A scalable route to first-order response properties with correlated sampling phaseless auxiliary-field quantum Monte Carlo

Leon Otis,[†] Sri Gudivada,[†] Marvin Friede,[‡] and James Shee*,[†],¶

†Department of Chemistry, Rice University, Houston, TX 77005-1892, USA

‡Mulliken Center for Theoretical Chemistry, University of Bonn, 53115 Bonn, Germany

¶Department of Physics and Astronomy, Rice University, Houston, TX 77005-1892, USA

E-mail: james.shee@rice.edu

Abstract

To make useful connections with experimental measurements, correlated electronic structure theories must accurately predict chemical properties in addition to energies. We present a finite-difference based algorithm to compute first-order response properties with phaseless auxiliary-field quantum Monte Carlo (ph-AFQMC) that relies on a branching correlated sampling approach. Focusing on electric dipole moments, we show that mean-field trial wave functions are sufficient to obtain high accuracy relative to CCSD(T) and experimental measurements for a set of 21 molecules, ranging in size from 2 to 18 atoms in their equilibrium geometries. As with energies, the quality of predicted dipole moments can be systematically improved with the use of correlated trial wave functions, even (or especially) in strongly correlated regimes. We show that the challenges faced by low-order perturbation theories in predicting the dipole moment of hydrogen fluoride across its dissociation coordinate are overcome with ph-AFQMC when using relatively simple trials. The key advantage of our approach over those

previously reported for ph-AFQMC is its scalability to large system sizes with a phaseless bias no worse than that of a typical ground-state energy calculation; we routinely converge dipole moments for systems with more than one thousand basis functions.

The accurate calculation of physical properties, typically related to the response of a molecule or material to weak electromagnetic fields, remains a significant challenge for approximate electronic structure methods. In the context of semi-empirical density functional theory development, substantial effort has been devoted to obtaining accurate energies, at times at the expense of accuracy for other observables, e.g., based on the electron density. With regard to *ab initio* wavefunction methods, many groups focusing on computing properties other than the energy have developed formalisms such as coupled perturbed Hartree-Fock ^{2–5} and the lambda equations for coupled cluster gradients and density matrices. ^{6–10}

Quantum Monte Carlo (QMC) methods, which have earned a reputation for highly accurate energetics, ^{11–14} face a number of challenges in pursuit of other properties. ^{15–20} One type of QMC method, called auxiliary-field QMC (AFQMC) is a promising approach for a variety of correlated many-particle systems. ^{14,21–23} AFQMC calculations of properties, to date, include studies of forces, ^{24,25} dipole moments, ^{26–29} densities ³⁰ and correlation functions. ^{31–36} One approach to obtaining properties in AFQMC (including observables corresponding to operators that do not commute with the Hamiltonian) is the back-propagation technique, ³⁷ which introduces a second imaginary-time propagation of the trial state when computing expectation values. A potential limitation on the accuracy of this approach comes from the increased bias caused by the phaseless approximation when a second constrained projection is performed. ^{26,38} More recently, another route to AFQMC properties has been developed based on automatic differentiation (AD-AFQMC). ²⁸ Initial applications ^{28,29} of AD-AFQMC to dipole moments have obtained improved agreement with experiment compared to back-propagation, but AD-AFQMC is currently limited to systems of a few hundred basis functions due to memory constraints.

To overcome these challenges, we introduce an accurate and scalable approach to first-

order properties in AFQMC based on finite differences (FD-AFQMC). This strategy rests on the use of correlated sampling to resolve small energy differences with a tractable number of Monte Carlo samples. In this study we focus on the computation of electric dipole moments, and demonstrate the accuracy of FD-AFQMC across a range of molecular sizes, including systems currently inaccessible to AD-AFQMC and canonical CCSD(T).

We first briefly review the main aspects of AFQMC before discussing the use of correlated sampling and finite differences. AFQMC is a projective method based on the application of the imaginary time propagator to some initial state $|\psi_I\rangle$:

$$|\psi_0\rangle = \lim_{\tau \to \infty} e^{-\tau \hat{H}} |\psi_I\rangle.$$
 (1)

The imaginary time propagation is discretized into small steps, $\Delta \tau$, and the Trotter-Suzuki decomposition³⁹ is applied to the exponential of the Hamiltonian to separate the one- and two-body parts. The electron-electron interaction \hat{V} can be written as a sum of squares of one-body operators:

$$\hat{V} = \frac{1}{2} \sum_{\gamma} \hat{L}_{\gamma}^2,\tag{2}$$

where

$$\hat{L}_{\gamma} = \sum_{pq} L_{pq}^{\gamma} \hat{a}_{p}^{\dagger} \hat{a}_{q}, \tag{3}$$

 \hat{a}_p^{\dagger} and \hat{a}_q are the usual one-electron creation and annihilation operators with p and q as orbital indices. The \hat{L}_{γ} can usually be obtained with a modified Cholesky decomposition ⁴⁰ and introducing them enables the exponential of the interaction term to be rewritten using the Hubbard-Stratonovich transformation: ^{41,42}

$$e^{-\Delta\tau\hat{V}} = \int \exp[\sqrt{-\Delta\tau} \sum_{\gamma} x_{\gamma} \hat{L}_{\gamma}] \prod_{\gamma} \frac{\exp(-\frac{1}{2}x_{\gamma}^{2})}{\sqrt{2\pi}} dx_{\gamma}$$
 (4)

The auxiliary fields x_{γ} introduced by the transformation can be sampled from a normal distribution, allowing a stochastic realization of the projection with random walkers. In practice,

the AFQMC projection is often constrained by the phaseless approximation imposed by a precomputed trial wave function.²¹ The phaseless approximation avoids numerical instabilities from the phase problem by enforcing that the walker weights remain real-valued and positive. Any potential bias from the constrained projection can be mitigated in principle by systematically improving the quality of the trial.

Correlated sampling procedures seek to reduce statistical noise in the differences of stochastically estimated quantities.⁴³ At a very general level, when computing a Monte Carlo estimate of the difference of two quantities A - B, the variance of the difference is:

$$var(A - B) = var(A) + var(B) - 2Cov(A, B) = var(A) + var(B) - 2(\langle AB \rangle - \langle A \rangle \langle B \rangle).$$
 (5)

The covariance Cov(A, B) is zero if A and B are independently sampled, which maximizes var(A - B) while in the opposite limit of perfect correlation in the sampling, 2Cov(A, B) = var(A) + var(B) and var(A - B) is zero with no noise in the estimate of A - B.

In the context of AFQMC, correlated sampling is implemented by using the same set of auxiliary fields and pairing the propagation of walkers for two (or more) similar systems. ⁴⁴ Recent work ⁴⁵ has shown that population control of the walkers can be approximately incorporated in a correlated fashion through multiple schemes. In this work, we use the "static" approach where the branching and annihilation importance sampling decisions based on the walker weights are based on a primary AFQMC calculation and used for the other correlated secondary calculations. The use of correlated sampling in AFQMC remains at a relatively early stage with some initial applications to ionization potentials, ⁴⁴ electron binding energies, ⁴⁶ diatomic dissociation energies, ⁴⁷ nuclear forces, ^{45,48} and noncovalent interactions. ⁴⁹

The dipole moment can be expressed as a derivative of the energy U with respect to electric field E:

$$\mu = -\left(\frac{\partial U}{\partial E}\right)_{E=0} \approx -\frac{U(E = \delta E) - U(E = 0)}{\delta E}.$$
 (6)

The derivative can be approximated with a difference of energies between a zero-field calculation and one with small field δE , and this approach can be repeated for each component of the dipole moment in general. In principle, the finite-difference approximation of the energy derivative improves as the magnitude of the perturbing field δE goes to zero; however, in this limit the energy difference in the numerator will be small and the suppression of statistical noise becomes critical. This motivates our present use of correlated sampling; indeed, we find that finite-difference dipole moments cannot be converged without it.

To summarize methodological details of the calculations to follow, we employ an electric field strength of 10^{-5} a.u. unless stated otherwise. With a few exceptions, we use restricted Hartree-Fock (RHF) trial wave functions obtained from PySCF, 50,51 which we also use for HF, MP2, and coupled cluster dipole moment calculations. The frozen-core approximation is used throughout our calculations. We also compute a set of dipole moments from DFT using ORCA 52,53 to compare against our results from wavefunction methods. Our geometries are such that the molecular center of charge is at the origin. In cases where the orientation of the dipole moment is clear from symmetry, we only need to use an electric field perturbation in that direction. When we consider larger and less symmetric molecules, we have perturbations along the x, y, and z directions along with the unperturbed case for a total of four AFQMC calculations. In our results, we show the magnitude of the dipole moment vector, computed from the set of Cartesian components, and it is this quantity that is compared with different levels of theory and experimental measurements. To characterize the overall deviation of AFQMC values from reference values for a set of molecules we compute the root mean squared regularized error (RMSRE) 54 defined as:

RMSRE =
$$\sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(\frac{\mu_i - \mu_{i,ref}}{\max(\mu_{i,ref}, 1 D)}\right)^2}.$$
 (7)

As a first check of the accuracy of our methodology, we consider a set of small and highly symmetric molecules that have previously been studied with AD-AFQMC.²⁹ As shown in Table 1, we see that our FD-AFQMC technique is quite accurate with dipole moments often within a few hundredths of a Debye of the values from automatic differentiation as well as experiment. These results were obtained with modest computational effort, using only 512 walkers propagating for 40 Ha⁻¹. We also note that we have found it completely infeasible to resolve the required energy differences with uncorrelated sampling even in these small systems, highlighting the importance of correlated sampling to our finite-difference strategy.

Table 1: Dipole moments in Debye for a set of small molecules. Stochastic uncertainties on the last digits are given in parentheses. All calculations use the aug-cc-pVQZ basis set. RHF, AD-AFQMC, and experimental values are taken from Ref. 29.

Method	H_2O	NH_3	HCl	HBr	CO	$\mathrm{CH_{2}O}$
RHF	1.985	1.617	1.235	0.947	-0.292	2.867
AD-AFQMC-RHF	1.815(5)	1.479(15)	1.085(8)	0.826(8)	0.025(8)	2.506(10)
FD-AFQMC-RHF	1.879(5)	1.491(8)	1.120(6)	0.716(2)	0.033(10)	2.448(20)
Experiment	$1.855^{\frac{5}{5}}$	$1.477^{\frac{56}{6}}$	$1.093^{\frac{57}{57}}$	$0.826^{\frac{57}{57}}$	$0.122^{\frac{58}{58}}$	$2.333^{\frac{59}{59}}$

Now we turn our attention to computing dipole moments of larger molecules, including those beyond the current reach of AD-AFQMC. To devise a manageable set of test molecules, we have considered experimental values listed in the Computational Chemistry Comparison and Benchmark Database (CCCBDB). ⁶⁰ We selected molecules with more than 12 atoms with listed experimental uncertainties of under 0.05 Debye. These experimental values can also be found in various primary literature sources cited on the CCCBDB page. ^{61–72} We performed DFT geometry optimizations of our chosen molecules using the ω B97M-V functional ⁷³ and the def2-QZVPP basis set with ORCA. ^{52,53} For FD-AFQMC calculations on this set, 2048 walkers were propagated for 40 Ha⁻¹ in all cases.

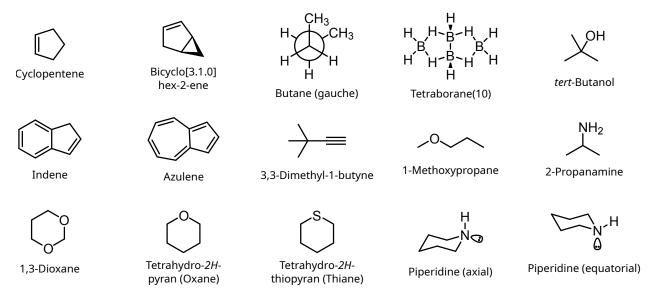


Figure 1: Test set of larger molecules from CCCBDB 60 with experimentally measured dipole moments.

For this set of larger molecules (Figure 1), we first compare dipole moments computed from different theoretical methods. Table 2 shows good agreement between FD-AFQMC and CCSD(T) with differences of only a few hundredths of a Debye except in the case of azulene. Taking CCSD(T) as the reference, Table 3 shows that the accuracy of FD-AFQMC across the whole set is very similar to that of CCSD with a RMSRE of about 4 percent, and can be achieved with $\mathcal{O}(N^4)$ (per sample) scaling with system size vs the $\mathcal{O}(N^6)$ scaling of CCSD. MP2 performs very well for this test set with an error of only 3 percent in its worst case, but has been found to yield significantly larger errors in Ref. 54, especially in the case of spin-polarized systems. In terms of absolute errors, FD-AFQMC, CCSD and MP2 are all highly accurate with average deviations of only about 0.01 or 0.02 Debye.

Table 2: Dipole moments in Debye for a collection of larger molecules at different levels of theory. All calculations used a cc-pVTZ basis set. Stochastic uncertainties of FD-AFQMC are given in parentheses.

Molecule	FD-AFQMC-RHF	HF	MP2	CCSD	CCSD(T)
Cyclopentene	0.165(4)	0.185	0.178	0.169	0.169
2-Propanamine	1.237(3)	1.284	1.229	1.232	1.219
Tetraborane	0.481(4)	0.529	0.473	0.490	0.478
Gauche Butane	0.086(3)	0.088	0.088	0.088	0.088
Bicyclo[3.1.0]hex-2-ene	0.259(6)	0.249	0.253	0.237	0.242
1,3-Dioxane	1.998(4)	2.229	1.986	2.032	1.971
1-Methoxypropane	1.086(4)	1.227	1.073	1.106	1.066
tert-Butanol	1.477(4)	1.628	1.473	1.501	1.465
3,3-Dimethyl-1-butyne	0.647(6)	0.691	0.568	0.588	0.585
Tetrahydro- $2H$ -pyran	1.411(5)	1.568	1.402	1.433	1.392
Tetrahydro- $2H$ -thiopyran	1.762(7)	1.933	1.758	1.742	1.722
Equatorial Piperidine	0.822(6)	0.891	0.826	0.843	0.823
Axial Piperidine	1.110(3)	1.193	1.115	1.129	1.105
Indene	0.671(10)	0.685	0.640	0.624	0.620
Azulene	1.131(30)	1.452	0.950	1.138	0.981

Table 3: Statistics over the set of larger molecules comparing to CCSD(T), including root mean square regularized error, maximum percentage error, mean absolute deviation, and maximum absolute deviation. Absolute deviations are in Debye. All methods use the cc-pVTZ basis set and FD-AFQMC uses a RHF trial.

Method	RMSRE (%)	MaxE (%)	MeanAD (D)	MaxAD (D)
FD-AFQMC-RHF	4.14	13.22	0.028	0.132
RHF	14.99	47.12	0.127	0.471
MP2	1.33	3.08	0.013	0.036
CCSD	4.451	15.68	0.029	0.157

We also compare FD-AFQMC against experiment and consider the effects of basis set size. As seen in Figure 2, while the use of larger basis sets can increase the dipole moment predicted by FD-AFQMC, there is not a consistent trend across all the molecules in our test set. Some cases with larger errors at the cc-pVTZ level, such as tert-butanol and tetrahydro-2H-pyran, are noticeably improved by increasing the basis set to aug-cc-pVQZ. However, use of aug-cc-pVQZ can lead to overestimation of the dipole moment and worsen the level of agreement with experiment, most prominently in the cases of cyclopentene and equatorial piperidine. Besides accuracy considerations, our results are also a demonstration of our

method's ability to scale to large numbers of basis functions. Our largest molecule, azulene, has 412 basis functions with cc-pVTZ and 1168 with aug-cc-pVQZ, well beyond the current capabilities of the automatic differentiation approach.

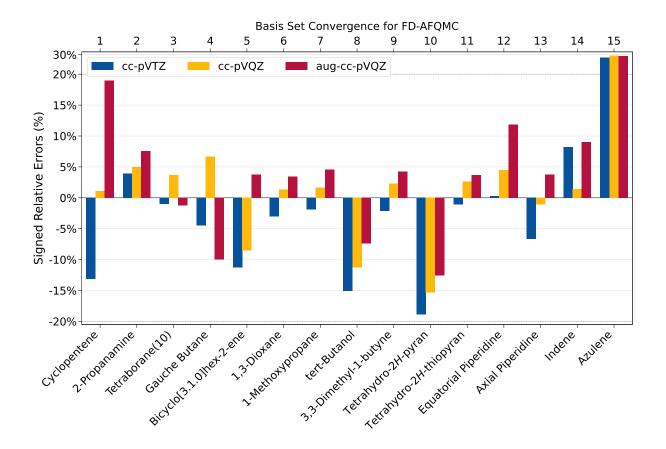


Figure 2: Signed relative error of FD-AFQMC-RHF to experiment in different basis sets.

In both absolute and percentage terms, the errors of FD-AFQMC relative to experiment for our large molecule set are encouragingly small. In Table 4, aug-cc-pVQZ values typically differ from experiment by less than 0.1 D and the largest error is 0.256 D in the case of azulene. The percentage error across the entire set in Table 5 for FD-AFQMC is comparable to those of the most successful DFT functionals we have tested (note the high accuracy of the hybrid functionals B3LYP and PBE0), and somewhat better than those of ω B97X-V and ω B97M-V. For azulene, where the choice of basis set has little effect, we investigate the role of the trial function in AFQMC and apply our finite-difference approach with a CASSCF

wave function from a (10e,10o) active space of π and π^* orbitals. We see in Figure 3 that a better trial wave function leads to a significant improvement in the dipole moment to a value of 0.969(5) D, within 0.1 D of the experimental value. The azulene results with RHF and CASSCF trials provide a demonstration of the systematic improvability of our approach to dipole moments.

Table 4: Dipole moments in Debye for the collection of larger molecules. FD-AFQMC values use the aug-cc-pVQZ basis set and RHF trial. Stochastic and experimental uncertainties are in parentheses.

Molecule	FD-AFQMC	Experiment
Cyclopentene	0.226(4)	$0.190(6)^{61}$
2-Propanamine	1.280(6)	$1.190(30)^{62}$
Tetraborane	0.480(7)	$0.486(2)^{63}$
Gauche Butane	0.081(4)	$0.090(2)^{64}$
Bicyclo[3.1.0]hex-2-ene	0.303(6)	$0.292(15)^{61}$
1,3-Dioxane	2.131(8)	$2.060(20)^{65}$
1-Methoxypropane	1.158(6)	$1.107(13)^{66}$
tert-Butanol	1.611(8)	$1.740(40)^{67}$
3,3-Dimethyl-1-butyne	0.689(5)	$0.661(4)^{68}$
Tetrahydro- $2H$ -pyran	1.521(4)	$1.740(20)^{65}$
Tetrahydro- $2H$ -thiopyran	1.846(4)	$1.781(10)^{69}$
Equatorial Piperidine	0.917(6)	$0.820(20)^{70}$
Axial Piperidine	1.234(8)	$1.189(15)^{70}$
Indene	0.676(14)	$0.620(20)^{71}$
Azulene	1.138(33)	$0.882(2)^{72}$

Table 5: Statistics over the set of larger molecules comparing to experiment, including root mean square regularized error and maximum percentage error. FD-AFQMC values use the aug-cc-pVQZ basis and RHF trial, while DFT calculations use the def2-QZVPP basis.

Method	RMSRE (%)	MaxE (%)
FD-AFQMC-RHF	8.70	25.60
PBE	8.86	20.17
B3LYP	7.02	14.58
PBE0	8.24	19.08
$\omega \mathrm{B}97\mathrm{X}\text{-}\mathrm{V}$	13.36	46.07
$\omega \mathrm{B}97\mathrm{M}\text{-}\mathrm{V}$	12.67	43.82

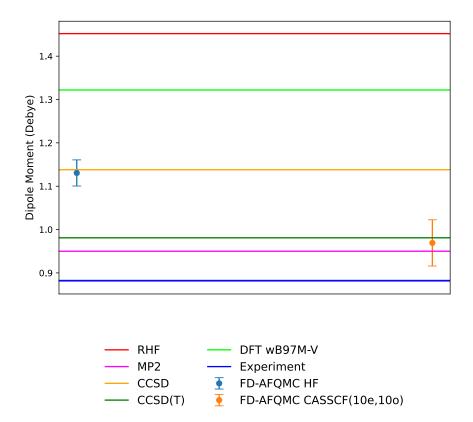


Figure 3: Dipole moment of azulene. The DFT calculation used a def2-QZVPP basis set while all others used cc-pVTZ. 99% weight of the CAS wave function corresponding to 726 determinants was used in the AFQMC trial.

Given the high accuracy of many methods on our test set, we finally consider FD-AFQMC's performance at non-equilibrium geometries with the notoriously challenging dissociation of hydrogen fluoride (FH). Previous benchmarking ⁵⁴ of DFT functionals on this test system has demonstrated that many functionals struggle to accurately predict the dipole moment throughout the dissociation coordinate. Hybrid functionals such as B3LYP may overestimate the maximum dipole moment and predict an overly slow decay at longer separations. We see in Figure 4 that this molecule also challenges MP2 with a significant spike in the dipole moment around 1.35 Å followed by an overly rapid decay. FD-AFQMC with a UHF trial agrees with the FCI reference more closely throughout the dissociation than UHF or MP2, but still overestimates the dipole moment at intermediate separation lengths.

The limitations of FD-AFQMC-UHF can be alleviated with improved trial wave functions, as shown in the inset of Figure 4. Trial wave functions of only 25 determinants from

(8e,5o) CASSCF calculations are sufficient to greatly improve the level of agreement between FD-AFQMC and FCI, and any remaining error can be eliminated with trials from a larger (6e,9o) active space. Our ability to converge to the FCI reference with improved trial wave functions provides some assurance that FD-AFQMC can obtain accurate properties even in cases with more challenging electronic correlation – in this case, a subtle competition between ionic and covalent electronic configurations as the single bond is broken.⁷⁴

However, we did encounter some unexpected challenges in converging the FD-AFQMC dipole moments along the FH dissociation coordinate. For example, we observed greater sensitivity of FD-AFQMC to noise and potential loss of correlation at longer stretches, relative to our studies of equilibrium geometries. With UHF trials, these challenges were overcome by employing smaller propagation times of only up to 20 Ha⁻¹, which was still sufficient to converge the dipole moments before any significant growth of stochastic noise due to the gradual loss of correlation between the primary and secondary systems. In addition, we removed any outlier spikes in the energy difference trajectories of more than 10⁻⁴ Ha. Some FD-AFQMC calculations with CASSCF trial wavefunctions were found to be more sensitive to noise than those with UHF trials; we addressed this by using a slightly larger perturbing field strength of 10⁻³ a.u. to successfully resolve energy differences and obtain accurate dipole moments. At very long bond lengths beyond 3.5 Å, we encounter severe noise with CASSCF trials and incorrectly obtain non-zero dipole moments. Challenges with noise at long bond lengths have also been observed in the computation of forces ⁴⁸ and future work will seek to improve correlated sampling's performance in this regime.

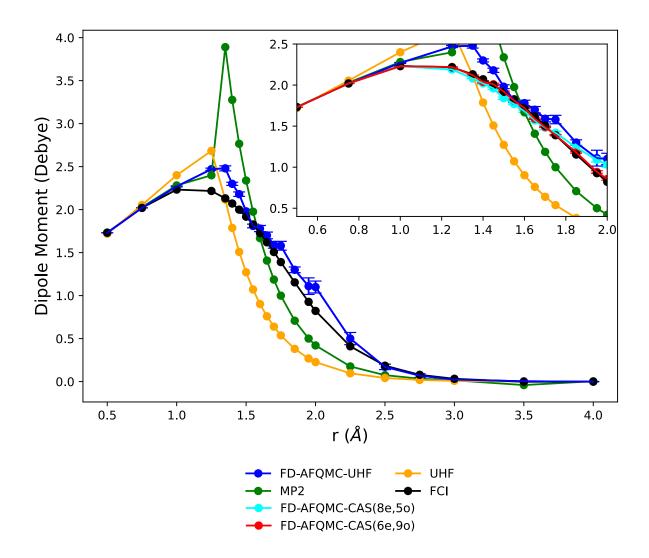


Figure 4: Dipole moments of the hydrogen fluoride diatomic along the dissociation coordinate. All methods use the 6-31G basis set, in order to compare with FCI values. FD-AFQMC-CAS results are shown in the inset. 25 determinants were used for FD-AFQMC-CAS(8e,5o) and 3546 were used for FD-AFQMC-CAS(6e,9o). All FD-AFQMC calculations use 2048 walkers.

In conclusion, we have introduced a finite-difference based approach within AFQMC to response properties and demonstrated that FD-AFQMC, with the aid of correlated sampling, can obtain highly accurate electric dipole moments even for relatively large molecular sizes and in regimes of strong electron correlation. While simple Hartree-Fock trial wave functions are already sufficient in most situations, we have also shown that straightforward improvements in the trial quality can ensure accurate dipole moments in challenging cases such as the azulene molecule and the dissociation of hydrogen fluoride. FD-AFQMC is scalable

to system sizes beyond the reach of CCSD, CCSD(T), and AFQMC approaches to properties based on automatic differentiation, and we can routinely compute dipole moments for asymmetric molecules and large basis sets. In future studies, our correlated sampling methodology could potentially be improved through a dynamic reweighting scheme ⁴⁵ or the averaging of energy differences from multiple short trajectories, ⁴⁴ in combination with other finite difference formulations. FD-AFQMC could also be extended to provide high quality reference values for dipole moments of electronically excited states. ⁷⁵ We also plan to apply the correlated sampling FD-AFQMC protocol to other first-order response properties such as magnetic dipole moments, using gauge invariant atomic orbitals, ⁷⁶ and to other properties such as quadrupole moments and polarizabilities.

1 Acknowledgments

We thank Brad Ganoe for helpful discussions. This work was supported in part by the Big-Data Private-Cloud Research Cyberinfrastructure MRI-award funded by NSF under grant CNS-1338099 and by Rice University's Center for Research Computing (CRC). This research also used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725. J. Shee acknowledges support from the Robert A. Welch Foundation, Award Number C-2212. This work used Expanse at the San Diego Supercomputing Center through allocation CHE240177 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services and Support (ACCESS) program, which is supported by U.S. National Science Foundation grants 2138259, 2138286, 2138307, 2137603, and 2138296.

2 Supplementary information

Cartesian coordinates of optimized structures, table of DFT dipole moment values

References

- (1) Medvedev, M. G.; Bushmarinov, I. S.; Sun, J.; Perdew, J. P.; Lyssenko, K. A. Density functional theory is straying from the path toward the exact functional. *Science* **2017**, 355, 49–52.
- (2) McWeeny, R. Some Recent Advances in Density Matrix Theory. Rev. Mod. Phys. 1960, 32, 335–369.
- (3) McWeeny, R. Perturbation Theory for the Fock-Dirac Density Matrix. *Phys. Rev.* **1962**, 126, 1028–1034.
- (4) Diercksen, G.; McWeeny, R. Self-Consistent Perturbation Theory. I. General Formulation and Some Applications. *The Journal of Chemical Physics* **1966**, *44*, 3554–3560.
- (5) Wolinski, K.; Hinton, J. F.; Pulay, P. Efficient implementation of the gauge-independent atomic orbital method for NMR chemical shift calculations. *Journal of the American* Chemical Society 1990, 112, 8251–8260.
- (6) Bartlett, R. J.; Musiał, M. Coupled-cluster theory in quantum chemistry. Rev. Mod. Phys. 2007, 79, 291–352.
- (7) Scheiner, A. C.; Scuseria, G. E.; Rice, J. E.; Lee, T. J.; Schaefer, I., Henry F. Analytic evaluation of energy gradients for the single and double excitation coupled cluster (CCSD) wave function: Theory and application. The Journal of Chemical Physics 1987, 87, 5361–5373.
- (8) Salter, E. A.; Trucks, G. W.; Bartlett, R. J. Analytic energy derivatives in many-body methods. I. First derivatives. *The Journal of Chemical Physics* **1989**, *90*, 1752–1766.
- (9) Gauss, J.; Stanton, J. F.; Bartlett, R. J. Coupled-cluster open-shell analytic gradients: Implementation of the direct product decomposition approach in energy gradient calculations. The Journal of Chemical Physics 1991, 95, 2623–2638.

- (10) Gauss, J.; Stanton, J. F. Coupled-cluster calculations of nuclear magnetic resonance chemical shifts. *The Journal of Chemical Physics* **1995**, *103*, 3561–3577.
- (11) Foulkes, W. M. C.; Mitas, L.; Needs, R. J.; Rajagopal, G. Quantum Monte Carlo simulations of solids. *Rev. Mod. Phys.* **2001**, *73*, 33–83.
- (12) Austin, B. M.; Zubarev, D. Y.; Lester, W. A. J. Quantum Monte Carlo and Related Approaches. *Chemical Reviews* **2012**, *112*, 263–288.
- (13) Dubecký, M.; Mitas, L.; Jurečka, P. Noncovalent Interactions by Quantum Monte Carlo. Chemical Reviews 2016, 116, 5188–5215.
- (14) Motta, M.; Zhang, S. Ab initio computations of molecular systems by the auxiliary-field quantum Monte Carlo method. Wiley Interdisciplinary Reviews: Computational Molecular Science 2018, 8, e1364.
- (15) Assaraf, R.; Caffarel, M. Computing forces with quantum Monte Carlo. *The Journal of Chemical Physics* **2000**, *113*, 4028–4034.
- (16) Assaraf, R.; Caffarel, M. Zero-variance zero-bias principle for observables in quantum Monte Carlo: Application to forces. The Journal of Chemical Physics 2003, 119, 10536–10552.
- (17) Ríos, P. L.; Conduit, G. J. Tail-regression estimator for heavy-tailed distributions of known tail indices and its application to continuum quantum Monte Carlo data. *Phys. Rev. E* 2019, 99, 063312.
- (18) Tiihonen, J.; Clay, I., Raymond C.; Krogel, J. T. Toward quantum Monte Carlo forces on heavier ions: Scaling properties. The Journal of Chemical Physics 2021, 154, 204111.
- (19) Hood, R. Q.; Chou, M. Y.; Williamson, A. J.; Rajagopal, G.; Needs, R. J.; Foulkes, W.

- M. C. Quantum Monte Carlo Investigation of Exchange and Correlation in Silicon. *Phys. Rev. Lett.* **1997**, *78*, 3350–3353.
- (20) Kaiser, A.; Kümmel, S. Accurate electron densities from quantum Monte Carlo calculations using real-space grids. *The Journal of Chemical Physics* **2025**, *162*, 134108.
- (21) Zhang, S.; Krakauer, H. Quantum Monte Carlo Method using Phase-Free Random Walks with Slater Determinants. *Phys. Rev. Lett.* **2003**, *90*, 136401.
- (22) Lee, J.; Pham, H. Q.; Reichman, D. R. 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* **2022**, *18*, 7024–7042.
- (23) Shee, J.; Weber, J. L.; Reichman, D. R.; Friesner, R. A.; Zhang, S. On the potentially transformative role of auxiliary-field quantum Monte Carlo in quantum chemistry: A highly accurate method for transition metals and beyond. *The Journal of Chemical Physics* **2023**, *158*, 140901.
- (24) Motta, M.; Zhang, S. Communication: Calculation of interatomic forces and optimization of molecular geometry with auxiliary-field quantum Monte Carlo. *The Journal of Chemical Physics* **2018**, *148*, 181101.
- (25) Chen, S.; Zhang, S. Computation of forces and stresses in solids: Towards accurate structural optimization with auxiliary-field quantum Monte Carlo. *Phys. Rev. B* **2023**, 107, 195150.
- (26) Motta, M.; Zhang, S. Computation of Ground-State Properties in Molecular Systems: Back-Propagation with Auxiliary-Field Quantum Monte Carlo. *Journal of Chemical Theory and Computation* 2017, 13, 5367–5378.
- (27) Mahajan, A.; Lee, J.; Sharma, S. Selected configuration interaction wave functions in

- phaseless auxiliary field quantum Monte Carlo. *The Journal of Chemical Physics* **2022**, 156, 174111.
- (28) Mahajan, A.; Kurian, J. S.; Lee, J.; Reichman, D. R.; Sharma, S. Response properties in phaseless auxiliary field quantum Monte Carlo. *The Journal of Chemical Physics* **2023**, *159*, 184101.
- (29) Mahajan, A.; Thorpe, J. H.; Kurian, J. S.; Reichman, D. R.; Matthews, D. A.; Sharma, S. Beyond CCSD(T) Accuracy at Lower Scaling with Auxiliary Field Quantum Monte Carlo. *Journal of Chemical Theory and Computation* **2025**, *21*, 1626–1642.
- (30) Chen, S.; Motta, M.; Ma, F.; Zhang, S. Ab initio electronic density in solids by many-body plane-wave auxiliary-field quantum Monte Carlo calculations. *Phys. Rev. B* **2021**, 103, 075138.
- (31) Zhang, S.; Carlson, J.; Gubernatis, J. E. Constrained path Monte Carlo method for fermion ground states. *Phys. Rev. B* **1997**, *55*, 7464–7477.
- (32) Motta, M.; Galli, D. E.; Moroni, S.; Vitali, E. Imaginary time correlations and the phaseless auxiliary field quantum Monte Carlo. The Journal of Chemical Physics 2014, 140, 024107.
- (33) Vitali, E.; Shi, H.; Qin, M.; Zhang, S. Computation of dynamical correlation functions for many-fermion systems with auxiliary-field quantum Monte Carlo. *Phys. Rev. B* **2016**, *94*, 085140.
- (34) Qin, M.; Shi, H.; Zhang, S. Numerical results on the short-range spin correlation functions in the ground state of the two-dimensional Hubbard model. *Phys. Rev. B* **2017**, 96, 075156.
- (35) Lee, J.; Zhang, S.; Reichman, D. R. Constrained-path auxiliary-field quantum Monte Carlo for coupled electrons and phonons. *Phys. Rev. B* **2021**, *103*, 115123.

- (36) Church, M. S.; Rubenstein, B. M. Real-time dynamics of strongly correlated fermions using auxiliary field quantum Monte Carlo. The Journal of Chemical Physics 2021, 154, 184103.
- (37) Zhang, S.; Carlson, J.; Gubernatis, J. E. Constrained Path Quantum Monte Carlo Method for Fermion Ground States. *Phys. Rev. Lett.* **1995**, *74*, 3652–3655.
- (38) Lee, J.; Malone, F. D.; Morales, M. A.; Reichman, D. R. Spectral Functions from Auxiliary-Field Quantum Monte Carlo without Analytic Continuation: The Extended Koopmans' Theorem Approach. *Journal of Chemical Theory and Computation* 2021, 17, 3372–3387.
- (39) Suzuki, M. Generalized Trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems. *Communications in Mathematical Physics* **1976**, *51*, 183–190.
- (40) Pedersen, T. B.; Lehtola, S.; Fdez. Galván, I.; Lindh, R. The versatility of the Cholesky decomposition in electronic structure theory. WIREs Computational Molecular Science **2024**, 14, e1692.
- (41) Hubbard, J. Calculation of Partition Functions. Phys. Rev. Lett. 1959, 3, 77–78.
- (42) Stratonovich, R. L. A method for the computation of quantum distribution functions.

 Dokl. Akad. Nauk SSSR (N.S.) 1957, 115, 1097–1100.
- (43) Monte Carlo Methods; John Wiley & Sons, Ltd, 2008; Chapter 2, pp 7–34.
- (44) Shee, J.; Zhang, S.; Reichman, D. R.; Friesner, R. A. Chemical Transformations Approaching Chemical Accuracy via Correlated Sampling in Auxiliary-Field Quantum Monte Carlo. *Journal of Chemical Theory and Computation* **2017**, *13*, 2667–2680.
- (45) Chen, S.; Yang, Y.; Morales, M.; Zhang, S. Algorithm for branching and population control in correlated sampling. *Phys. Rev. Res.* **2023**, *5*, 043169.

- (46) Hao, H.; Shee, J.; Upadhyay, S.; Ataca, C.; Jordan, K. D.; Rubenstein, B. M. Accurate Predictions of Electron Binding Energies of Dipole-Bound Anions via Quantum Monte Carlo Methods. *The Journal of Physical Chemistry Letters* **2018**, *9*, 6185–6190.
- (47) Shee, J.; Rudshteyn, B.; Arthur, E. J.; Zhang, S.; Reichman, D. R.; Friesner, R. A. 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. *Journal of chemical theory and computation* **2019**, *15*, 2346–2358.
- (48) Goings, J. J.; Shin, K.; Noh, S.; Kyoung, W.; Kim, D.; Baek, J.; Roetteler, M.; Epifanovsky, E.; Zhao, L. Molecular Properties in Quantum-Classical Auxiliary-Field Quantum Monte Carlo: Correlated Sampling with Application to Accurate Nuclear Forces. 2025; https://arxiv.org/abs/2507.17992.
- (49) Awasthi, D.; Otis, L.; Huang, E.; Shee, J. Noncovalent interaction energies with phase-less auxiliary-field quantum Monte Carlo. 2025; ChemRxiv 10.26434/chemrxiv-2025-hvnfs.
- (50) 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.; others PySCF: the Python-based simulations of chemistry framework. Wiley Interdisciplinary Reviews: Computational Molecular Science 2018, 8, e1340.
- (51) Sun, Q.; Zhang, X.; Banerjee, S.; Bao, P.; Barbry, M.; Blunt, N. S.; Bogdanov, N. A.; Booth, G. H.; Chen, J.; Cui, Z.-H.; others Recent developments in the PySCF program package. *The Journal of chemical physics* **2020**, *153*.
- (52) Neese, F. The ORCA program system. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2012, 2, 73–78.

- (53) Neese, F. Software Update: The ORCA Program System—Version 6.0. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2025, 15, e70019.
- (54) Hait, D.; Head-Gordon, M. How Accurate Is Density Functional Theory at Predicting Dipole Moments? An Assessment Using a New Database of 200 Benchmark Values. Journal of Chemical Theory and Computation 2018, 14, 1969–1981.
- (55) Lide, D. R. CRC handbook of chemistry and physics: a ready-reference book of chemical and physical data; CRC press, 1995.
- (56) Shimizu, F. Stark Spectroscopy of NH3 ν 2 Band by 10-μ CO2 and N2O Lasers. The Journal of Chemical Physics 1970, 52, 3572–3576.
- (57) NIST Diatomic spectral database. 2005; https://www.nist.gov/pml/diatomic-spectral-database.
- (58) Meerts, W.; De Leeuw, F.; Dymanus, A. Electric and magnetic properties of carbon monoxide by molecular-beam electric-resonance spectroscopy. Chemical Physics 1977, 22, 319–324.
- (59) Theulé, P.; Callegari, A.; Rizzo, T. R.; Muenter, J. S. Fluorescence detected microwave Stark effect measurements in excited vibrational states of H2CO. *The Journal of Chemical Physics* **2003**, *119*, 8910–8915.
- (60) Computational Chemistry Comparison and Benchmark Database: Experimental Dipoles. 2022; https://cccbdb.nist.gov/diplistx.asp.
- (61) NIST Hydrocarbon spectral database. 2004; https://www.nist.gov/pml/hydrocarbon-spectral-database.
- (62) Mehrotra, S. C.; Griffin, L. L.; Britt, C. O.; Boggs, J. E. Microwave spectrum, structure, dipole moment, and quadrupole coupling constants of isopropylamine. *Journal of Molecular Spectroscopy* **1977**, *64*, 244–251.

- (63) Simmons, N. P. C.; Burg, A. B.; Beaudet, R. A. Microwave spectrum, structure, and dipole moment of tetraborane(10), B4H10. *Inorganic Chemistry* **1981**, *20*, 533–536.
- (64) Hüttner, W.; Majer, W.; Kästle, H. Ground-state rotational spectrum and spectroscopic parameters of the gauche butane conformer. *Molecular Physics* **1989**, *67*, 131–140.
- (65) Hellwege, K.-H., Hellwege, A. M., Eds. Molecular Constants from Microwave, Molecular Beam, and Electron Spin Resonance Spectroscopy; Springer-Verlag Berlin Heidelberg, Copyright 1974 Springer-Verlag Berlin Heidelberg.
- (66) Kato, H.; Nakagawa, J.; Hayashi, M. Microwave spectrum, structure, and dipole moment of the trans-trans isomer of methylpropylether. *Journal of Molecular Spectroscopy* 1980, 80, 272–278.
- (67) Digest of literature on dielectrics; The National Academies Press: Washington, DC, 1970.
- (68) Nugent, L. J.; Mann, D. E.; Lide, J., David R. Microwave Structure Determinations on Tertiary Butyl Acetylene and Tertiary Butyl Cyanide. The Journal of Chemical Physics 1962, 36, 965–971.
- (69) Kitchin, R.; Malloy, T. B.; Cook, R. L. The molecular conformation and dipole moment of thiane from the microwave spectrum. *Journal of Molecular Spectroscopy* **1975**, *57*, 179–188.
- (70) Parkin, J.; Buckley, P.; Costain, C. The microwave spectrum of piperidine: Equatorial and axial ground states. *Journal of Molecular Spectroscopy* **1981**, *89*, 465–483.
- (71) Caminati, W. Low-energy vibrations of indene. J. Chem. Soc., Faraday Trans. 1993, 89, 4153–4155.
- (72) Huber, S.; Grassi, G.; Bauder, A. Structure and symmetry of azulene as determined from microwave spectra of isotopomers. *Molecular Physics* **2005**, *103*, 1395–1409.

- (73) Mardirossian, N.; Head-Gordon, M. ωB97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. The Journal of Chemical Physics 2016, 144, 214110.
- (74) Mayhall, N. J.; Horn, P. R.; Sundstrom, E. J.; Head-Gordon, M. Spin-flip non-orthogonal configuration interaction: a variational and almost black-box method for describing strongly correlated molecules. *Phys. Chem. Chem. Phys.* 2014, 16, 22694–22705.
- (75) Damour, Y.; Quintero-Monsebaiz, R.; Caffarel, M.; Jacquemin, D.; Kossoski, F.; Scemama, A.; Loos, P.-F. Ground- and Excited-State Dipole Moments and Oscillator Strengths of Full Configuration Interaction Quality. *Journal of Chemical Theory and Computation* 2023, 19, 221–234.
- (76) Helgaker, T.; Jo/rgensen, P. An electronic Hamiltonian for origin independent calculations of magnetic properties. *The Journal of Chemical Physics* **1991**, *95*, 2595–2601.
- (77) Strande, S.; Cai, H.; Tatineni, M.; Pfeiffer, W.; Irving, C.; Majumdar, A.; Mishin, D.; Sinkovits, R.; Norman, M.; Wolter, N.; others *Practice and Experience in Advanced Research Computing 2021: Evolution Across All Dimensions*; 2021; pp 1–4.
- (78) Boerner, T. J.; Deems, S.; Furlani, T. R.; Knuth, S. L.; Towns, J. Practice and experience in advanced research computing 2023: Computing for the common good; 2023; pp 173–176.

TOC Graphic

