On the potentially transformative role of auxiliary-field quantum Monte Carlo in quantum chemistry: A highly accurate method for transition metals and beyond

James Shee,[1, a] John L. Weber,[2] David R. Reichman,[2] Richard A. Friesner,[2] and Shiwei Zhang[3, b]
[1] *Kenneth S. Pitzer Center for Theoretical Chemistry, Department of Chemistry, University of California, Berkeley, California 94720, USA*
[2] *Department of Chemistry, Columbia University, 3000 Broadway, New York, NY, 10027*
[3] *Center for Computational Quantum Physics, Flatiron Institute, 162 5th Avenue, New York, NY 10010*

Approximate solutions to the *ab initio* electronic structure problem have been a focus of theoretical and computational chemistry research for much of the past century, with the goal of predicting relevant energy differences to within "chemical accuracy" (1 kcal/mol). For small organic molecules, or in general for weakly correlated main group chemistry, a hierarchy of single-reference wavefunction methods have been rigorously established spanning perturbation theory and the coupled cluster (CC) formalism. For these systems, CC with singles, doubles, and perturbative triples (CCSD(T)) is known to achieve chemical accuracy, albeit at $\mathcal{O}(N^7)$ computational cost. In addition, a hierarchy of density functional approximations of increasing formal sophistication, known as Jacob's ladder, has been shown to systematically reduce average errors over large data sets representing weakly-correlated chemistry. However, the accuracy of such computational models is less clear in the increasingly important frontiers of chemical space including transition metals and f-block compounds, in which strong correlation can play an important role in reactivity. A stochastic method, phaseless auxiliary-field quantum Monte Carlo (ph-AFQMC), has been shown capable of producing chemically accurate predictions even for challenging molecular systems beyond the main-group, with relatively low $\mathcal{O}(N^3-N^4)$ cost and near-perfect parallel efficiency. Herein we present our perspectives on the past, present, and future of the ph-AFQMC method. We focus on its potential in transition metal quantum chemistry to be a highly accurate, systematically-improvable method which can reliably probe strongly correlated systems in biology and chemical catalysis, and provide reference thermochemical values (for future development of density functionals or interatomic potentials) when experiments are either noisy or absent. Finally, we discuss the present limitations of the method, and where we expect near term development to be most fruitful.

I. PRELIMINARIES

The scope of this perspective will be restricted to

1. solving what is commonly known as the electronic structure problem, i.e. finding the lowest eigenvalue of the Schrödinger equation (of a given symmetry) under the Born-Oppenheimer Hamiltonian.

2. quantum chemical calculations on classical computing devices. We posit that, in parallel with the rapid development of quantum algorithms and hardware, the continued advancement of classical approaches will enable robust, converged predictions – ideally from more than one method – for challenging molecular systems.

3. the background, present state, and future prospects of auxiliary-field quantum Monte Carlo (AFQMC) with an emphasis on chemical applications.

II. WHY MORE ACCURATE AND SCALABLE METHODS ARE STILL NEEDED IN 2022

The coupled cluster model with single, double, and perturbative triple excitations, CCSD(T),[1] is widely accepted as the "gold standard" method for small-molecule, main-group chemistry. Indeed, it has paved the way for the development of today's state-of-the-art semi-empirical density functional approximations. The most successful density functionals have $\mathcal{O}$(1-10) parameters optimized to minimize errors vs reference values, typically from CCSD(T) extrapolated to the complete basis set (CBS) limit, for main-group chemical reactions.[2,3] These computationally efficient density functionals, in turn, enabled the accurate parameterization of interatomic potentials amenable to large scale molecular simulations,[4] e.g., via fits to DFT torsional scans. Notable demonstrations of transformational technologies in drug discovery attest to the academic and commercial utility due to this heirarchy of computational methods.[5]

In recent years, significant efforts have been made to develop machine learning architectures that can either directly predict energies from structural features,[6–10] or correct mean-field or tight-binding electronic structure methods.[11–13] Machine learning has also been used to develop novel density functionals,[14] and expand

---

[a] Electronic mail: jshee@berkeley.edu
[b] Electronic mail: szhang@flatironinstitute.org

the reach of variational Monte Carlo with neural network many-body wavefunctions.[15] Although more data-efficient architectures for these applications are under rapid development,[16] often relying on the inclusion of physically motivated constraints,[17] these models generally require large amounts of training data and are thus inherently limited by the cost and accuracy of the high level methods upon which they are trained. The reliability of CCSD(T) and even many density functional approximations, as well as plentiful gas-phase experimental measurements, for small main-group molecules has enabled the development of efficient machine learning-based approaches for this region of chemical space.

However, the chemical elements found in organic molecules form a very small subset of the entire periodic table. The involvement of d and f orbitals in transition metals, lanthanides, and actinides opens up a world of reactive possibilities. Compounds with such elements play important roles in most areas of chemistry, ranging from inorganic chemistry, biochemistry, and materials science to radiochemistry, drug design, chemical catalysis, and emerging fields of molecular magnetism and quantum information science. Can computational chemists routinely perform predictive simulations – e.g. of redox events, reaction pathways, or binding free energies – with $\mathcal{O}(1)$ kcal/mol accuracy for chemical systems containing elements beyond the main group? We believe the answer is: "not yet".

The fundamental reason is a lack of accurate reference data in systems with d and f electrons, which becomes a bottleneck despite machine learning advances and progress in molecular mechanics-inspired potentials.[18,19] This is, in part, a consequence of the scarcity of reliable, gas-phase experimental data (though progress in experimental techniques is encouraging[20]), and reported measurements often are accompanied by large uncertainties. Meanwhile, the computational cost of exact theoretical approaches that in principle could yield predictions of quality comparable to experiments scales exponentially with system size. For approximate computational methods, several factors contribute to make systematically accurate calculations very challenging. d and f orbitals can give rise to strong static correlation which arises from energetic near-degeneracies in transition metal, lanthanide, and actinide complexes. This type of strong correlation, also known as multi-reference character,[21] can occur in molecules wherein spatially-localized electrons are magnetically coupled. Chemically relevant examples include bridged multi-metal clusters, reduced monometal complexes with redox-noninnocent ligands, or (poly)-radical sites in extended organic chromophores such as polyacenes.[21,22] This low-spin, partial recoupling phenomenon is more readily found in intermediate regions of bond-dissociation curves. Another complicating factor which is not typically encountered in organic molecules is that these types of orbitals can also enable complex bonding (e.g. σ, π, δ bonding) involving multiple dynamically-correlated electron pairs. Furthermore, the presence of both types of correlations is coupled with enhanced and potentially significant contributions from relativistic effects.

While CCSD(T) typically yields sub-kcal/mol accuracy for properties related to thermochemistry, kinetics, non-covalent interactions, etc. of small organic molecules, its accuracy has been questioned when it comes to transition metal systems[21,23,24] and also larger (especially polarizable) systems.[25,26] We believe that methodological advances within CC theory, including the development of multi-reference variants[27–29] and fine-tuned localization protocols,[30,31] can substantially improve the outlook for d- and f-electron systems. Many other quantum chemical methods have also been designed to treat systems with strong electronic correlations. Among them are multi-reference perturbation theories based on the driven similarity renormalization group,[32,33] non-orthogonal configuration interaction[34] with the inclusion of dynamic correlation,[35] multi-configurational density functional theory,[36] approaches based on the density matrix renormalization group,[37–39] and both selected[40,41] and stochastic[42] approaches to approximate full configuration interaction.

Regarding transition metal thermochemistry, it has not yet been convincingly demonstrated that an approximate quantum-chemical model can produce predictions on available computing resources which are accurate to within 1-2 kcal/mol of exact theoretical predictions or experimental measurements for a test set of $\mathcal{O}(100)$ cases, each containing $\mathcal{O}(10-100)$ atoms. To achieve such a goal will require continued development of the leading quantum chemical methods, and quite possibly, comparative use of more than one approximate method. The key challenge is to treat both static and dynamic correlations accurately and on an equal footing, while avoiding exponentially growing computational costs with system size. In our view, the phaseless auxiliary-field quantum Monte Carlo (ph-AFQMC) method stands out in this regard as a promising candidate.

III. AFQMC IN CONTEXT

Mature and robust electronic structure methods typically undergo decades of development, validation, and optimization. In this sense AFQMC is still relatively new compared to most methods in the standard quantum chemistry repertoire. In this section we chart the trajectory and maturation of AFQMC in the context of electronic structure theory.

Formal details are reviewed in, e.g., Refs. 43, 44, and 45. Very briefly, the lowest-energy eigenstate of a Hamiltonian can be obtained from an arbitrary initial wavefunction (with non-zero overlap with the ground state) by employing a Wick rotation, $it \to \tau$, in the time-dependent Schrödinger equation. Imaginary-time propagation will then exponentially damp the coefficients of all other eigenstates. In AFQMC, the propagator is mapped to an integral of one-body operators in auxiliary fields via the Hubbard-Stratonovich transformation. In a seminal work, Zhang and Krakauer introduced a computational framework for fermionic systems which uses Monte Carlo to sample a manifold of non-orthogonal Slater determinants while avoiding the exponential decay of signal to noise (with respect to system size or imaginary-time projection). A key idea was to reformulate the field-theoretic representation of path integrals of actions into open-ended random walks in a space of DFT-like solutions in stochastic auxiliary fields, which afforded natural connections to the formalism familiar to the electronic structure community. This makes possible[46] the introduction of a second ingredient, a constraint which imposes a gauge condition on the phases of the Slater determinants, yielding a polynomial scaling algorithm at the expense of a systematically improvable bias.[47] The resulting method is non-perturbative and naturally multi-reference.

A unique computational advantage of AFQMC, when compared with more traditional quantum chemical methods, is that it can be implemented in an essentially embarrassingly parallel way, with random walkers divided into subsets and propagated on different compute nodes. In recent years, AFQMC has undergone intense optimization efforts. Recognizing that the entire algorithm can be executed with matrix computations, ph-AFQMC has been dramatically accelerated via the use of graphical processing units (GPUs).[48,49] Another significant advance is the recent implementation of a single-parameter localized orbital approximation, which does not depend on single-reference methods to construct highly compact orbital spaces.[50] An all-electron, localized-orbital ph-AFQMC calculation of the Fe(acac)$_3$ complex with around 1000 basis functions (cc-pVTZ) and a trial wavefunction with ∼100 determinants requires only 3 hours of wall-time, running on 100 nodes (on current Summit resources). Additional algorithmic advances include efficient ways to utilize multi-determinant trial wavefunctions,[48,51] a correlated sampling approach to more efficiently converge energy differences,[52] tensor decompositions,[53,54] frozen core and downfolding techniques,[55] stochastic resolution-of-the-identity strategies,[56] constraint release methods, nuclear gradients and geometry optimization,[57,58] and the back-propagation algorithm to estimate observables which do not commute with the Hamiltonian.[59]

In the limit of an exact trial wavefunction, ph-AFQMC will recover the exact ground-state energy. However, in practice, the most important user input is the choice of approximate trial wavefunction for the constraint. There is mounting evidence that for typical main group compounds, accurate thermochemistry and even total energies within 1 kcal/mol (1.6 mHa) of exact reference values can be obtained from ph-AFQMC calculations with complete active space self consistent field (CASSCF) trial wavefunctions. For the ground-state of the carbon dimer, C_2, at equilibrium, ph-AFQMC with truncated CASSCF expansions with less than 100 determinants is able to converge to the exact total energy.[60] In the hydrogen chain, accurate results were obtained in a size-consistent manner across the entire regime of bondlengths with generalized Hartree-Fock (GHF) trial wave functions[61] which better preserve symmetry.[62]

Selected CI wavefunctions, which can approximately solve active spaces far exceeding 18 orbitals,[63] have emerged recently as promising trial wavefunctions for ph-AFQMC.[64] For C_2, convergence to the exact total energy has been shown with a selected CI trial (8e90o active space) with 10^{3-4} determinants retained.[65] In Fig. 1 we demonstrate that convergence to chemical accuracy can be achieved with selected CI trial wave functions in an active space of 8e16o with at least an order of magnitude fewer determinants. For a fixed number of determinants, it appears preferable to represent the trial wavefunction in its natural orbital basis – detailed investigations of these important subtleties are ongoing in our research groups and others.

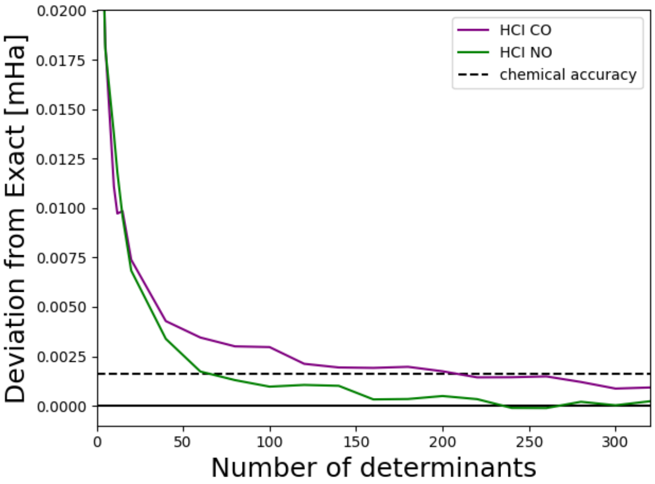

FIG. 1: Convergence of ph-AFQMC total energies with respect to the number of determinants kept in selected CI trial wavefunctions ($\epsilon_1 = 10^{-4}$, 8e16o active space) for C_2 at R=1.24253 Å. The cc-pVQZ basis is used; HCI refers to selected CI without orbital optimization in the canonical (CO) or natural orbital (NO) basis. The reference value is from a FCIQMC calculation.[66] All statistical error bars are < 0.6 mHa.

We note that other types of trial wavefunctions have been tested,[44,67–72] along with strategies which do not employ the phaseless constraint[73,74] or release it after an initial equilibration period. For large molecules and/or strongly correlated states, the use of a trial wavefunction to implement the phaseless constraint is likely to be necessary to overcome the otherwise exponentially decaying signal to noise ratio. Thus, the search for robust and scalable procedures to produce trial wavefunctions which can yield chemically-accurate ph-AFQMC predictions for large and diverse chemical systems is arguably the most momentous outstanding challenge in this field.

IV. SALIENT FRONTIERS OF QUANTUM CHEMISTRY AND THE PROMISE OF AFQMC

A. Transition Metals

$3d$ transition metal compounds are particularly challenging from an electron correlation perspective (vs $4d$ and $5d$ elements), as the relatively weak ligand field splitting can result in small energy gaps between low and high spin states. In compounds with high symmetry or low coordination number, there tends to be more degenerate or nearly-degenerate orbitals, and static correlation may arise from competing spin states,[21] leading to a wide range of predictions from DFT functionals and a breakdown of the usual route to systematic improvement (ascending Jacob's ladder for small organic molecules). In addition, bonding in organometallic complexes can involve both dispersion effects[75] and correlations between multiple electron pairs (e.g., in a metal-carbonyl bond simultaneous σ-donation and π-backbonding is a six-electron process). Such complicated dynamical correlation effects are typically not found in small organic molecules, and require a theoretical description beyond second-order perturbation theory.[21] As the high scaling of canonical CCSD and CCSD(T) render these methods prohibitively costly, those seeking accurate predictions from wavefunction methods frequently turn to localized orbital approximations, the most popular being the domain-based pair natural orbital approach.[76,77] However, the reliability of DLPNO-CCSD(T) for some transition metal systems has been drawn into question,[78,79] though extrapolation schemes appear promising.[30,31]

In light of this, we believe that another *ab initio*, potentially benchmark-quality method, which is systematically improvable and computationally feasible for realistic transition metal systems, would be highly desirable to complement and validate CC predictions. In a series of papers focused on $3d$ transition metal compounds, the present authors have demonstrated that ph-AFQMC can achieve average errors of around 1 kcal/mol and maximum errors of roughly <3 kcal/mol for atomic ionization potentials,[48] bond dissociation energies of metal-nonmetal diatomics[80] and 4-6 coordinate complexes,[81] and vertical and adiabatic ionization energies of metallocenes.[82] The mean average errors from experimental measurements are shown in Table I.

In these studies, CASSCF trial wavefunctions limited (due to resource constraints) to <18 active orbitals were used. For many systems with strong static and/or dynamic correlations, it is unlikely that this type of trial wavefunction provides sufficient flexibility to conclusively demonstrate convergence of total energies with respect to the phaseless bias. Indeed, as shown explicitly in Fig. 2 for FeO, one of the most difficult cases encountered, even with selected CI trials with all orbitals active (computationally feasible in the def2-SVP basis), convergence of the total energies to within chemical accuracy is relatively laborious. However, the energy *difference* does convergence rapidly with the percentage of the trial wavefunction retained, which is consistent with the chemically-accurate thermochemical predictions (ionization potentials, bond dissociations, etc.) shown in our transition metal publications to date. To give a sense of the computational efficiency of this result, Fe, O, and FeO with 90% of the trial wavefunction weight retained corresponds to 2, 1, and 120 determinants (and likely to fewer in the natural orbital bases).

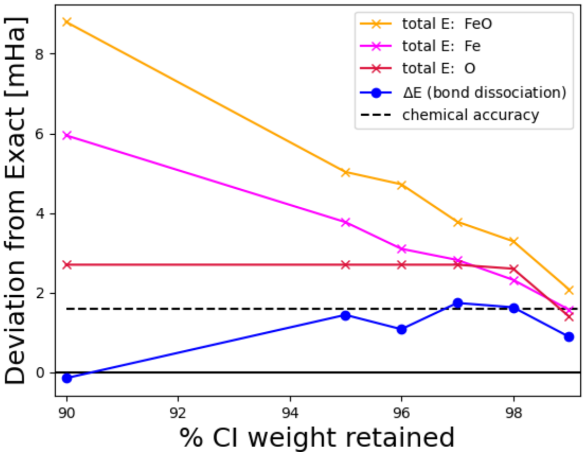

FIG. 2: Convergence of ph-AFQMC total energies with respect to the percentage of determinants kept in the SHCI trial wavefunction ($\epsilon_1 = 10^{-4}$, full active space) for FeO, Fe, and O toward the near-exact, ASCIPT2 reference values.[24] The ph-AFQMC energy difference, namely the bond dissociation energy, is within chemical accuracy over the range of percentages investigated, i.e. converges much more quickly than the total energies. All statistical error bars are <0.3 mHa. For comparison, CCSD and CCSD(T) dissociation energies are in error by 20.6 and 4.2 mHa, respectively.[24]

Furthermore, taking advantage of the fact that most quantities comparable to experiment are energy differences, a correlated sampling algorithm[52] proved very effective in accelerating convergence (with respect to statistical error bars) of the energy differences directly. Single-precision floating point arithmetic greatly accelerated the calculations on GPUs, though for larger systems than the ones investigated previously, mixed precision algorithms will likely be required.

For closed-shell and decidedly single-reference transition metal complexes, appropriately-performed DLPNO-CCSD(T) calculations are expected to produce reliable reference values, which can be used, e.g., to assess the performance of various density functional approximations. Such high quality data sets have recently been reported, such as MOR41[84] and ROST61.[85] However, for many electrochemically-relevant, open-shelled transition metal systems, there is reason to suspect that methods based on CCSD(T) may not be sufficiently accurate. For example, the CC ansatz is based on a single-determinant reference, yet UHF states are often

TABLE I: Mean averaged error [kcal/mol] of ph-AFQMC calculations from gas-phase experiments for five different $3d$ transition metal-containing test sets: 10 atomic ionization potentials (AIP10, Ref. 48), 41 bond dissociation energies for metal-nonmetal diatomics (BDE41, Ref. 80), 34 ligand dissociation energies of 4-6 coordination compounds (LDE34, Ref. 81), 6 metallocene adiabatic ionization potentials (MCIP6, Ref. 82).

Test set	MAE
AIP10	0.5(0.7)
BDE41	1.4(0.4)
LDE34	1.1(0.3)
MCIP6	1.7(1.0)

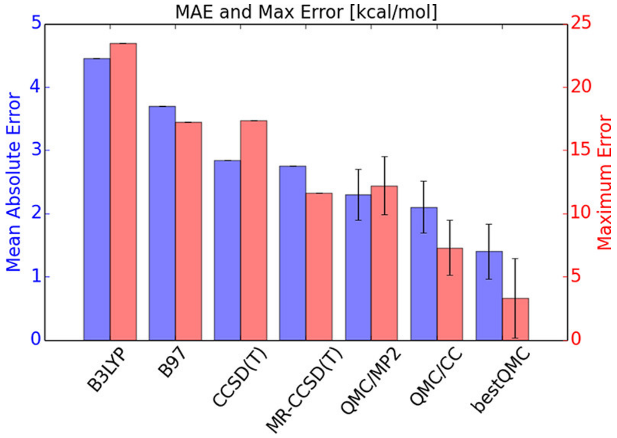

FIG. 3: Mean absolute and maximum errors of various computational methods for the BDE41 dataset, with respect to experimental measurements. "QMC/X" indicates ph-AFQMC in a triple-ζ basis extrapolated to the CBS limit with triple- and quadruple-ζ calculations using method X (MP2 or CCSD(T)). "BestQMC" indicates that some CBS extrapolations used X=ph-AFQMC. Taken from Ref. 80.

heavily spin-contaminated and it is not uncommon that ROHF calculations do not converge (or land on qualitatively wrong solutions) for spin-states of high multiplicity. Furthermore, predictions (total energies and even energy differences) from CCSD(T) can differ substantially from those from CCSD. In our view, ph-AFQMC and its localized-orbital implementation (which, unlike the DLPNO approximation, is not expected to be less efficient for multi-reference systems) can contribute a new route for accurate computations in situations of moderate to strong static correlation. A proof of concept study comparing ph-AFQMC and CCSD(T) for electrochemically relevant complexes is currently underway. Selected CI trial wavefunctions, including those generated via an iterative orbital optimization procedure,[86] with $\mathcal{O}$(10-100) active orbitals are now feasible to obtain with modest computational resources. This represents a promising route toward systematic elimination of the phaseless bias.

When converged predictions from more than one systematically improvable, non-empirical, and formally independent methods agree, the probability that all are identically wrong is vanishingly small. This multimethod approach was convincingly invoked to unambiguously characterize the ground-state charge and spin order of a microscopic model system of cuprate high-temperature superconductors.[87] In the same spirit, electronic structure properties of transition metal ions and oxide diatomics have been extensively benchmarked with a large number of first-principles methods.[83] AFQMC played an important role in both of these projects, and we believe that it will be a powerful tool going forward for quantum chemical applications. For example, when ph-AFQMC and CCSD(T) predictions agree for cases which do *not* exhibit notable multireference character, the consensus prediction is stronger than that from one method alone. This type of combined approach can potentially be the new "gold standard" and can be especially useful when experimental values do not exist.

B. The interplay of electron correlation and relativity

For heavier elements, i.e. those with occupied 4-5d and/or f orbitals, computationally tractable approximations to the full Dirac equation must be considered, complicating the job of quantum chemical modelling. An important challenge is to treat electron-electron correlations and both scalar and spin-dependent relativistic effects in an accurate and balanced manner. Currently available computational tools for such systems are significantly less robust vs nonrelativistic approaches. For example, variants of DFT incorporating scalar relativistic effects and spin-orbit coupling (SOC) to some extent yield average errors of 10 and 20 kcal/mol for thermochemistry involving relatively small actinide and lanthanide complexes, respectively.[88–90] Recent progress in all-electron basis set development and composite correlated wavefunction approaches appear more promising.[91–94]

Ref. 95 presents a relativistic many-body formalism to perform ph-AFQMC calculations on two-component Hamiltonians which include an explicit treatment of SOC. The non-relativistic ph-AFQMC framework is generalized to include additional spin-flip sectors, which doubles the problem size. The size of the Hilbert space can be reduced via use of effective core potentials and the frozen-core approximation. With single determinant trial wavefunctions, encouraging accuracy in chemical and solid-state properties has been demonstrated. For instance, the electron affinity of the Pb atom is computed to within 0.1 eV vs experiment, the bond dissociation energies of heavier halide diatomics to sub 2 kcal/mol, and the cohesive energy of solid Bi to within 0.1 eV vs experiment.

Current work involves the computation of zero-field splittings for transition metal complexes, shown in Fig. 5. A key challenge is that, unlike typical thermochemical quantities, magnetic coupling constants and zero-field splitting energy scales are on the order of 10-100 wavenumbers (roughly 0.03-0.3 kcal/mol). For large

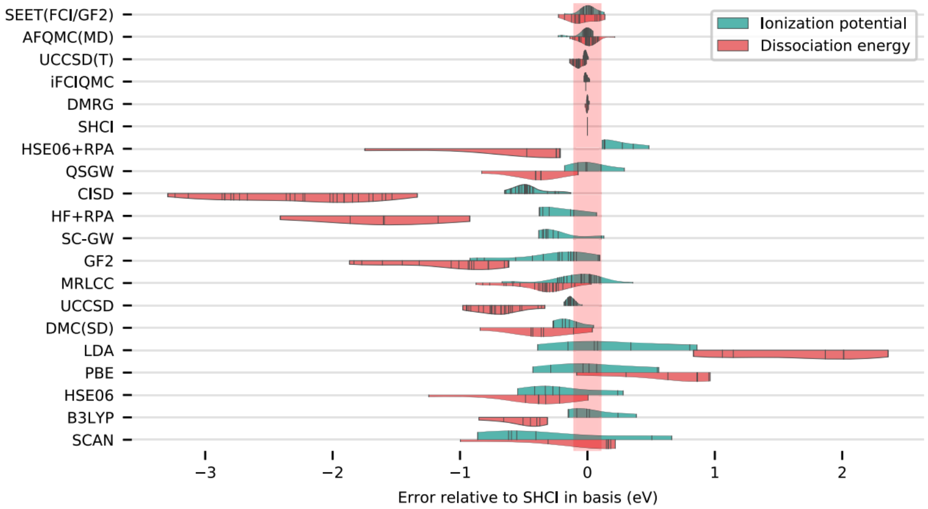

FIG. 4: Kernel density estimation plot of the single-reference subset of metal-oxide dissociation energies and atomic ionization potentials with respect to SHCI reference calculations. Taken from Ref. 83.

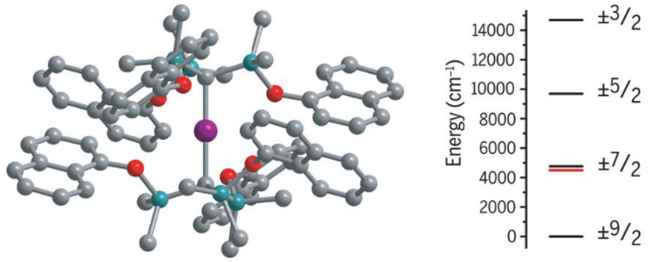

FIG. 5: (left) Molecular structure of $Co(C(SiMe_2ONaph)_3)_2$. Purple, gray, turquoise, and red spheres represent Co, C, Si, and O, respectively. Hydrogen atoms have been omitted for clarity. (right) The calculated splitting of the ground state by spin-orbit coupling. The red line is the experimentally determined energy of the $M_J = \pm 7/2$ state. Taken from Ref. 96.

molecular systems and/or those with substantial electron correlation effects, converging statistical error bars to below these thresholds implies a hefty computational cost (the error bars go as $\frac{1}{\sqrt{S}}$, where S is the number of samples). Correlated sampling is one possible route, and improved algorithms are currently being developed with such applications in mind. Alternatively, embedding approaches to downfold the full Hilbert space into a computational tractable and chemically relevant subspace[97] are compelling alternatives to solving the full molecular problem explicitly.

C. Spin gaps in photochemistry and magnetically coupled systems

The relative energy splittings among states of different spin multiplicities are critical quantities in photochemical processes. For example, photoinduced spin-conserving electronic excitations from closed-shell ground-states can be followed by inter-system crossing to a non-emissive triplet state with a relatively longer lifetime such that subsequent photochemistry can be accomplished. Alternatively, under certain conditions photoexcitation can be followed by singlet fission which produces two triplet excitons from a single photon.[98] This process, and its reverse known as triplet-triplet annihilation photon upconversion,[99] have the potential to dramatically increase solar cell efficiency.

However, the spin gaps relevant to the processes described above are typically difficult, if not impossible, to characterize experimentally. For this reason, triplet states (in general, excited states with vanishing oscillator strength with respect to a singlet ground-state, such as doubly excited singlet states) are frequently referred to as “dark”. The accurate prediction of electronic spectra including both bright and dark states represents a challenge for most popular quantum chemical methods, especially those based on DFT. For example, it is well known that relative spin state energetics are sensitive to the exchange-correlation functional employed,[77] as the inclusion of exact Hartree-Fock exchange artificially favors higher-spin states. In addition, linear response TDDFT[100] can rigorously only describe single excitations (though we note that other DFT-based approaches,[101,102] while not formally rigorous, have been found capable of yielding accurate energetics for pure doubly-excited states).

ph-AFQMC can compute spin gaps with consistently high accuracy even for relatively large molecules and/or when multi-reference states are involved. Singlet-triplet gaps of biradicaloids involving open-shell singlet states, such as the benzyne isomers shown in Fig. 6, and of polyacenes, which are among the most popular and versatile photocatalysts, have been predicted accurately using single-determinant trial wavefunctions.[70] We then demonstrated that ph-AFQMC can be a valuable tool for *in silico* design of novel upconverting annihilators. By comparing triplet energies computed with ph-AFQMC with S_1 energies, we used the thermodynamic criteria of $S_1 \geq 2\ T_1$ and $T_1^{sensitizer} > T_1^{annihilator}$ to screen myriad candidate annihilators, and our prediction that phenyl-substituted benzathiadiazole would upconvert (and that other BTDs would not) was confirmed by experimental collaborators (see Figures 7 and 8).[103] We note that the triplet-quintet gap of an iron porphyrin was accurately computed with ph-AFQMC,[68] and that a simple algorithm for computing excited-states with the same symmetry as the ground-state has been proposed.[60] (An alternate approach for AFQMC computations of general excitations has been shown to perform well in solids[104] but has not been tested in molecules).

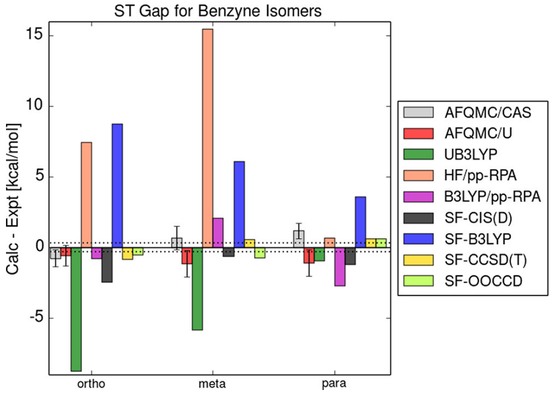

FIG. 6: Singlet-triplet gaps for ortho-, meta-, and para-benzynes relative to experimental values. Taken from Ref. 70.

These encouraging applications, which demonstrate that spin-gaps involving even multi-configurational low-spin states can be accurately predicted with a scalable algorithm, suggest that ph-AFQMC has the potential to tackle many grand-challenge problems in bioinorganic chemistry involving transition metal catalysts and magnetically-coupled metal clusters. Prominent systems include the oxygen evolving complex in Photosystem II,[105,106] and iron-sulfur centers[107,108] in the electron transport chain and nitrogenase enzymes. Such systems involve intricate electrochemical processes critical to their biological function which are not completely understood. We envision that ph-AFQMC will soon be used to elucidate the subtle reaction mechanisms involved in these natural processes, which will enable the rational design of new synthetic catalysts. Finally, accurate predictions of the spin ladders of molecular magnets – which typically contain transition metal,[109] lanthanide,[110,111] or actinide[112] elements – will enable the extraction of important properties such as magnetization barriers and magnetic susceptibilities.

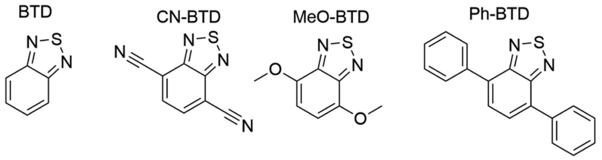

FIG. 7: Molecular structures of substituted benzothiadiazole.

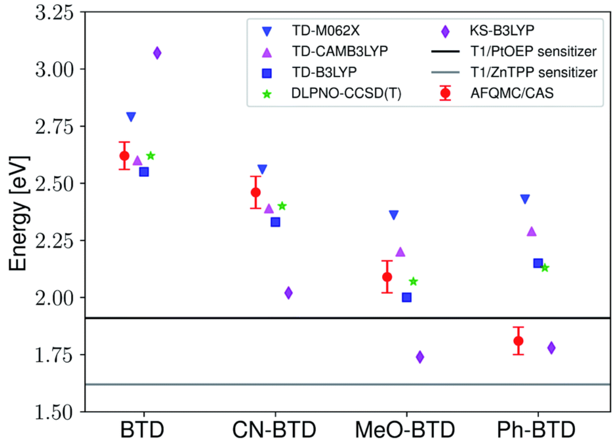

FIG. 8: Singlet-triplet gaps for BTD series. DLPNO-CCSD(T) indicates the T_0 triples procedure. Ph-BTD was found experimentally to upconvert when paired with the PtOEP sensitizer, and not with the ZnTPP sensitizer (sensitizer triplet energies are shown with the horizontal lines). Taken from Ref. 103.

V. OUTLOOK

In our view, the following properties set ph-AFQMC apart from other quantum chemical methods: the computational cost scales quartically with system size for Gaussian basis sets (and is cubic with a localized orbital implementation, or with planewave basis), the resulting accuracy is in principle systematically improvable, the algorithm exhibits near-perfect parallel efficiency and can leverage hardware such as GPUs, and the method has the flexibility to provide a balanced and accurate treatment of both static and dynamic electron correlation.

What remains to be done to have a truly robust and scalable benchmark-quality method for weakly to strongly correlated molecular systems? As mentioned, AFQMC is a relatively new member of the quantum chemistry toolbox, and as such there are a number of possible future research directions, and fruitful opportunities for development and optimization. We highlight several of these below:

1. Explore more efficient ways to reach the CBS limit. This is essential for comparing with both experimental measurements and (e.g., one-shot triple- or quadruple-ζ) DFT calculations. Approaches based on similarity transformations and explicitly correlated procedures[113–115] currently scale at least as $\mathcal{O}(N^6)$; while this scaling is acceptable for meth-

ods based on CCSD, additional approximations would be appropriate to have practical utility for large systems with AFQMC. Composite extrapolation strategies, i.e. those based on $E_{QMC}(\text{CBS}) \sim E_{QMC}(\text{small}) + [E_X(\text{large}) - E_X(\text{small})]$, where X might be, e.g., DLPNO-CCSD(T) or multireference perturbation theory,[32,116] are also promising. Finally, we note the recent optimization of specialized basis sets for correlated calculations of solid-state[117,118] and molecular[119–121] systems.

2. The combination of correlated sampling and a mechanism for branching and population control (i.e., duplicating walkers with large weights and annihilating walkers with small weights with appropriate probability) is desirable. Algorithms are currently under investigation, and have indeed shown significantly improved performance in correlated sampling in lattice models and in solids. Adapting and further developing such approaches in molecular systems will be valuable.

3. It will be useful to build up a much larger dataset to better clarify how the bias resulting from the phaseless constraint depends on the trial wavefunction employed, for a variety of different chemical systems. Is there a general, automatable protocol to generate optimal trial wave functions? What is the computationally cheapest trial wavefunction, or set of trial wavefunctions, required to converge ph-AFQMC to within chemical accuracy for a given system? This effort can complement and create synergy with the development of better trial wave function ansatzes (pseudo-BCS, symmetry-projected wave functions, non-orthogonal Slater determinants, self-consistency, etc).

4. Pople introduced the idea of a "theoretical model chemistry"[122,123] which, although approximate, makes uniform approximations and ideally satisfies a list of pre-defined properties advantageous for the description of chemical reactions. Note that, in this sense, ph-AFQMC can give rise to various model chemistries as determined by, e.g., the choice of trial wavefunction, much like the CC framework can give rise to various models such as CCSD or CCSD(T). While of course it would be desirable (and is increasingly possible, as we have shown with C_2 and FeO) to converge every single calculation independently with respect to the phaseless bias, we invite the community to consider Pople's perspective, and to make efforts to undertake broad assessments using diverse thermochemical datasets of pre-defined ph-AFQMC model chemistries.

5. Explore alternatives to the cosine projection of the phaseless constraint. During early development of ph-AFQMC, several versions were tested including the half-plane and harmonic potential, which yielded indistinguishable results from the cosine projection in jellium and simple molecules.[124] These results were all obtained with plane-wave basis functions, and different basis and Hubbard-Stratonovich choices can have an effect on the accuracy. Efforts in this direction are underway, including strategies to perform constraint-release combining phaseless with the Metropolis algorithm.

While the AFQMC community at present is expanding along other avenues – coupled electron-phonon systems,[125] finite-temperature algorithms,[126–129] quantum model systems,[87,130,131] *ab initio* calculations of solids,[118,132–135] hybrid algorithms leveraging quantum devices[72] – we hope we have made the case that ph-AFQMC has an important role to play in the field of molecular quantum chemistry. From highly accurate thermochemical predictions for transition metals and f-block compounds to mechanistic investigations of biological processes or chemical catalysis where one or more steps involve strongly correlated electronic effects, ph-AFQMC has the potential to play a unique and possibly transformative role. We view ph-AFQMC as complementary to other *ab initio* approaches, and encourage the joint use of multiple independent methods to solve scientific challenges and to provide reference values especially in the absence of experiment. In this spirit, we reiterate that in today's era of artificial intelligence and machine learning, high quality training data is the new gold. In the near future, we envision that routine ph-AFQMC predictions will enable the development of new semi-empirical density functionals and interatomic potentials, along with those based on machine-learning.

ACKNOWLEDGMENTS

J.S. acknowledges funding from the National Institute of General Medical Sciences of the National Institutes of Health under award number F32GM142231. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory. This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under contract no. DE-AC05-00OR22725. The Flatiron Institute is a division of the Simons Foundation. We thank many collaborators and colleagues who have contributed to this work and/or significantly influenced our perspective on the subject, including Martin Head-Gordon, Benjamin Rudshteyn, Evan Arthur, Mario Motta, Andreas Hansen, Hung Vuong, Pierre Devlaminck, Brandon Eskridge, Miguel Morales, Fengjie Ma, Hao Shi, and Henry Krakauer.

DATA AVAILABILITY

The data that supports the findings of this study are available within the article [and its supplementary material].

[1] R. J. Bartlett and M. Musiał, "Coupled-cluster theory in quantum chemistry," Reviews of Modern Physics **79**, 291 (2007).

[2] N. Mardirossian and M. Head-Gordon, "Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals," Molecular Physics **115**, 2315–2372 (2017).

[3] L. Goerigk, A. Hansen, C. Bauer, S. Ehrlich, A. Najibi, and S. Grimme, "A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions," Physical Chemistry Chemical Physics **19**, 32184–32215 (2017).

[4] C. Lu, C. Wu, D. Ghoreishi, W. Chen, L. Wang, W. Damm, G. A. Ross, M. K. Dahlgren, E. Russell, C. D. Von Bargen, *et al.*, "Opls4: Improving force field accuracy on challenging regimes of chemical space," Journal of Chemical Theory and Computation **17**, 4291–4300 (2021).

[5] R. Abel, L. Wang, E. D. Harder, B. Berne, and R. A. Friesner, "Advancing drug discovery through enhanced free energy calculations," Accounts of Chemical Research **50**, 1625–1632 (2017).

[6] J. Behler, "Four generations of high-dimensional neural network potentials," Chemical Reviews **121**, 10037–10072 (2021).

[7] T. W. Ko, J. A. Finkler, S. Goedecker, and J. Behler, "A fourth-generation high-dimensional neural network potential with accurate electrostatics including non-local charge transfer," Nature Communications **12**, 1–11 (2021).

[8] L. D. Jacobson, J. M. Stevenson, F. Ramezanghorbani, D. Ghoreishi, K. Leswing, E. D. Harder, and R. Abel, "Transferable neural network potential energy surfaces for closed-shell organic molecules: Extension to ions," Journal of Chemical Theory and Computation **18**, 2354–2366 (2022).

[9] O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schutt, A. Tkatchenko, and K.-R. Müller, "Machine learning force fields," Chemical Reviews **121**, 10142–10186 (2021).

[10] J. P. Janet and H. J. Kulik, "Predicting electronic structure properties of transition metal complexes with neural networks," Chemical Science **8**, 5137–5152 (2017).

[11] M. Welborn, L. Cheng, and T. F. Miller III, "Transferability in machine learning for electronic structure via the molecular orbital basis," Journal of Chemical Theory and Computation **14**, 4772–4779 (2018).

[12] T. Husch, J. Sun, L. Cheng, S. J. Lee, and T. F. Miller III, "Improved accuracy and transferability of molecular-orbital-based machine learning: Organics, transition-metal complexes, non-covalent interactions, and transition states," The Journal of Chemical Physics **154**, 064108 (2021).

[13] Z. Wang, S. Ye, H. Wang, J. He, Q. Huang, and S. Chang, "Machine learning method for tight-binding hamiltonian parameterization from ab-initio band structure," npj Computational Materials **7**, 1–10 (2021).

[14] R. Pederson, B. Kalita, and K. Burke, "Machine learning and density functional theory," Nature Reviews Physics **4**, 357–358 (2022).

[15] D. Pfau, J. S. Spencer, A. G. Matthews, and W. M. C. Foulkes, "Ab initio solution of the many-electron Schrödinger equation with deep neural networks," Physical Review Research **2**, 033429 (2020).

[16] S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt, and B. Kozinsky, "E (3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials," Nature Communications **13**, 1–11 (2022).

[17] R. Nagai, R. Akashi, and O. Sugino, "Completing density functional theory by machine learning hidden messages from molecules," npj Computational Materials **6**, 1–8 (2020).

[18] S. Dajnowicz, D. Ghoreishi, K. Modugula, W. Damm, E. D. Harder, R. Abel, L. Wang, and H. S. Yu, "Advancing free-energy calculations of metalloenzymes in drug discovery via implementation of LFMM potentials," Journal of Chemical Theory and Computation **16**, 6926–6937 (2020).

[19] M. Buchhorn, R. J. Deeth, and V. Krewald, "Revisiting the fundamental nature of metal-ligand bonding: An impartial and automated fitting procedure for angular overlap model parameters," Chemistry–A European Journal **28**, e202103775 (2022).

[20] M. D. Morse, "Predissociation measurements of bond dissociation energies," Accounts of Chemical Research **52**, 119–126 (2018).

[21] J. Shee, M. Loipersberger, D. Hait, J. Lee, and M. Head-Gordon, "Revealing the nature of electron correlation in transition metal complexes with symmetry breaking and chemical intuition," The Journal of Chemical Physics **154**, 194109 (2021).

[22] Y. Yang, E. R. Davidson, and W. Yang, "Nature of ground and electronic excited states of higher acenes," Proceedings of the National Academy of Sciences **113**, E5098–E5107 (2016).

[23] Y. A. Aoto, A. P. de Lima Batista, A. Kohn, and A. G. de Oliveira-Filho, "How to arrive at accurate benchmark values for transition metal compounds: Computation or experiment?" Journal of Chemical Theory and Computation **13**, 5291–5316 (2017).

[24] D. Hait, N. M. Tubman, D. S. Levine, K. B. Whaley, and M. Head-Gordon, "What levels of coupled cluster theory are appropriate for transition metal systems? a study using near-exact quantum chemical values for 3d transition metal binary compounds," Journal of Chemical Theory and Computation **15**, 5370–5385 (2019).

[25] Y. S. Al-Hamdani, P. R. Nagy, A. Zen, D. Barton, M. Kállay, J. G. Brandenburg, and A. Tkatchenko, "Interactions between large molecules pose a puzzle for reference quantum mechanical methods," Nature Communications **12**, 1–12 (2021).

[26] B. D. Nguyen, G. P. Chen, M. M. Agee, A. M. Burow, M. P. Tang, and F. Furche, "Divergence of many-body perturbation theory for noncovalent interactions of large molecules," Journal of Chemical Theory and Computation **16**, 2258–2273 (2020).

[27] D. I. Lyakh, M. Musiał, V. F. Lotrich, and R. J. Bartlett, "Multireference nature of chemistry: The coupled-cluster view," Chemical Reviews **112**, 182–243 (2012).

[28] J. Paldus, "Externally and internally corrected coupled cluster approaches: an overview," Journal of Mathematical Chemistry **55**, 477–502 (2017).

[29] G. J. Aroeira, M. M. Davis, J. M. Turney, and H. F. Schaefer III, "Coupled cluster externally corrected by adaptive configuration interaction," Journal of Chemical Theory and Computation **17**, 182–190 (2020).

[30] A. Altun, F. Neese, and G. Bistoni, "Extrapolation to the limit of a complete pair natural orbital space in local coupled-cluster calculations," Journal of Chemical Theory and Computation **16**, 6142–6149 (2020).

[31] M. Drosou, C. A. Mitsopoulou, and D. A. Pantazis, "Reconciling local coupled cluster with multireference approaches for transition metal spin-state energetics," Journal of Chemical Theory and Computation (2022).

[32] C. Li and F. A. Evangelista, "Multireference theories of electron correlation based on the driven similarity renormalization group," Annual Review of Physical Chemistry **70**, 245–273 (2019).

[33] S. Battaglia, L. Fransén, I. Fdez. Galván, and R. Lindh, "Regularized CASPT2: an intruder-state-free approach," Journal of Chemical Theory and Computation (2022).

[34] N. J. Mayhall, P. R. Horn, E. J. Sundstrom, and M. Head-Gordon, "Spin–flip non-orthogonal configuration interaction: a variational and almost black-box method for describing strongly correlated molecules," Physical Chemistry Chemical Physics **16**,

22694–22705 (2014).

[35] H. G. Burton and A. J. Thom, "Reaching full correlation through nonorthogonal configuration interaction: A second-order perturbative approach," Journal of Chemical Theory and Computation **16**, 5586–5600 (2020).

[36] P. Sharma, J. J. Bao, D. G. Truhlar, and L. Gagliardi, "Multiconfiguration pair-density functional theory," Annual Review of Physical Chemistry **72**, 541–564 (2021).

[37] S. R. White and R. L. Martin, "Ab initio quantum chemistry using the density matrix renormalization group," The Journal of Chemical Physics **110**, 4127–4130 (1999).

[38] G. K.-L. Chan and S. Sharma, "The density matrix renormalization group in quantum chemistry," Annual Review of Physical Chemistry **62**, 465–481 (2011).

[39] L. Freitag and M. Reiher, "The density matrix renormalization group for strong correlation in ground and excited states," Quantum Chemistry and Dynamics of Excited States: Methods and Applications , 205–245 (2020).

[40] S. Sharma, A. A. Holmes, G. Jeanmairet, A. Alavi, and C. J. Umrigar, "Semistochastic heat-bath configuration interaction method: Selected configuration interaction with semistochastic perturbation theory," Journal of Chemical Theory and Computation **13**, 1595–1604 (2017).

[41] N. M. Tubman, C. D. Freeman, D. S. Levine, D. Hait, M. Head-Gordon, and K. B. Whaley, "Modern approaches to exact diagonalization and selected configuration interaction with the adaptive sampling CI method," Journal of Chemical Theory and Computation **16**, 2139–2159 (2020).

[42] G. H. Booth, A. J. Thom, and A. Alavi, "Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space," The Journal of Chemical Physics **131**, 054106 (2009).

[43] M. Motta and S. Zhang, "Ab initio computations of molecular systems by the auxiliary-field quantum Monte Carlo method," Wiley Interdisciplinary Reviews: Computational Molecular Science **8**, e1364 (2018).

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

[45] S. Zhang, "Ab initio electronic structure calculations by auxiliary-field quantum Monte Carlo," Handbook of Materials Modeling: Methods: Theory and Modeling , 123–149 (2020).

[46] E. Pavarini, E. Koch, and S. Zhang, "Many-body methods for real materials," Tech. Rep. (Theoretische Nanoelektronik, 2019).

[47] S. Zhang and H. Krakauer, "Quantum Monte Carlo method using phase-free random walks with Slater determinants," Physical Review Letters **90**, 136401 (2003).

[48] J. Shee, E. J. Arthur, S. Zhang, D. R. Reichman, and R. A. Friesner, "Phaseless auxiliary-field quantum Monte Carlo on graphical processing units," Journal of Chemical Theory and Computation **14**, 4109–4121 (2018).

[49] F. D. Malone, S. Zhang, and M. A. Morales, "Accelerating auxiliary-field quantum Monte Carlo simulations of solids with graphical processing units," Journal of Chemical Theory and Computation **16**, 4286–4297 (2020).

[50] J. L. Weber, H. Vuong, P. A. Devlaminck, J. Shee, J. Lee, D. R. Reichman, and R. A. Friesner, "A localized-orbital energy evaluation for auxiliary-field quantum Monte Carlo," Journal of Chemical Theory and Computation (2022).

[51] A. Mahajan and S. Sharma, "Efficient local energy evaluation for multi-Slater wave functions in orbital space quantum Monte Carlo," The Journal of Chemical Physics **153**, 194108 (2020).

[52] J. Shee, S. Zhang, D. R. Reichman, and R. A. Friesner, "Chemical transformations approaching chemical accuracy via correlated sampling in auxiliary-field quantum Monte Carlo," Journal of Chemical Theory and Computation **13**, 2667–2680 (2017).

[53] M. Motta, J. Shee, S. Zhang, and G. K.-L. Chan, "Efficient ab initio auxiliary-field quantum Monte Carlo calculations in Gaussian bases via low-rank tensor decomposition," Journal of Chemical Theory and Computation **15**, 3510–3521 (2019).

[54] F. D. Malone, S. Zhang, and M. A. Morales, "Overcoming the memory bottleneck in auxiliary field quantum Monte Carlo simulations with interpolative separable density fitting," Journal of Chemical Theory and Computation **15**, 256–264 (2018).

[55] W. Purwanto, S. Zhang, and H. Krakauer, "Frozen-orbital and downfolding calculations with auxiliary-field quantum Monte Carlo," Journal of Chemical Theory and Computation **9**, 4825–4833 (2013).

[56] J. Lee and D. R. Reichman, "Stochastic resolution-of-the-identity auxiliary-field quantum Monte Carlo : Scaling reduction without overhead," The Journal of Chemical Physics **153**, 044131 (2020).

[57] M. Motta and S. Zhang, "Communication: Calculation of interatomic forces and optimization of molecular geometry with auxiliary-field quantum Monte Carlo," The Journal of Chemical Physics **148**, 181101 (2018).

[58] S. Chen and S. Zhang, "A structural optimization algorithm with stochastic forces and stresses," arXiv preprint arXiv:2204.12074 (2022).

[59] M. Motta and S. Zhang, "Computation of ground-state properties in molecular systems: Back-propagation with auxiliary-field quantum Monte Carlo," Journal of Chemical Theory and Computation **13**, 5367–5378 (2017).

[60] W. Purwanto, S. Zhang, and H. Krakauer, "Excited state calculations using phaseless auxiliary-field quantum Monte Carlo : Potential energy curves of low-lying C_2 singlet states," The Journal of Chemical Physics **130**, 094107 (2009).

[61] M. Motta, D. M. Ceperley, G. K.-L. Chan, J. A. Gomez, E. Gull, S. Guo, C. A. Jiménez-Hoyos, T. N. Lan, J. Li, F. Ma, *et al.*, "Towards the solution of the many-electron problem in real materials: Equation of state of the hydrogen chain with state-of-the-art many-body methods," Physical Review X **7**, 031059 (2017).

[62] H. Shi and S. Zhang, "Symmetry in auxiliary-field quantum Monte Carlo calculations," Physical Review B **88**, 125132 (2013).

[63] A. A. Holmes, N. M. Tubman, and C. Umrigar, "Heat-bath configuration interaction: An efficient selected configuration interaction algorithm inspired by heat-bath sampling," Journal of Chemical Theory and Computation **12**, 3674–3680 (2016).

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

[65] J. Lee, H. Q. Pham, and D. R. Reichman, "Twenty years of auxiliary-field quantum Monte Carlo in quantum chemistry: An overview and assessment on main group chemistry and bond-breaking," arXiv preprint arXiv:2208.01280 (2022).

[66] S. M. Greene, R. J. Webber, J. E. Smith, J. Weare, and T. C. Berkelbach, "Full configuration interaction excited-state energies in large active spaces from randomized subspace iteration," arXiv preprint arXiv:2201.12164 (2022).

[67] E. J. Landinez Borda, J. Gomez, and M. A. Morales, "Non-orthogonal multi-Slater determinant expansions in auxiliary field quantum Monte Carlo," The Journal of Chemical Physics **150**, 074105 (2019).

[68] J. Lee, F. D. Malone, and M. A. Morales, "Utilizing essential symmetry breaking in auxiliary-field quantum Monte Carlo : Application to the spin gaps of the C_{36} fullerene and an iron porphyrin model complex," Journal of Chemical Theory and Computation **16**, 3019–3027 (2020).

[69] M. Qin, H. Shi, and S. Zhang, "Coupling quantum Monte Carlo and independent-particle calculations: Self-consistent constraint for the sign problem based on the density or the density matrix," Physical Review B **94**, 235119 (2016).

[70] J. Shee, E. J. Arthur, S. Zhang, D. R. Reichman, and R. A. Friesner, "Singlet–triplet energy gaps of organic biradicals and polyacenes with auxiliary-field quantum Monte Carlo," Journal of Chemical Theory and Computation **15**, 4924–4932 (2019).

[71] Z.-Y. Xiao, H. Shi, and S. Zhang, "Pseudo-BCS wave function from density matrix decomposition: Application in auxiliary-field quantum Monte Carlo," Physical Review Research **3**, 013065 (2021).

[72] W. J. Huggins, B. A. O'Gorman, N. C. Rubin, D. R. Reichman, R. Babbush, and J. Lee, "Unbiasing fermionic quantum Monte Carlo with a quantum computer," Nature **603**, 416–420 (2022).

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

[74] A. Mahajan and S. Sharma, "Taming the sign problem in auxiliary-field quantum Monte Carlo using accurate wave functions," Journal of Chemical Theory and Computation **17**, 4786–4798 (2021).

[75] M. Bursch, E. Caldeweyher, A. Hansen, H. Neugebauer, S. Ehlert, and S. Grimme, "Understanding and quantifying london dispersion effects in organometallic complexes," Accounts of Chemical Research **52**, 258–266 (2018).

[76] C. Riplinger, B. Sandhoefer, A. Hansen, and F. Neese, "Natural triple excitations in local coupled cluster calculations with pair natural orbitals," The Journal of Chemical Physics **139**, 134101 (2013).

[77] B. M. Floser, Y. Guo, C. Riplinger, F. Tuczek, and F. Neese, "Detailed pair natural orbital-based coupled cluster studies of spin crossover energetics," Journal of Chemical Theory and Computation **16**, 2224–2235 (2020).

[78] I. Sandler, J. Chen, M. Taylor, S. Sharma, and J. Ho, "Accuracy of DLPNO-CCSD(T): Effect of basis set and system size," The Journal of Physical Chemistry A **125**, 1553–1563 (2021).

[79] Q. M. Phung, C. Martín-Fernández, J. N. Harvey, and M. Feldt, "Ab initio calculations for spin-gaps of non-heme iron complexes," Journal of Chemical Theory and Computation **15**, 4297–4304 (2019).

[80] J. Shee, B. Rudshteyn, E. J. Arthur, S. Zhang, D. R. Reichman, and R. A. Friesner, "On achieving high accuracy in quantum chemical calculations of 3d 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 **15**, 2346–2358 (2019).

[81] B. Rudshteyn, D. Coskun, J. L. Weber, E. J. Arthur, S. Zhang, D. R. Reichman, R. A. Friesner, and J. Shee, "Predicting ligand-dissociation energies of 3d coordination complexes with auxiliary-field quantum Monte Carlo," Journal of Chemical Theory and Computation **16**, 3041–3054 (2020).

[82] B. Rudshteyn, J. L. Weber, D. Coskun, P. A. Devlaminck, S. Zhang, D. R. Reichman, J. Shee, and R. A. Friesner, "Calculation of metallocene ionization potentials via auxiliary field quantum Monte Carlo : Toward benchmark quantum chemistry for transition metals," Journal of Chemical Theory and Computation **18**, 2845–2862 (2022).

[83] K. T. Williams, Y. Yao, J. Li, L. Chen, H. Shi, M. Motta, C. Niu, U. Ray, S. Guo, R. J. Anderson, *et al.*, "Direct comparison of many-body methods for realistic electronic Hamiltonians," Physical Review X **10**, 011041 (2020).

[84] S. Dohm, A. Hansen, M. Steinmetz, S. Grimme, and M. P. Checinski, "Comprehensive thermochemical benchmark set of realistic closed-shell metal organic reactions," Journal of Chemical Theory and Computation **14**, 2596–2608 (2018).

[85] L. R. Maurer, M. Bursch, S. Grimme, and A. Hansen, "Assessing density functional theory for chemically relevant open-shell transition metal reactions," Journal of Chemical Theory and Computation **17**, 6134–6151 (2021).

[86] J. E. Smith, B. Mussard, A. A. Holmes, and S. Sharma, "Cheap and near exact casscf with large active spaces," Journal of Chemical Theory and Computation **13**, 5468–5478 (2017).

[87] B.-X. Zheng, C.-M. Chung, P. Corboz, G. Ehlers, M.-P. Qin, R. M. Noack, H. Shi, S. R. White, S. Zhang, and G. K.-L. Chan, "Stripe order in the underdoped region of the two-dimensional Hubbard model," Science **358**, 1155–1160 (2017).

[88] L. E. Aebersold and A. K. Wilson, "Considering density functional approaches for actinide species: The An66 molecule set," The Journal of Physical Chemistry A **125**, 7029–7037 (2021).

[89] S. Grimmel, G. Schoendorff, and A. K. Wilson, "Gauging the performance of density functionals for lanthanide-containing molecules," Journal of Chemical Theory and Computation **12**, 1259–1266 (2016).

[90] L. E. Aebersold, S. H. Yuwono, G. Schoendorff, and A. K. Wilson, "Efficacy of density functionals and relativistic effective core potentials for lanthanide-containing species: The Ln54 molecule set," Journal of Chemical Theory and Computation **13**, 2831–2839 (2017).

[91] D. A. Pantazis and F. Neese, "All-electron scalar relativistic basis sets for the lanthanides," Journal of Chemical Theory and Computation **5**, 2229–2238 (2009).

[92] K. A. Peterson, "Correlation consistent basis sets for actinides. i. the Th and U atoms," The Journal of Chemical Physics **142**, 074105 (2015).

[93] R. Feng and K. A. Peterson, "Correlation consistent basis sets for actinides. ii. the atoms Ac and Np–Lr," The Journal of Chemical Physics **147**, 084108 (2017).

[94] S. C. North and A. K. Wilson, "Ab initio composite approaches for heavy element energetics: Ionization potentials for the actinide series of elements," The Journal of Physical Chemistry A **126**, 3027–3042 (2022).

[95] B. Eskridge, H. Krakauer, H. Shi, and S. Zhang, "Ab initio calculations in atoms, molecules, and solids, treating spin–orbit coupling and electron interaction on an equal footing," The Journal of Chemical Physics **156**, 014107 (2022).

[96] P. C. Bunting, M. Atanasov, E. Damgaard-Møller, M. Perfetti, I. Crassee, M. Orlita, J. Overgaard, J. van Slageren, F. Neese, and J. R. Long, "A linear cobalt(II) complex with maximal orbital angular momentum from a non-Aufbau ground state," Science **362**, eaat7319 (2018).

[97] B. Eskridge, H. Krakauer, and S. Zhang, "Local embedding and effective downfolding in the auxiliary-field quantum Monte Carlo method," Journal of Chemical Theory and Computation **15**, 3949–3959 (2019).

[98] M. B. Smith and J. Michl, "Singlet fission," Chemical Reviews **110**, 6891–6936 (2010).

[99] J. Zhou, Q. Liu, W. Feng, Y. Sun, and F. Li, "Upconversion luminescent materials: advances and applications," Chemical Reviews **115**, 395–465 (2015).

[100] A. Dreuw and M. Head-Gordon, "Single-reference ab initio methods for the calculation of excited states of large molecules," Chemical Reviews **105**, 4009–4037 (2005).

[101] D. Hait and M. Head-Gordon, "Excited state orbital optimization via minimizing the square of the gradient: General approach and application to singly and doubly excited states via density functional theory," Journal of Chemical Theory and Computation **16**, 1699–1710 (2020).

[102] A. T. Gilbert, N. A. Besley, and P. M. Gill, "Self-consistent field calculations of excited states using the maximum overlap method (MOM)," The Journal of Physical Chemistry A **112**, 13164–13171 (2008).

[103] J. L. Weber, E. M. Churchill, S. Jockusch, E. J. Arthur, A. B. Pun, S. Zhang, R. A. Friesner, L. M. Campos, D. R. Reichman, and J. Shee, "In silico prediction of annihilators for triplet–triplet annihilation upconversion via auxiliary-field quantum Monte Carlo," Chemical Science **12**, 1068–1079 (2021).

[104] F. Ma, S. Zhang, and H. Krakauer, "Excited state calculations in solids by auxiliary-field quantum Monte Carlo," New Journal of Physics **15**, 093017 (2013).

[105] M. Askerka, G. W. Brudvig, and V. S. Batista, "The O2-evolving complex of photosystem II: Recent insights from quantum mechanics/molecular mechanics (QM/MM), extended X-ray absorption fine structure (EXAFS), and femtosecond X-ray crystallography data," Accounts of Chemical Research **50**, 41–48 (2017).

[106] V. Krewald, M. Retegan, F. Neese, W. Lubitz, D. A. Pantazis, and N. Cox, "Spin state as a marker for the structural evolution of nature's water-splitting catalyst," Inorganic Chemistry **55**, 488–501 (2016).

[107] S. Sharma, K. Sivalingam, F. Neese, and G. K. Chan, "Low-energy spectrum of iron–sulfur clusters directly from many-particle quantum mechanics," Nature Chemistry **6**, 927–933 (2014).

[108] Z. Li, S. Guo, Q. Sun, and G. K. Chan, "Electronic landscape of the p-cluster of nitrogenase as revealed through many-electron quantum wavefunction simulations," Nature Chemistry **11**, 1026–1033 (2019).

[109] M. Atanasov, D. Aravena, E. Suturina, E. Bill, D. Maganas, and F. Neese, "First principles approach to the electronic structure, magnetic anisotropy and spin relaxation in mononuclear 3d-transition metal single molecule magnets," Coordination Chemistry Reviews **289**, 177–214 (2015).

[110] D. N. Woodruff, R. E. Winpenny, and R. A. Layfield, "Lanthanide single-molecule magnets," Chemical Reviews **113**, 5110–5148 (2013).

[111] S. T. Liddle and J. van Slageren, "Improving f-element single molecule magnets," Chemical Society Reviews **44**, 6655–6669 (2015).

[112] K. R. Meihaus and J. R. Long, "Actinide-based single-molecule magnets," Dalton Transactions **44**, 2517–2528 (2015).

[113] W. Kutzelnigg, "r12-dependent terms in the wave function as closed sums of partial wave amplitudes for large l," Theoretica Chimica Acta **68**, 445–469 (1985).

[114] T. Yanai and T. Shiozaki, "Canonical transcorrelated theory with projected Slater-type geminals," The Journal of Chemical Physics **136**, 084107 (2012).

[115] Q. Ma and H.-J. Werner, "Explicitly correlated local coupled-cluster methods using pair natural orbitals," Wiley Interdisciplinary Reviews: Computational Molecular Science **8**, e1371 (2018).

[116] K. Andersson, P.-Å. Malmqvist, and B. O. Roos, "Second-order perturbation theory with a complete active space self-consistent field reference function," The Journal of Chemical Physics **96**, 1218–1226 (1992).

[117] H.-Z. Ye and T. C. Berkelbach, "Correlation-consistent Gaussian basis sets for solids made simple," Journal of Chemical Theory and Computation **18**, 1595–1606 (2022).

[118] M. A. Morales and F. D. Malone, "Accelerating the convergence of auxiliary-field quantum Monte Carlo in solids with optimized Gaussian basis sets," The Journal of Chemical Physics **153**, 194111 (2020).

[119] S. R. White, "Hybrid grid/basis set discretizations of the Schrödinger equation," The Journal of Chemical Physics **147**, 244102 (2017).

[120] S. R. White and E. M. Stoudenmire, "Multisliced gausslet basis sets for electronic structure," Physical Review B **99**, 081110 (2019).

[121] Y. Qiu and S. R. White, "Hybrid gausslet/Gaussian basis sets," The Journal of Chemical Physics **155**, 184107 (2021).

[122] J. A. Pople, J. S. Binkley, and R. Seeger, "Theoretical models incorporating electron correlation," International Journal of Quantum Chemistry **10**, 1–19 (1976).

[123] J. A. Pople, "Nobel lecture: Quantum chemical models," Reviews of Modern Physics **71**, 1267 (1999).

[124] S. Zhang, H. Krakauer, W. A. Al-Saidi, and M. Suewattana, "Quantum simulations of realistic systems by auxiliary fields," Computer Physics Communications **169**, 394–399 (2005).

[125] J. Lee, S. Zhang, and D. R. Reichman, "Constrained-path auxiliary-field quantum Monte Carlo for coupled electrons and phonons," Physical Review B **103**, 115123 (2021).

[126] Y.-Y. He, M. Qin, H. Shi, Z.-Y. Lu, and S. Zhang, "Finite-temperature auxiliary-field quantum Monte Carlo : Self-consistent constraint and systematic approach to low temperatures," Physical Review B **99**, 045108 (2019).

[127] Y. Liu, M. Cho, and B. Rubenstein, "Ab initio finite temperature auxiliary field quantum Monte Carlo," Journal of Chemical Theory and Computation **14**, 4722–4732 (2018).

[128] T. Shen, Y. Liu, Y. Yu, and B. M. Rubenstein, "Finite temperature auxiliary field quantum Monte Carlo in the canonical ensemble," The Journal of Chemical Physics **153**, 204108 (2020).

[129] J. Lee, M. A. Morales, and F. D. Malone, "A phaseless auxiliary-field quantum Monte Carlo perspective on the uniform electron gas at finite temperatures: Issues, observations, and benchmark study," The Journal of Chemical Physics **154**, 064109 (2021).

[130] E. Vitali, H. Shi, A. Chiciak, and S. Zhang, "Metal-insulator transition in the ground state of the three-band Hubbard model at half filling," Physical Review B **99**, 165116 (2019).

[131] M. Qin, C.-M. Chung, H. Shi, E. Vitali, C. Hubig, U. Schollwöck, S. R. White, S. Zhang, *et al.*, "Absence of superconductivity in the pure two-dimensional Hubbard model," Physical Review X **10**, 031016 (2020).

[132] F. Ma, W. Purwanto, S. Zhang, and H. Krakauer, "Quantum Monte Carlo calculations in solids with downfolded Hamiltonians," Physical Review Letters **114**, 226401 (2015).

[133] S. Zhang, F. D. Malone, and M. A. Morales, "Auxiliary-field quantum Monte Carlo calculations of the structural properties of nickel oxide," The Journal of Chemical Physics **149**, 164102 (2018).

[134] F. D. Malone, A. Benali, M. A. Morales, M. Caffarel, P. R. Kent, and L. Shulenburger, "Systematic comparison and cross-validation of fixed-node diffusion Monte Carlo and phaseless auxiliary-field quantum Monte Carlo in solids," Physical Review B **102**, 161104 (2020).

[135] S. Chen, M. Motta, F. Ma, and S. Zhang, "Ab initio electronic density in solids by many-body plane-wave auxiliary-field quantum Monte Carlo calculations," Physical Review B **103**, 075138 (2021).