp-AFQMC.

3.3. Isomerization of $Cu_2O_2^{2+}$. Enzymes like tyrosinase and catechol oxidase feature a binuclear $Cu_2O_2^{2+}$ active site. The copper atoms serve to activate molecular oxygen and can bind to it in a variety of ways.[66] A particular reaction pathway connecting two such structural isomers shown in Figure 6 has

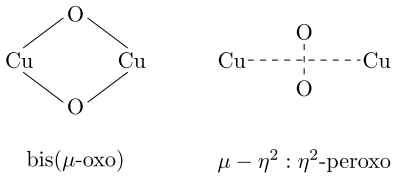


Figure 6. $Cu_2O_2^{2+}$ isomers corresponding to $f = 0$ (left) and $f = 1$ (right).

been studied extensively using several wave function methods and density functional theory (DFT) functionals.[67−75] It is known to be a challenging electronic structure problem with a large spread in calculated isomerization energetics obtained from different techniques. While a thorough study of this system would necessitate the inclusion of solvation and ligand effects, in this paper, we only focus on the core and compare various methods with the results obtained using fp-AFQMC. We also gauge the impact of semicore correlation and scalar relativity on relative energies.

The isomerization pathway is obtained by simply linearly interpolating between the two structures shown in Figure 6 using a single parameter f as described in ref 67. To allow a direct comparison with previous studies, we performed fp-AFQMC calculations using the same double-ζ quality basis that has often been employed, referred to here as BS1. This basis used the Stuttgart pseudopotential and associated basis functions for copper atoms, while an ANO basis with 10s6p3d/4s3p2d contraction for oxygen atoms. We froze the semicore 3s and 3p electrons on copper and the core 1s electrons on oxygen at the HF level leading to a correlation space (32e, 108o). Multi-Slater and CCSD wave functions were used as $|\psi_l\rangle$ and $|\psi_r\rangle$ states, respectively (details in the Supporting Information). The calculated relative energies from fp-AFQMC and a few other methods are shown in Figure 7. Note that all of the methods shown used the same basis set. The correlation space was the same for all techniques except

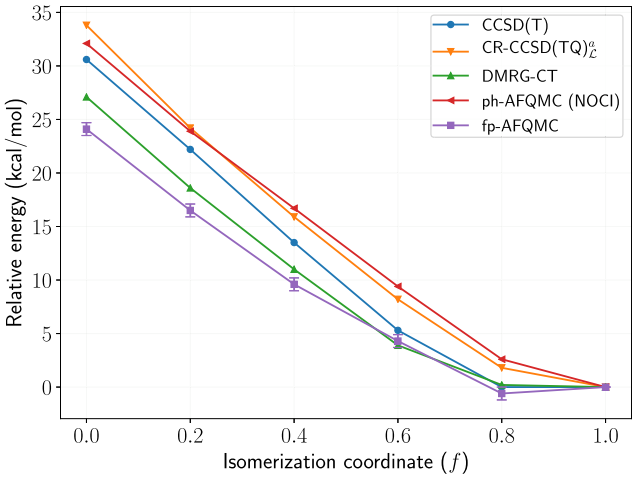


Figure 7. Relative energies along the isomerization curve of $Cu_2O_2^{2+}$.

DMRG-CT (strongly contracted canonical transformation), which froze the same number of orbitals as other methods but at the DMRG-SCF level. We also list the relative energies between the two isomers in Table 6. CCSD(T) overstabilizes the $\mu-\eta^2:\eta^2$-peroxo isomer by about 6.5(6) kcal/mol compared to fp-AFQMC. Looking at the absolute energies (provided in the Supporting Information), this difference is mainly due to CCSD(T) undercorrelating the bis(μ-oxo) isomer by about 8 mH. On the other hand, the perturbative treatment of triples in CCSD(T) seems to provide a good description of the correlation in the $\mu-\eta^2:\eta^2$-peroxo isomer, slightly overcorrelating it, suggesting that this state does not have a significant biradical character. CR-CCSD(TQ)$^a_{\mathcal{L}}$ and ph-AFQMC with a nonorthogonal CI trial both overstabilize the $\mu-\eta^2:\eta^2$-peroxo isomer even further. The DMRG-CT curve is in relatively good agreement with the fp-AFQMC results.

To understand the effect of semicore correlation and the basis set on the relative energies of bis(μ-oxo) and $\mu-\eta^2:\eta^2$-peroxo isomers, we performed additional calculations reported in Table 6. Correlating the copper 3s and 3p semicore electrons and the oxygen 1s electrons, leading to a correlation space of (52e, 118o), substantially increases the gap in BS1. This is also true of CCSD(T) energies. We also performed fp-AFQMC calculations in the bigger triple-ζ ANO-RCC basis set with a 21s15p10d6f4g/6s5p3d2f1g contraction on copper and 14s9p4d3f/4s3p2d1f on oxygen using the X2C scalar relativistic Hamiltonian. We will refer to this basis set as BS2. The fp-AFQMC gap without semicore correlation in BS2 is larger than the BS1 gap. CCSD(T) calculations suggest a much smaller impact of semicore correlations on the isomerization barrier in BS2 compared to BS1. This could be due to an inadequate representation of semicore electrons and scalar relativistic effects in BS1. Previous studies have noted the sizeable influence of scalar relativity on relative energies as well as their poor description in the ECP of BS1.[70,72] CCSD(T) calculations in the quadruple-ζ quality ANO-RCC basis set (BS3) suggest that the relative energy is nearly converged with the basis set size for BS3. The fp-AFQMC/CCSD(T) energy was obtained by adding semicore and basis set corrections to the fp-AFQMC BS3 value and is the best estimate of the relative energy presented here. Nonrelativistic CCSD(T) and strongly contracted n-electron valence perturbation theory (SC-NEVPT2) calculations performed in the quadruple-ζ ANO basis set (termed BS3-NR), also reported in Table 6,

https://doi.org/10.1021/acs.jctc.1c00371
J. Chem. Theory Comput. 2021, 17, 4786−4798

Table 6. Relative Energy $\Delta E = E(f = 0) - E(f = 1)$ (in kcal/mol) for $Cu_2O_2^{2+}$ [a]

basis set type	relativistic effects	method	correlation space	ΔE
Cu: Stuttgart ECP	included in the ECP	CCSD(T)	(32e, 108o)	30.6
O: ANO triple-ζ (BS1)		CR-CCSD(TQ)$_{\mathcal{L}}^{a}$[67]	(32e, 108o)	33.8
		DMRG-CT[71]	(32e, 108o)	27.1
		ph-AFQMC (NOCI)[74]	(32e, 108o)	32.1
		fp-AFQMC	(32e, 108o)	24.1(6)
		CCSD(T)	(52e, 118o)	34.5
		D_{2h}S-UHF [8 det][73]	(52e, 118o)	50[b]
		ph-AFQMC (NOCI)[74]	(52e, 118o)	41.9
		fp-AFQMC	(52e, 118o)	31.0(6)
ANO triple-ζ (BS2)	X2C	CCSD(T)	(32e, 158o)	34.8
		CCSD(T)	(52e, 166o)	33.6
		fp-AFQMC	(32e, 158o)	29(1)
ANO quadruple-ζ (BS3)	X2C	CCSD(T)	(52e, 290o)	33.0
		fp-AFQMC/CCSD(T)	(52e, 290o)	27(1)
ANO quadruple-ζ (BS3-NR)	not included	CCSD(T)	(32e, 280o)	37.3
		CCSD(T)	(52e, 290o)	38.1
		SC-NEVPT2[75]	(48e, 288o)	41.3(8)

[a]Refer to the main text for full descriptions of the basis sets. [b]Estimated from the plot in ref 73.

confirm the importance of including scalar relativity in this system. Relative energies obtained from nonrelativistic calculations are significantly larger than relativistic ones, while semicore correlation effects are relatively small for CCSD(T). Based on this evidence, it seems likely that semicore correlation effects are relatively small in this system provided an appropriate basis set is employed.

4. CONCLUSIONS

We have presented efficient ways to use high-quality wave functions, namely multi-Slater, CCSD, and symmetry-projected mean-field states, in near-exact free projection AFQMC. Appropriate combinations of wave functions reduce the amount of projection time required to achieve convergence and mitigate the sign problem. We analyzed how the trial states affect the statistical efficiency of Monte Carlo sampling using illustrative examples. We also found some encouraging evidence suggesting that focal point basis set and semicore corrections may allow widening the scope of this approach when used in combination with other methods. We provided benchmark energy values for the challenging problems of automerization of cyclobutadiene and isomerization of $Cu_2O_2^{2+}$ in moderately sized basis sets.

The method presented here can be improved in several ways that we plan to pursue in the future. To tackle systems with low-lying excited states that are difficult to project out deterministically using the wave functions presented here, a transcorrelated method with trial states including Jastrow factors may be effective. Another possibility is to search for constraints like the phaseless approximation and assess their performance for different trial states. Our experiments with the systems considered here suggest that direct Gaussian sampling is very efficient for moderately correlated systems. Importance sampling is likely necessary for more strongly correlated systems. A more systematic study of focal point corrections is warranted to assess their scope of applicability. It will also be fruitful to explore alternative ways to extrapolate energies to the basis set limit.

Improvements to the software implementation along the lines of those already introduced for ph-AFQMC are possible. Significant cost reductions can be gained by making the full use of symmetries in the ab initio Hamiltonian.[76] Our current implementation only takes advantage of the spin symmetry. The use of graphical processing units has been suggested to provide drastic speedups in AFQMC calculations.[38,77] Such performance improvements should allow applications to larger systems than those presented here at a fraction of the cost. In addition to performance improvements, we are also interested in looking for ways to calculate properties besides ground-state energies like correlation functions and nuclear forces within the framework presented here.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.1c00371.

Details of wave functions used and absolute fp-AFQMC energies (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Ankit Mahajan − *Department of Chemistry, University of Colorado, Boulder, Colorado 80302, United States;* orcid.org/0000-0002-2138-3798; Email: ankitmahajan76@gmail.com

Author

Sandeep Sharma − *Department of Chemistry, University of Colorado, Boulder, Colorado 80302, United States;* orcid.org/0000-0002-6598-8887

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.1c00371

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The funding for this project was provided by the National Science Foundation through the Grant CHE-1800584. S.S. was also partly supported through the Sloan research fellowship. All calculations were performed on the Blanca and Summit

clusters at CU Boulder. The authors thank Takeshi Yanai for providing the basis set information and Miguel Morales for the ph-AFQMC energies for $Cu_2O_2^{2+}$. The authors also thank Joonho Lee for useful discussions.

REFERENCES

(1) Ceperley, D.; Chester, G. V.; Kalos, M. H. Monte Carlo simulation of a many-fermion study. *Phys. Rev. B* **1977**, *16*, No. 3081.
(2) Nightingale, M. P.; Umrigar, C. J. *Quantum Monte Carlo Methods in Physics and Chemistry*; Springer Science & Business Media, 1998.
(3) Foulkes, W.; Mitas, L.; Needs, R.; Rajagopal, G. Quantum Monte Carlo simulations of solids. *Rev. Mod. Phys.* **2001**, *73*, No. 33.
(4) Becca, F.; Sorella, S. *Quantum Monte Carlo Approaches for Correlated Systems*; Cambridge University Press, 2017.
(5) Troyer, M.; Wiese, U.-J. Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations. *Phys. Rev. Lett.* **2005**, *94*, No. 170201.
(6) Ceperley, D. M.; Alder, B. J. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.* **1980**, *45*, No. 566.
(7) Kalos, M. H.; Levesque, D.; Verlet, L. Helium at zero temperature with hard-sphere and other forces. *Phys. Rev. A* **1974**, *9*, No. 2178.
(8) Trivedi, N.; Ceperley, D. M. Ground-state correlations of quantum antiferromagnets: A Green-function Monte Carlo study. *Phys. Rev. B* **1990**, *41*, No. 4552.
(9) Sugiyama, G.; Koonin, S. Auxiliary field Monte-Carlo for quantum many-body ground states. *Ann. Phys.* **1986**, *168*, 1−26.
(10) White, S. R.; Scalapino, D. J.; Sugar, R. L.; Loh, E.; Gubernatis, J. E.; Scalettar, R. T. Numerical study of the two-dimensional Hubbard model. *Phys. Rev. B* **1989**, *40*, No. 506.
(11) Sorella, S.; Baroni, S.; Car, R.; Parrinello, M. A novel technique for the simulation of interacting fermion systems. *Europhys. Lett.* **1989**, *8*, 663.
(12) Hamann, D.; Fahy, S. Energy measurement in auxiliary-field many-electron calculations. *Phys. Rev. B* **1990**, *41*, No. 11352.
(13) Zhang, S.; Carlson, J.; Gubernatis, J. E. Constrained path Monte Carlo method for fermion ground states. *Phys. Rev. B* **1997**, *55*, No. 7464.
(14) Rom, N.; Charutz, D.; Neuhauser, D. Shifted-contour auxiliary-field Monte Carlo: circumventing the sign difficulty for electronic-structure calculations. *Chem. Phys. Lett.* **1997**, *270*, 382−386.
(15) Al-Saidi, W.; Zhang, S.; Krakauer, H. Auxiliary-field quantum Monte Carlo calculations of molecular systems with a Gaussian basis. *J. Chem. Phys.* **2006**, *124*, No. 224101.
(16) Motta, M.; Zhang, S. Ab initio computations of molecular systems by the auxiliary-field quantum Monte Carlo method. *WIREs Comput. Mol. Sci.* **2018**, *8*, No. e1364.
(17) Booth, G. H.; Thom, A. J. W.; Alavi, A. Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space. *J. Chem. Phys.* **2009**, *131*, No. 054106.
(18) Booth, G. H.; Grüneis, A.; Kresse, G.; Alavi, A. Towards an exact description of electronic wavefunctions in real solids. *Nature* **2013**, *493*, 365.
(19) Zhang, S.; Krakauer, H. Quantum Monte Carlo Method using Phase-Free Random Walks with Slater Determinants. *Phys. Rev. Lett.* **2003**, *90*, No. 136401.
(20) Cleland, D.; Booth, G. H.; Alavi, A. Communications: Survival of the fittest: Accelerating convergence in full configuration-interaction quantum Monte Carlo. *J. Chem. Phys.* **2010**, *132*, No. 041103.
(21) Dobrautz, W.; Luo, H.; Alavi, A. Compact numerical solutions to the two-dimensional repulsive Hubbard model obtained via nonunitary similarity transformations. *Phys. Rev. B* **2019**, *99*, No. 075119.
(22) Sorella, S. The phase diagram of the Hubbard model by Variational Auxiliary Field quantum Monte Carlo. arXiv:2101.07045. https://arxiv.org/abs/2101.07045v1, 2021.
(23) Purwanto, W.; Zhang, S.; Krakauer, H. An auxiliary-field quantum Monte Carlo study of the chromium dimer. *J. Chem. Phys.* **2015**, *142*, No. 064302.
(24) Hlubina, R.; Sorella, S.; Guinea, F. Ferromagnetism in the two dimensional t- t hubbard model at the van hove density. *Phys. Rev. Lett.* **1997**, *78*, No. 1343.
(25) Shi, H.; Jiménez-Hoyos, C. A.; Rodríguez-Guzmán, R.; Scuseria, G. E.; Zhang, S. Symmetry-projected wave functions in quantum Monte Carlo calculations. *Phys. Rev. B* **2014**, *89*, No. 125129.
(26) Wouters, S.; Verstichel, B.; Van Neck, D.; Chan, G. K.-L. Projector quantum Monte Carlo with matrix product states. *Phys. Rev. B* **2014**, *90*, No. 045104.
(27) Huron, B.; Malrieu, J. P.; Rancurel, P. Iterative perturbation calculations of ground and excited state energies from multiconfigurational zerothorder wavefunctions. *J. Chem. Phys.* **1973**, *58*, 5745−5759.
(28) Giner, E.; Scemama, A.; Caffarel, M. Using perturbatively selected configuration interaction in quantum Monte Carlo calculations. *Can. J. Chem.* **2013**, *91*, 879−885.
(29) Evangelista, F. A. Adaptive multiconfigurational wave functions. *J. Chem. Phys.* **2014**, *140*, No. 124114.
(30) Holmes, A. A.; Tubman, N. M.; Umrigar, C. J. Heat-Bath Configuration Interaction: An Efficient Selected Configuration Interaction Algorithm Inspired by Heat-Bath Sampling. *J. Chem. Theory Comput.* **2016**, *12*, 3674−3680.
(31) Tubman, N. M.; Lee, J.; Takeshita, T. Y.; Head-Gordon, M.; Whaley, K. B. A deterministic alternative to the full configuration interaction quantum Monte Carlo method. *J. Chem. Phys.* **2016**, *145*, No. 044112.
(32) Čížek, J. On the correlation problem in atomic and molecular systems. Calculation of wavefunction components in Ursell-type expansion using quantum-field theoretical methods. *J. Chem. Phys.* **1966**, *45*, 4256−4266.
(33) Bartlett, R. J.; Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **2007**, *79*, No. 291.
(34) Scuseria, G. E.; Jiménez-Hoyos, C. A.; Henderson, T. M.; Samanta, K.; Ellis, J. K. Projected quasiparticle theory for molecular electronic structure. *J. Chem. Phys.* **2011**, *135*, No. 124108.
(35) Jiménez-Hoyos, C. A.; Henderson, T. M.; Tsuchimochi, T.; Scuseria, G. E. Projected hartree-fock theory. *J. Chem. Phys.* **2012**, *136*, No. 164109.
(36) Beebe, N. H.; Linderberg, J. Simplifications in the generation and transformation of two-electron integrals in molecular calculations. *Int. J. Quantum Chem.* **1977**, *12*, 683−705.
(37) Koch, H.; de Merás, A. S.; Pedersen, T. B. Reduced scaling in electronic structure calculations using Cholesky decompositions. *J. Chem. Phys.* **2003**, *118*, 9481−9484.
(38) Malone, F. D.; Zhang, S.; Morales, M. A. Overcoming the memory bottleneck in auxiliary field quantum Monte Carlo simulations with interpolative separable density fitting. *J. Chem. Theory Comput.* **2019**, *15*, 256−264.
(39) Shee, J.; Arthur, E. J.; Zhang, S.; Reichman, D. R.; Friesner, R. A. Phaseless auxiliary-field quantum Monte Carlo on graphical processing units. *J. Chem. Theory Comput.* **2018**, *14*, 4109−4121.
(40) Motta, M.; Shee, J.; Zhang, S.; Chan, G. K.-L. Efficient ab initio auxiliary-field quantum Monte Carlo calculations in Gaussian bases via low-rank tensor decomposition. *J. Chem. Theory Comput.* **2019**, *15*, 3510−3521.
(41) Lee, J.; Reichman, D. R. Stochastic resolution-of-the-identity auxiliary-field quantum Monte Carlo: Scaling reduction without overhead. *J. Chem. Phys.* **2020**, *153*, No. 044131.
(42) Shi, H.; Zhang, S. Some recent developments in auxiliary-field quantum Monte Carlo for real materials. *J. Chem. Phys.* **2021**, *154*, No. 024107.
(43) Mahajan, A.; Sharma, S. Efficient local energy evaluation for multi-Slater wave functions in orbital space quantum Monte Carlo. *J. Chem. Phys.* **2020**, *153*, No. 194108.

https://doi.org/10.1021/acs.jctc.1c00371
J. Chem. Theory Comput. 2021, 17, 4786−4798

(44) Assaraf, R.; Moroni, S.; Filippi, C. Optimizing the Energy with Quantum Monte Carlo: A Lower Numerical Scaling for JastrowSlater Expansions. *J. Chem. Theory Comput.* **2017**, *13*, 5273−5281.
(45) Chang, C.-C.; Rubenstein, B. M.; Morales, M. A. Auxiliary-field-based trial wave functions in quantum Monte Carlo calculations. *Phys. Rev. B* **2016**, *94*, No. 235144.
(46) Sun, Q.; Berkelbach, T. C.; Blunt, N. S.; Booth, G. H.; Guo, S.; Li, Z.; Liu, J.; McClain, J. D.; Sayfutyarova, E. R.; Sharma, S.; Wouters, S.; Chan, K.-L. G. PySCF: the Python-based simulations of chemistry framework. *WIREs Comput. Mol. Sci.* **2018**, *8*, No. e1340.
(47) https://github.com/pauxy-qmc/pauxy (Accessed July 1, 2021).
(48) Sharma, S.; Holmes, A. A.; Jeanmairet, G.; Alavi, A.; Umrigar, C. J. Semistochastic Heat-Bath Configuration Interaction Method: Selected Configuration Interaction with Semistochastic Perturbation Theory. *J. Chem. Theory Comput.* **2017**, *13*, 1595−1604.
(49) Smith, J. E.; Mussard, B.; Holmes, A. A.; Sharma, S. Cheap and near exact CASSCF with large active spaces. *J. Chem. Theory Comput.* **2017**, *13*, 5468−5478.
(50) https://github.com/sanshar/VMC/tree/dqmc (Accessed July 1, 2021).
(51) https://github.com/ankit76/fp_afqmc (Accessed July 1, 2021).
(52) Purwanto, W.; Al-Saidi, W.; Krakauer, H.; Zhang, S. Eliminating spin contamination in auxiliary-field quantum Monte Carlo: Realistic potential energy curve of F2. *J. Chem. Phys.* **2008**, *128*, No. 114309.
(53) East, A. L.; Allen, W. D. The heat of formation of NCO. *J. Chem. Phys.* **1993**, *99*, 4638−4650.
(54) Császár, A. G.; Allen, W. D.; Schaefer, H. F., III In pursuit of the ab initio limit for conformational energy prototypes. *J. Chem. Phys.* **1998**, *108*, 9751−9764.
(55) Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. Estimates of the ab initio limit for $\pi-\pi$ interactions: The benzene dimer. *J. Am. Chem. Soc.* **2002**, *124*, 10887−10893.
(56) Phung, Q. M.; Feldt, M.; Harvey, J. N.; Pierloot, K. Toward highly accurate spin state energetics in first-row transition metal complexes: a combined CASPT2/CC approach. *J. Chem. Theory Comput.* **2018**, *14*, 2446−2455.
(57) Williams, K. T.; Yao, Y.; Li, J.; Chen, L.; Shi, H.; Motta, M.; Niu, C.; Ray, U.; Guo, S.; Anderson, R. J.; et al. Direct comparison of many-body methods for realistic electronic Hamiltonians. *Phys. Rev. X* **2020**, *10*, No. 011041.
(58) Eriksen, J. J.; Anderson, T. A.; Deustua, J. E.; Ghanem, K.; Hait, D.; Hoffmann, M. R.; Lee, S.; Levine, D. S.; Magoulas, I.; Shen, J.; et al. The ground state electronic energy of benzene. *J. Phys. Chem. Lett.* **2020**, *11*, 8922−8929.
(59) Lee, J.; Malone, F. D.; Reichman, D. R. The performance of phaseless auxiliary-field quantum Monte Carlo on the ground state electronic energy of benzene. *J. Chem. Phys.* **2020**, *153*, No. 126101.
(60) Lyakh, D. I.; Lotrich, V. F.; Bartlett, R. J. The tailored-CCSD (T) description of the automerization of cyclobutadiene. *Chem. Phys. Lett.* **2011**, *501*, 166−171.
(61) Dang, D.-K.; Zimmerman, P. M. Fully variational incremental CASSCF. *J. Chem. Phys.* **2021**, *154*, No. 014105.
(62) Shen, J.; Piecuch, P. Combining active-space coupled-cluster methods with moment energy corrections via the CC (P; Q) methodology, with benchmark calculations for biradical transition states. *J. Chem. Phys.* **2012**, *136*, No. 144104.
(63) Vitale, E.; Alavi, A.; Kats, D. FCIQMC-tailored distinguishable cluster approach. *J. Chem. Theory Comput.* **2020**, *16*, 5621−5634.
(64) Bhaskaran-Nair, K.; Demel, O.; Pittner, J. Multireference state-specific Mukherjees coupled cluster method with noniterative triexcitations. *J. Chem. Phys.* **2008**, *129*, No. 184105.
(65) Whitman, D. W.; Carpenter, B. K. Limits on the activation parameters for automerization of cyclobutadiene-1, 2-d2. *J. Am. Chem. Soc.* **1982**, *104*, 6473−6474.
(66) Solomon, E. I.; Baldwin, M. J.; Lowery, M. D. Electronic structures of active sites in copper proteins: contributions to reactivity. *Chem. Rev.* **1992**, *92*, 521−542.
(67) Cramer, C. J.; Włoch, M.; Piecuch, P.; Puzzarini, C.; Gagliardi, L. Theoretical models on the Cu_2O_2 torture track: Mechanistic implications for oxytyrosinase and small-molecule analogues. *J. Phys. Chem. A* **2006**, *110*, 1991−2004.
(68) Malmqvist, P. Å.; Pierloot, K.; Shahi, A. R. M.; Cramer, C. J.; Gagliardi, L. The restricted active space followed by second-order perturbation theory method: Theory and application to the study of CuO_2 and Cu_2O_2 systems. *J. Chem. Phys.* **2008**, *128*, No. 204109.
(69) Marti, K. H.; Ondík, I. M.; Moritz, G.; Reiher, M. Density matrix renormalization group calculations on relative energies of transition metal complexes and clusters. *J. Chem. Phys.* **2008**, *128*, No. 014104.
(70) Kurashige, Y.; Yanai, T. High-performance ab initio density matrix renormalization group method: Applicability to large-scale multireference problems for metal compounds. *J. Chem. Phys.* **2009**, *130*, No. 234114.
(71) Yanai, T.; Kurashige, Y.; Neuscamman, E.; Chan, G. K.-L. Multireference quantum chemistry through a joint density matrix renormalization group and canonical transformation theory. *J. Chem. Phys.* **2010**, *132*, No. 024105.
(72) Liakos, D. G.; Neese, F. Interplay of correlation and relativistic effects in correlated calculations on transition-metal complexes: The $(Cu_2O_2)^{2+}$ core revisited. *J. Chem. Theory Comput.* **2011**, *7*, 1511−1523.
(73) Jiménez-Hoyos, C. A.; Rodríguez-Guzmán, R.; Scuseria, G. E. Multi-component symmetry-projected approach for molecular ground state correlations. *J. Chem. Phys.* **2013**, *139*, No. 204102.
(74) Landinez Borda, E. J.; Gomez, J.; Morales, M. A. Non-orthogonal multi-Slater determinant expansions in auxiliary field quantum Monte Carlo. *J. Chem. Phys.* **2019**, *150*, No. 074105.
(75) Blunt, N. S.; Mahajan, A.; Sharma, S. Efficient multireference perturbation theory without high-order reduced density matrices. *J. Chem. Phys.* **2020**, *153*, No. 164120.
(76) Motta, M.; Zhang, S.; Chan, G. K.-L. Hamiltonian symmetries in auxiliary-field quantum Monte Carlo calculations for electronic structure. *Phys. Rev. B* **2019**, *100*, No. 045127.
(77) Malone, F. D.; Zhang, S.; Morales, M. A. Accelerating Auxiliary-Field Quantum Monte Carlo Simulations of Solids with Graphical Processing Units. *J. Chem. Theory Comput.* **2020**, *16*, 4286−4297.

https://doi.org/10.1021/acs.jctc.1c00371
J. Chem. Theory Comput. 2021, 17, 4786−4798