I. DETAILED METHODOLOGY

A. AFQMC

The AFQMC method [1–3] estimates the ground-state properties of a many-fermion system by statistically sampling the wave function $|\psi_g\rangle \propto e^{-\beta\hat{H}}|\psi^0\rangle$, where $|\psi^0\rangle$ is an initial wave function which is nonorthogonal to the ground state. The projection is carried out iteratively by small time step $e^{-\Delta\tau\hat{H}}$, with $\beta = n\Delta\tau$ sufficiently large to project out all excited states. The propagator is represented as $e^{-\Delta\tau\hat{H}} = \int d\boldsymbol{x} p(\boldsymbol{x})\hat{B}(\boldsymbol{x})$ where $\hat{B}(\boldsymbol{x})$ is a one-body operator which depends on the vector $\boldsymbol{x}$, and $p(\boldsymbol{x})$ is a probability distribution. This representation maps a many-body system into an ensemble of one-body systems, with the ensemble then sampled by Monte Carlo (MC) techniques. We use open-ended random walks in Slater determinant space to sample the imaginary time projection and represent the ground state wave function: $|\psi_g\rangle = \int d\phi\, c_\phi|\phi\rangle$, where the Slater determinants in the integral are non-orthogonal $\langle\phi'|\phi\rangle \neq 0$. A gauge constraint, implemented approximately with a trial wave function $|\psi_T\rangle$, is applied on the sampled Slater determinants [1, 2] to control the sign or phase problem.

In this work, we present results obtained from the AFQMC method implemented for Gaussian basis sets [4, 5]. We set the linear dependence threshold to be 10^{-8} for the one-electron basis [6] and use the modified Cholesky decomposition [7] with a threshold 10^{-6} for the Coulomb interaction. Most of the calculations use projection time $\beta = 35\ E_{Ha}^{-1}$ and time step $\Delta\tau = 0.005\ E_{Ha}^{-1}$. The convergence error from finite β is negligible and extrapolations are performed when the Trotter error is larger than Monte Carlo uncertainty. The reported error bars are estimated by one standard deviation statistical errors.

Truncated CASSCF wave functions were used as $|\psi_T\rangle$ here. Fast update procedure [8, 9] allows the use of multi-determinant CAS trial wave functions with sublinear cost. (For example, in ScO TZ basis, the cost of the AFQMC calculation with a $|\psi_T\rangle$ of 163 determinants is 2.1× that of a single determinant calculation.) Typically around 10 CAS orbitals are used to generate the $|\psi_T\rangle$ in both atoms and molecules. Following procedures in past AFQMC calculations using CASSCF [10, 11], we truncate the wave function by discarding determinants with the smallest weights, up to an integrated weight of $\delta = 10^{-3}$, which results in ~ 100 determinants in most cases. Because of the fast update algorithm, we could check the effect on the AFQMC results of increasing the CAS space to the next level with little cost. In VO and FeO, a noticeable difference was seen outside the statistical error, and we increased the CAS space to 12, resulting in ~ 1700 and ~ 1600 determinants, respectively, in their $|\psi_T\rangle$.

In solids, CASSCF trial wave functions would not be applicable straightforwardly in a size-consistent manner. There have been many benchmark studies of AFQMC using single determinant trial wave functions (e.g. Refs. [10–12]). To give an idea of the dependence of the constraint error in AFQMC for the specific systems here, the computed total energies in TiO, VO, CrO, MnO change by about −4.9, −1.5, +6.8, and −9.7 milli-Hartrees from the reported multi-det $|\psi_T\rangle$ results when a single determinant UHF trial wave function is used (TZ basis, with MC error bars 1-2 milli-Hartrees). Besides using a single Slater determinant, there are a number of possibilities to systematically improve the trial wave function, including generalized Hartree-Fock [13], symmetry projection and symmetry-adapted multi-determinants [6, 14, 15], Hartree-Fock-Bogoliubov form [16, 17], self-consistent trial wave functions [18, 19].

B. Configuration interaction

Configuration interaction with singles and doubles excitations (CISD) was used as implemented in the PySCF package. CISD approximates the many-body wave function as a sum of Slater determinants, constructed from a reference Slater determinant, which was taken from restricted open shell Hartree-Fock. In CISD, the wave function is given as

$$\begin{aligned}|\Psi_{CISD}\rangle = (&c_0 \\ &+ \sum_{ij,\sigma} C^{(s)}_{ai} c^\dagger_{a,\sigma} c_{i,\sigma} \\ &+ \sum_{ijkl,\sigma,\sigma'} C^{(d)}_{abij,\sigma,\sigma'} c^\dagger_{a,\sigma} c^\dagger_{b,\sigma'} c_{i,\sigma} c_{j\sigma'} \\ &)|\Psi_{HF}\rangle,\end{aligned}$$

where $c^\dagger$ and c are creation/destruction operators, respectively, a, b refer to virtual orbitals, i, j refer to occupied orbitals, and all C parameters are variationally optimized. CISD scales approximately as $\mathcal{O}(N_e^6)$ and is known to not

TABLE I. Calculations in the database. In each cell, the symbols correspond to the basis performed as follows: d: vdz, t: vtz, q: vqz, 5: v5z, c: cbs. A dash means that there were no calculations for that system using that technique.

	System																						
Method	Cr	Cr+	CrO	Cu	Cu+	CuO	Fe	Fe+	FeO	Mn	Mn+	MnO	O	O+	Sc	Sc+	ScO	Ti	Ti+	TiO	V	V+	VO
AFQMC(MD)	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5
B3LYP	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5
CISD	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5
DMC(SD)	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c
DMRG	dt	dt	d	dt	dt	d	dt	dt	d	dt	dt	d	dt	dt	dt	dt	d	dt	dt	d	dt	dt	d
GF2	dtq	dtq	dtq	—	—	dtq	dtq	dtq	dtq	dtq	dtq	dt	dtq	dtq	dtq	dtq	dtq	dtq	dtq	dtq	dtq	dtq	dtq
HF+RPA	t	t	t	t	t	—	dt	t	t	t	t	t	dtq	dtq	t	t	t	t	t	t	t	t	t
HSE06	dtq5	dtq5	dtq5	dtq5	dtq5	—	ddtq5	dtq5	dtq5	dtq5	dtq5	dtq5	ddtq5q	ddtq5q	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5
LDA	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5
MRLCC	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5
PBE	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5
PBE+RPA	t	t	t	t	t	—	t	t	t	t	t	t	dtq	dtq	t	t	t	t	t	t	t	t	t
QSGW	t	t	t	—	—	—	dt	t	t	t	t	t	dtq	dtq	t	t	t	t	t	t	t	t	t
SC-GW	d	d	—	—	—	—	d	d	—	d	d	—	d	d	d	d	—	d	d	—	d	d	—
SCAN	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5
SEET(FCI/GF2)	dtq	dtq	dq	—	—	dtq	dtq	dtq	dtq	dtq	dtq	—	dtq	dtq	dtq	dtq	dtq	dtq	dtq	dtq	dtq	dtq	dtq
SHCI	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5
UCCSD	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5
UCCSD(T)	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5	dtq5
iFCIQMC	dtq	dtq	—	dtq	dtq	—	dtq	dtq	—	dtq	dtq	—	dtq	dtq	dtq	dtq	d	dtq	dtq	—	dtq	dtq	—

be size extensive.

C. Coupled Cluster

Unrestricted coupled cluster was used as implemented in the PySCF package.[20] The reference state was restricted open shell Hartree-Fock. We found that an unrestricted reference state led to worse results by values up to 10 mHa for the total energy of the molecules. In UCCSD, the wave function is approximated as

$$|\Psi_{CCSD}\rangle = e^{\hat{T}}|\Psi_{HF}\rangle, \tag{1}$$

where the $\hat{T}$ operator contains one and two body operators. The exponential *ansatz* ensures that the technique is size extensive in contrast to CISD. UCCSD scales approximately as $\mathcal{O}(N_e^6)$ [21].

UCCSD(T) evaluates the perturbative effect of including three-body operators from a UCCSD reference, and is often called the gold standard of quantum chemistry when used for equilibrium properties. It scales approximately as $\mathcal{O}(N_e^7)$. Despite the steep formal scaling, the prefactor is quite small. Thus compared to the other accurate methods in this paper, namely, SHCI, DMRG, FCIQMC, AFQMC, and SEET, the UCCSD(T) calculations were the least expensive by a significant amount.

D. Density Functional Theory

Density functional theory (DFT) in the restricted open shell Kohn-Sham approach was used as implemented in the PySCF package.[20] Level 6 grids were used to improve the accuracy, and the resultant state was carefully checked to ensure that it was the DFT ground state, since often the self consistent field process converged to the incorrect state. The basis set error in DFT is very small, typically with less than 1 mHartree difference between the vtz and v5z basis sets. This is because the basis in DFT only has to express the occupied Kohn-Sham orbitals accurately, and is not used to describe electron correlation. Strictly speaking, the DFT energy is only comparable to the many-body solution in the complete basis set, because the functionals are designed to approximate the correlation energy in the basis set limit.

E. DMRG

The density matrix renormalization group (DMRG) [22] provides a variational ansatz for the wavefunction of the matrix product state form,

$$|\Psi\rangle = \sum_{n_1 n_2 \ldots n_k} [\mathbf{A}^{n_1}\mathbf{A}^{n_2}\ldots\mathbf{A}^{n_k}]_{11}|n_1 n_2 \ldots n_k\rangle \tag{2}$$

where $\mathbf{A}^n$ is a matrix of variational parameters for each orbital and $|n_1 n_2 \ldots n_k\rangle$ is an occupancy vector. The above ansatz expresses the coefficient of any occupancy vector as a product of matrices, where the "bond" dimension of the matrix $\mathbf{A}^n$ is $M \times M$. M may be increased until the ansatz is exact, which happens for M approximately the square root of the full Hilbert space size. The cost of the calculation using the quantum chemistry Hamiltonian with quartic interactions is proportional to $M^3k^3 + M^2k^4$ where k is the number of orbitals [23–26]. In a localized basis, the M required for a given accuracy scales like $e^{V^{D/D+1}}$ where V is the volume of the system and D is the dimension [27]. Thus, when extending a system along one dimension, M is independent of system size, while when extending a system along all three dimensions, the computational scaling is $e^{V^{2/3}}$. In any dimension, this is therefore a savings over full configuration interaction, which scales like e^V.

As can be seen, the ansatz requires an ordering of the orbitals and also treats all orbitals on the same footing. The latter means that in practice the DMRG is often a good ansatz relative to many methods when there are active orbitals which needed to be treated in a balanced way. However, it is inefficient when there are many doubly occupied or empty orbitals. The atoms and molecules in this system fall into this latter single-reference category. Thus we do not expect the DMRG to be especially efficient, but it serves as a near-numerically exact method to benchmark other techniques more suitable for these systems.

The DMRG calculations in this work were carried out using a spin-adapted code (a slight modification of the above ansatz) which allows us to obtain pure spin states [28]. The orbital ordering was generated by the default genetic algorithm [29]. We used the two-site variant of the DMRG and carried out calculations systematically increasing M. The largest M we used ranged from 4000 - 10000. To verify the accuracy of the energy we carried out an extrapolation in the total energy. We did this either by the standard linear extrapolation in the energy against the discarded weight in the two-site algorithm, where the DMRG energies at different M were computed by sweeping backwards from the largest M down to smaller M's (backwards schedule) [24, 29, 30], or by extrapolating the energies of the largest M values against $1/M$. In the first case, the extrapolation error is usually reported as a fraction of the extrapolation distance between the lowest variational energy and the extrapolated energy. This extrapolated energy was consistent with the energy obtained by extrapolating against $1/M$, but in some cases the $1/M$ extrapolation was more linear, and we report that as the extrapolated energy. The DMRG energies at the largest M are variational. Where included, they provide the lowest variational energies for this benchmark.

F. FCIQMC

The FCIQMC method [31–33] directly samples the many body wavefunction by stochastic propagation of a population of discrete walkers in Slater determinant space, defined by a given single particle basis. The annihilation of walkers with anti-walkers is crucial to ensure that the average wavefunction has the correct configurational sign-structure, and can lead to a circumvention of the Fermion sign problem without uncontrolled approximations. When a determinant is occupied by a small number of walkers, it is not clear whether the determinant should be ultimately dominated by walkers or anti-walkers, and so spawning new walkers from such a determinant can cause incorrect sign information to propagate throughout the network. This problem is minimized by the systematically improvable approach of Initiator FCIQMC [34, 35], which only allows the creation of new walkers on currently unoccupied determinants, by parent walkers residing on a determinant with a population above some threshold (in this work taken to be 3 walkers). This increases the incidence of annihilation events and encourages the sign-coherent propagation of walkers, and results in a convergence to the exact wavefunction energy and properties as the number of walkers increases. Furthermore, small subspaces are identified to define a 'trial wavefunction' onto which the sampled wavefunction is projected to calculate the energy, as well as another small subspace in which exact propagation can occur (the semi-stochastic adaptation [36, 37]). These subspaces serve to minimize the stochastic errorbars of the estimators.

All FCIQMC calculations in this work were undertaken via the following process:

- A maximum walker population N_{walker} is chosen
- The walker population is initialized on a single determinant and then allowed to reach N_{walker}
- A short interval in imaginary time after the maximum population is reached, the trial wavefunction space is initialized by performing an exact deterministic diagonalization of the Hamiltonian in the subspace spanned by the N_{TWF} most populated determinants (N_{TWF} ~200 determinants in this work)
- At the same iteration, a subspace of the N_{SS} most populated determinants is identified and designated as the semi-stochastic space. Thereafter, the walkers residing in this subspace are exactly propagated. (N_{SS} ~10,000 determinants in this work)
- The walker population is left to evolve under initiator FCIQMC dynamics until the numerator and denominator of the trial wavefunction projected energy stabilize around a mean value
- The energy is taken to be the ratio of means of the numerator and denominator of the energy
- The stochastic error is estimated using the Flyvberg – Peterson analysis[38] for serially correlated data.

The stochastically sampled wavefunction is affected by the systematic error introduced by the initiator criterion for spawning. In order to reduce this error to within acceptable bounds, the above procedure is repeated for increasingly large values of N_{walker} until convergence of the energy estimate with respect to this parameter is achieved. This was approximately 15 million, 50 million, 100 million and 200 million walkers for the Tm/Tm^+ vdz, vtz, vqz and TmO vdz calculations respectively[39]. This computational effort very roughly corresponds to 100 core hours per million walkers, with a maximum of ~15,000 CPU hours used in order to converge to ScO (vdz) system to small random error bars (200 million walkers).

G. Fixed node diffusion Monte Carlo

Fixed node diffusion Monte Carlo was used as implemented in the QWalk package.[40] A single determinant was generated using PySCF density functional theory in the B3LYP approximation. This determinant gave the lowest upper bound energy to the ground state. We multiplied the determinant by a 3-body Jastrow factor, which then was optimized to minimize the total energy using the linear method.[41, 42] The resultant single determinant Slater-Jastrow wave function was used as a guiding function for diffusion Monte Carlo.[43] In this method, the diffusion Monte Carlo wave function is given by

$$|\Psi_{DMC}\rangle = e^{-\tau \hat{H}} |\Psi_T\rangle, \tag{3}$$

where $|\Psi_T\rangle$ is the trial wave function, in this case the Slater-Jastrow wave function. The T-moves scheme[44] was used to ensure an upper bound to the ground state energy. Timestep errors were extrapolated out using a linear fitting process and timesteps as low as 0.0025 Hartrees^{-1}.

Scaling in stochastic methods is complex. To obtain the total energy with a given stochastic uncertainty, DMC(SD) scales as $\mathcal{O}(N_e^3 + \epsilon N_e^4)$. The N_e^4 cost typically does not appear until the number of electrons is more than 400-500. This method has been applied on systems with more than 1000 electrons.

We also considered more accurate trial wave functions, which can improve the fixed node error. We constructed them from SHCI wave functions by running small selected CI, and choosing the determinants with the largest weights. These weights were then reoptimized in the presence of the Jastrow factor using the linear method.[41, 42]

H. GF2

The fully self-consistent second order Green's function theory (GF2) [45–49] includes all second-order skeleton diagrams dressed with the renormalized second-order propagators and bare interactions. GF2 is formulated as a low-order approximation to the exact Luttinger-Ward (LW) functional [50] and therefore is Φ-derivable, thermodynamically consistent, and conserving [51, 52].

For transition atoms, we solve all the non-linear equations self-consistently at non-zero temperature. At each iteration, the self-energy, Green's function, and Fock matrix are updated until convergence is reached, so that the converged solution is reference-independent. Because of stability problems in the self-consistency, the calculations for transition monoxides are only done with one-shot GF2 on top of unrestricted Hartree-Fock. The self-energy, Green's function, and Fock matrix are not iterated until self-consistency. All the calculations are done on a mesh that combines a sparse power-law grid with an explicit transform based on a Legendre expansion of the self-energy. [53]

I. MRLCC

The multi-reference linearized coupled-cluster (MRLCC) is a flavor of multi-reference perturbation theory. We consider a reference wavefunction $|\Psi_0\rangle$ obtained for example from a CAS-like calculation. The out-of-active-space dynamical correlation may be added by multi-reference perturbation theory, where the expressions of the four first contributions to the energy are

$$E_0 = \langle \Psi_0 | \hat{H}_0 | \Psi_0 \rangle \tag{4}$$

$$E_1 = \langle \Psi_0 | \hat{V} | \Psi_0 \rangle \tag{5}$$

$$E_2 = \langle \Psi_0 | \hat{V} | \Psi_1 \rangle \tag{6}$$

$$E_3 = \langle \Psi_1 | \hat{V} - E_1 | \Psi_1 \rangle \tag{7}$$

and where the first order correction to the wavefunction $|\Psi_1\rangle$ obeys

$$(E_0 - \hat{H}_0)|\Psi_1\rangle = \hat{V}|\Psi_0\rangle. \tag{8}$$

In Rayleigh-Schrödinger perturbation theory, the partitioning of the Hamiltonian $\hat{H} = \hat{H}_0 + \hat{V}$ is so that $|\Psi_0\rangle$ and E_0 are the eigenvector and eigenvalue of the zeroth order Hamiltonian $\hat{H}_0$. As is well known, this leaves some flexibility for the choice of $\hat{H}_0$, leading to different perturbation theories having different properties: the use of the

Fock operator yields CASPT[54], and the use of the Dyall Hamiltonian[55] yields NEVPT[56, 57]. In this work, we show derivations and results for the MRLCC perturbation theory[58–61], which uses the Fink Hamiltonian[62, 63]:

$$\hat{H}_0 = \Big(\sum_{mn} t^n_m \hat{E}^n_m + \sum_{mnop} v^{op}_{mn} \hat{E}^{op}_{mn} \Big)_{\Delta=0}, \tag{9}$$

where the spin-free excitation operators are written with a hat, t and v are tensors, and the m, n, o, p indices refer to any molecular orbital. The notation $\Delta = 0$ indicates that only terms that do not change the number of electrons in the core, active and virtual spaces are taken into $\hat{H}_0$. It follows that:

$$\hat{V} = \Big(\sum_{mn} t^n_m \hat{E}^n_m + \sum_{mnop} v^{op}_{mn} \hat{E}^{op}_{mn} \Big)_{\Delta\neq0}, \tag{10}$$

and one can readily see from Eqs. (4) and (5) that the zeroth order energy is the energy of the reference wavefunction and that $E_1 = 0$, which goes to say that this zeroth order Hamiltonian is somewhat close to the exact $\hat{H}$ and is good in the context of perturbation theory (this is also the case for NEVPT).

In the internally contracted scheme, the first order correction to the wavefunction is expressed as a sum over eight class contributions $|\Psi^c_1\rangle$ expanded on perturber wavefunctions that are connected to the reference wavefunction:

$$|\Psi^c_1\rangle = \sum_I d^c_I \ \hat{E}^c_I |\Psi_0\rangle, \tag{11}$$

with coefficients $\mathbf{d}^c$. In this representation the Fink Hamiltonian is block diagonal (and so is the Dyall Hamiltonian), and the coefficients $\mathbf{d}^c$ in Eq. (11) are found by solving Eq. (8) subsequently for each class. Projection of Eq. (8) onto the basis of a class yields:

$$\mathbf{A}^c \mathbf{d}^c = \mathbf{S}^c \mathbf{w}^c, \tag{12}$$

with

$$A^c_{IJ} = \langle \Psi_0 | \hat{E}^{c\dagger}_I \left[(E_0 - \hat{H}_0), \hat{E}^c_J \right] | \Psi_0 \rangle \tag{13}$$

$$S^c_{IJ} = \langle \Psi_0 | \hat{E}^{c\dagger}_I \hat{E}^c_J | \Psi_0 \rangle \tag{14}$$

where realizing that $(E_0 - \hat{H}_0)|\Psi_0\rangle = 0$ allows the introduction of the commutator (see for example Ref. 61).

The terms to manipulate to compute $\mathbf{A}$, $\mathbf{S}$ and E_2, E_3 involve long strings of creation/annihilation operators, and one can use the Wick's theorem to simplify the expressions. This will result in series of tensor contractions involving one and two electron integrals (stemming from $\hat{H}_0$ and $\hat{V}$), and RDMs up to fourth order. The application of the Wick's theorem to the strings of operators is done with the "Second Quantization Algebra" symbolic algebra Python library [**?**], which we modified to fit our needs. The scripts also automate the generation of the C code used to solve the resulting equations and to calculate the energies E_2 and E_3.

J. QSGW

Typically RPA total energies are done from a simple one-body reference Hamiltonian H_0 such as PBE, HSE06, or Hartree-Fock. Since different choices yield different results, there are significant and unavoidable ambiguities. First of all, there is a fundamental issue: When employing a non-self-consistent Green's function, the different formulas for the total energy do differ in practice: the Galitskii-Migal and the RPA total energies are not equal any more [64].

The starting-point dependence can be surmounted by iterating G to self-consistency, that is, by finding a G generated by GW that is the same as the G that generates it ($G^{\rm out}$=$G^{\rm in}$). But it has long been known that full self-consistency can be quite poor in solids [65, 66]. A recent re-examination of some semiconductors [67] confirms that the dielectric function (and concomitant QP levels) indeed worsen when G is self-consistent, for reasons explained in Appendix A in Ref. [68]. Fully scGW becomes more problematic in transition metals [69]. Finally, scGW is a conserving approximation in the Green's function G, but W loses its usual physical meaning as a response function.

An alternative is the Quasiparticle Self-Consistent GW approximation[70] (QSGW). It is similar to scGW, but at each cycle the dynamical self-energy is rendered static and hermitian, forming a new noninteracting G_0 by making the substitution

$$V^{\rm xc} = \frac{1}{2} \sum_{ij} |\psi_i\rangle \left\{ {\rm Re}[\Sigma(\varepsilon_i)]_{ij} + {\rm Re}[\Sigma(\varepsilon_j)]_{ij} \right\} \langle \psi_j |, \tag{15}$$

where Re{} stands for the Hermitian part of the operator. This process is carried through to self-consistency. It was formally justified [68, 70] as a construction to optimize the noninteracting Green's function G_0 by minimizing some measure of the difference $|G - G_0|$. More recently it has been justified as minimizing the the gradient of the Baym-Kadanoff functional in the space of all G_0 [71].

QSGW is nevertheless an approximate self-consistent procedure, which relies on effective one-electron wavefunctions instead of the full Green's function. As a consequence, the difference between the expressions for the total energy (Galitskii-Migdal or RPA) still persists once the self-consistency has been reached. In this work, we defined the QSGW total energy as the one obtained from the RPA expression based on QSGW eigenvalues and wavefunctions, consistently with Ref. 72. Indeed, the RPA total energy, through the adiabatic connection, is capable of incorporating the correlated part of the kinetic energy [73], whereas the Galitskii-Migdal is not. The correlated part of the kinetic energy is sizeable and it is important to include it properly.

It is well known that GW overestimates the correlation energy. A prior study of weakly correlated molecular dimers[74] showed that (1) the RPA tends to systematically overestimate the correlation energy, and (2) the error is connected with short-ranged correlations. This tendency is also found here, as noted in the main text. The ionization energy and the dimer formation energy, both of which benefit from partial cancellation of errors in short-range correlation, are much better described. We also find that the RPA total energy based on QSGW, with its optimal choice for G_0, perform significantly better than RPA based on other G_0, e.g. PBE or Hartree Fock, as will be shown elsewhere.

Finally, it is has been established, using less optimal forms for G_0, that low-order diagrammatic corrections (especially second order screened exchange) significantly reduce errors in the correlation energy [75, 76]. It is shown elsewhere [77] how ladders dramatically improve the dielectric function in TM oxide crystals, so it is reasonable to expect that correlation energies computed from it will see a similar improvement. Density-functional approximations for the exchange-correlation kernel significantly improve on heats of formation of dimers from *sp* elements [78].

K. RPA

We calculated the RPA total energy using the Tamm-Dancoff approximation (specifically Eq. (9) from Ref. [79]), using several flavors of G_0: PBE, Hartree-Fock, HSE06, and QSGW, for the TM atom and M+O dimer. All the RPA calculations in this paper use the MOLGW code [80].

L. sc-GW

The GW method [81] evaluates a subset of terms of a diagrammatic weak coupling series in the interaction V deterministically. The GW approximation can be understood as a first-order approximation to Hedin's series of renormalized propagators and interactions. It is expressed in terms of self-energies Σ, Green's functions G, screened interactions W and polarizations P by the self-consistent solution of the equations $G = G_0 + G_0\Sigma G$, $W = V + VPW$, where G_0 denotes the Hartree-Fock Green's function, and Σ and P are computed as $\Sigma = -GW$ and $P = GG$. The main difference to GF2 is that, in GW, both propagators and interactions are renormalized, whereas GF2 only renormalizes the propagators. However, GF2 obtains all second order contributions, whereas the second order exchange is missing from GW. Our results are converged to self-consistency, using a finite-temperature imaginary time formulation evaluated at temperatures low enough that the system is in its ground state. Sparse imaginary time and Matsubara frequency grids based on Chebyshev polynomials and the intermediate representation (IR) [82–85] is used, which significantly reduced the computational cost. Our code is based on the ALPS libraries [86, 87].

Self-consistent GW is a $\Phi-$ [51, 52] and Ψ-derivable [88] weak coupling method, in the sense that it neglects some diagrams of order V^2. Achieving full self-consistency requires the storage and manipulation of W, which is a frequency-dependent four-index tensor. The necessity of handling this object numerically restricts the method in our implementation to relatively small system sizes.

M. Self-energy embedding theory (SEET)

SEET [89–96] is a finite temperature Green's function embedding method. The embedding construction allows us to describe the weakly and strongly correlated orbitals at different levels of theory. The weakly correlated orbitals are

treated by a low level, most often a perturbative method (here Green's function second order (GF2) [46–48, 97–99] or a single iteration of GF2). The strongly correlated orbitals are treated with a high level, usually non-perturbative method. When multiple strongly correlated orbitals are present, they are separated into several intersecting or non intersecting subsets A_i. Each of these subsets contains M_i^A orbitals and $M = \sum_i M_i^A + M^R$, where M is the total number of orbitals in the problem and M^R are all the orbitals that are not contained in the groups of the strongly correlated orbitals. The orbitals from each of the subsets A_i are used to construct Anderson Impurity Models (AIM) that are then solved by a non-perturbative method, here full configuration interaction (FCI) [100–102]. The intersubset interactions are treated most commonly at a perturbative level. Orbitals chosen to each of the A_i subsets can be chosen based on several criteria such as occupancies of natural orbitals (NOs) or energies of molecular orbitals (MOs), for details see Refs. [48,50].

A general SEET functional can be written as

$$\begin{aligned}\Phi_{\mathrm{MIX}}^{\mathrm{SEET}} &= \Phi_{\mathrm{weak}}^{\mathrm{tot}} + \textstyle\sum_i^{\binom{n}{k}} (\Phi_{\mathrm{strong}}^{\mathrm{A_i^k}} - \Phi_{\mathrm{weak}}^{\mathrm{A_i^k}}) \\ &\pm \textstyle\sum_{k=K-1}^{k=1} \sum_i^{\binom{n}{k}} (\Phi_{\mathrm{strong}}^{\mathrm{B_i^k}} - \Phi_{\mathrm{weak}}^{\mathrm{B_i^k}}),\end{aligned} \tag{16}$$

where $\Phi_{\mathrm{weak}}^{\mathrm{tot}}$, in this work is a GF2 solution for the whole orbital space, $\Phi_{\mathrm{strong}}^{\mathrm{A_i^k}}$ is obtained from the solution of AIM for the strongly correlated subset of orbitals A_i, $\Phi_{\mathrm{weak}}^{\mathrm{A_i^k}}$ is the solution of the subset A_i with a weakly correlated method used to remove the double counting. The terms $(\Phi_{\mathrm{strong}}^{\mathrm{B_i^k}} - \Phi_{\mathrm{weak}}^{\mathrm{B_i^k}})$ are present in case of intersecting subsets A_i and are necessary to remove the double counting, for details see Ref. 94. We denote a particular SEET calculation as SEET(method strong/method weak)-m([$\mathrm{M^A}$o]/basis) since self-energies from intersecting orbital subspaces containing M^A orbitals are treated with "method strong". The whole system is treated with "method weak" and the orbitals from subsets A_i are transformed to a certain orbital basis denoted here as "basis". In this paper, we most commonly use the basis of molecular orbitals. The details of the finite temperature imaginary time GF2 grid as well as the frequency grid can be found in Ref. 53, 103, and 104.

N. SHCI

The semistochastic heat-bath configuration iteration (SHCI) method [105–108], is an efficient instance of the general class of methods wherein a selected configuration interaction is performed followed by a perturbative correction (SCI+PT). SCI+PT methods have two stages. In the first stage a set of "important" determinants are selected the Hamiltonian is diagonalized in the subspace of these determinants, $\mathcal{V}$, to obtain the the lowest few eigenstates (or the lowest state if one is interested in the ground state only). In the second stage, a second-order perturbation theory is used to calculate the energy contributions of all determinants that do not belong to the space $\mathcal{V}$ but have a non-zero Hamiltonian matrix element with at least one of the determinants in $\mathcal{V}$. Such methods have been used for about 50 years [109–111] and continue to be a subject of interest [112–120] to the present day.

We briefly describe the two innovations that account for the time- and memory-efficiency of SHCI. A more detailed description can be found in 108.

1. During the variational and the perturbative steps, straightforward implementations of SCI+PT scan all determinants connected to at least one of the determinants in $\mathcal{V}$ and select those determinants ($|D_a\rangle$) for which the absolute value of the 2^{nd}-order perturbative contribution to the energy

$$\left| \frac{\left(\sum_{D_i \in \mathcal{V}} H_{ai} c_i\right)^2}{E_V - E_a} \right| > \epsilon, \tag{17}$$

where the subscript a denotes a determinant not currently present in $\mathcal{V}$, E_V is the energy of the current variational wavefunction, E_a is the energy of determinant D_a, and ϵ is a parameter that controls the number of determinants selected. During the variational stage, this is done iteratively to build up the variational wavefunction starting from a single determinant. During the perturbative stage, $\epsilon = 0$. Instead, SHCI modifies the selection criterion to [105],

$$\max_{D_i \in \mathcal{V}} |H_{ai} c_i| > \epsilon, \tag{18}$$

which greatly reduces the cost by taking advantage of the fact that most of the matrix elements, H_{ai}, are 2-body excitations, which depend only on the indices of the 4 orbitals whose occupations change and not on the other

occupied orbitals of a determinant. Thus by presorting the absolute values of all possible matrix elements of the 2-body excitations in descending order, the scan over determinants D_a can be terminated when $|H_{ai}|$ drops below ϵ/c_i. A similar idea is used to speed up the selection of 1-body excitations as well. This enables a procedure in which *only the important determinants which will be included in the variational wavefunction, or make significant contributions to the perturbative correction, are ever looked at*, resulting in orders of magnitude saving over a naive implementation of the SCI+PT algorithm! Different values of ϵ are used during the variational and the perturbative stages of the calculation, which we denote by ϵ_1 and ϵ_2.

2. The first innovation greatly speeds up both the variational and the perturbative steps of the algorithm. However, the perturbative step has a very large memory requirement when $\mathcal{V}$ has a large number of determinants (say 10^9) because all determinants that are connected to those in $\mathcal{V}$ must be stored. [121] We have developed a 2-step [106] and later a 3-step [108] semistochastic perturbative approach that both completely overcomes this memory bottleneck and is faster than the deterministic approach.

In addition to these two major methodological improvements, the SHCI method uses auxiliary arrays to speed up the computation of the Hamiltonian matrix [108]. Further, it makes extensive use of hashing and techniques such as variable-byte encoding, hardware atomic operations, dynamic load-balancing and thread pooling to achieve a high efficiency in the use of computer time and memory.

The convergence of the variational and perturbative energies depends significantly on the orbitals used. The convergence obtained from using Hartree-Fock orbitals can be improved by using natural orbitals obtained from an SHCI calculation with a fairly large value of ϵ_1, and can be further improved by using orbitals that minimize the SHCI variational energy using a modified version of the algorithm described in Ref. 122.

We typically choose $\epsilon_2 = 10^{-6}\epsilon_1$, so that a single parameter, ϵ_1 controls the accuracy of the calculation. The energy at the $\epsilon_1 = 0$ limit is obtained using a quadratic fit to the energies versus the perturbative correction [107]. Note that although the SHCI algorithm has a perturbative component, systematically improvable approximations to the exact energy in the chosen basis are obtained by performing calculations with progressively smaller values of ϵ_1 until the total energy (variational energy plus perturbative correction), or its extrapolation versus the perturbative correction, is converged to the desired tolerance.

II. BASIS SETS, BOND LENGTHS, AND EFFECTIVE CORE POTENTIALS

All calculations used the effective core potentials and associated aug-ccpVnZ gaussian basis sets of Trail and Needs[123–125]. These effective core potentials are produced from explicitly correlated multi-configuration Hartree-Fock calculations, and include contributions from core-core and core-valence correlations. Augmented double-, triple-, quadruple-, and quintuple-zeta basis sets were used.

The molecules were computed with bond lengths in Å as follows: ScO: 1.668, TiO: 1.623,VO: 1.591, CrO: 1.621, MnO: 1.648, FeO: 1.616, CuO: 1.725.

III. DATA

Each contributing author provided a separate data file detailing the results of their calculations for both atomic and molecular systems. Each row of each data file indicates the system considered, the method and the basis set used, and the resulting total energies, along with any associated systematic or stochastic error. Additional method-specific information is also included (for example, the size of the active orbital space used in auxiliary-field quantum Monte Carlo). Not all methods completed each calculation using each level of basis set quality, due to the high computational expense of some methods.

The CSV headers required are:

- charge: charge of the molecule/atom (either 0 or 1)
- molecule: (for molecules) the name of the molecule: example “VO”
- atom: (for atoms) the name of the atom: example “V”
- pseudopotential: always “trail” for this test set

- pyscf-version: "new" for after pyscf 1.5. A small improvement in accuracy was implemented after this version. Should be "new" for all calculations
- method: a string representing the method used for the calculation. example "PBE"
- totalenergy: total energy in Hartrees
- totalenergy-stocherr : Stochastic error estimate
- totalenergy-syserr : Systematic error estimate, if available

A script called 'gather.py' retrieves all data from all directories. Further scripts handle plotting and error estimation.

IV. ENERGY COMPARISON FOR SEVERAL ACCURATE METHODS

There are 5 methods (DMRG, iFCIQMC, UCCSD(T), AFQMC(MD) and SEET(FCI/GF2)) for which the total energies have an rms deviation of 4 mHa or less relative to the SHCI reference. The maximum absolute error, the rms error and the number of systems treated are shown in Table IV for these methods. Tables IV and IV show the corresponding quantities for the ionization energy and the dissociation energy.

TABLE II. Total energy errors relative to SHCI of the five methods that agree best with SHCI.

Method	# systems	max abs error	rms error
DMRG	39	0.000590	0.000181
iFCIQMC	49	0.001611	0.000639
UCCSD(T)	92	0.006811	0.002309
AFQMC(MD)	92	0.007470	0.003540
SEET(FCI/GF2)	59	0.013656	0.004001

TABLE III. Ionization energy errors relative to SHCI of the five methods that agree best with SHCI.

Method	# systems	max abs error	rms error
DMRG	14	0.0.000512	0.000232
iFCIQMC	21	0.000965	0.000567
UCCSD(T)	28	0.001222	0.000675
AFQMC(MD)	28	0.008400	0.002888
SEET(FCI/GF2)	18	0.006580	0.002466

TABLE IV. Dissociation energy errors relative to SHCI of the five methods that agree best with SHCI.

Method	# systems	max abs error	rms error
DMRG	7	0.000678	0.000327
iFCIQMC	1	0.000488	0.000488
UCCSD(T)	28	0.005108	0.002990
AFQMC(MD)	28	0.007880	0.002590
SEET(FCI/GF2)	14	0.008356	0.004152

V. BASIS SET EXTRAPOLATION

Methods not operating directly in the complete basis set limit must be extrapolated using finite basis calculations. The extrapolated energy $E_m(\mathrm{CBS})$ for method m is estimated from

$$E_m(\mathrm{CBS}) = E_{\mathrm{HF}}(\mathrm{CBS}) + \Delta(\mathrm{CBS}). \tag{19}$$

where $E_{\mathrm{HF}}(\mathrm{CBS})$ is obtained by fitting the HF energies to the form

$$E_{\mathrm{HF}}(n) = E_{\mathrm{HF}}(\mathrm{CBS}) + b\exp(-cn). \tag{20}$$

and $\Delta(\mathrm{CBS})$ is obtained by fitting the correlation energies to

$$E_m(n) - E_{\mathrm{HF}}(n) = \Delta(\mathrm{CBS}) + \frac{\gamma}{n^3}, \tag{21}$$

where n is the cardinal index of the basis. For density functional theory, we extrapolated using an exponential *ansatz* as in Eqn 20. The largest difference between quadruple and quintuple zeta energies is less than 0.4 mHa.

An example extrapolation for all materials and SHCI correlation energies is shown in Fig 1. Extrapolations for the total energy, ionization energy and dissociation energy are shown in Tables V, VI and VII respectively. The ionization and dissociation energies converge much more rapidly than the total energies. Of the 4 basis set extrapolations considered, cbs45 is the most accurate. Table VIII shows the rms deviation of the total, ionization and binding energies of each basis set and extrapolation relative to the corresponding cbs45 energies.

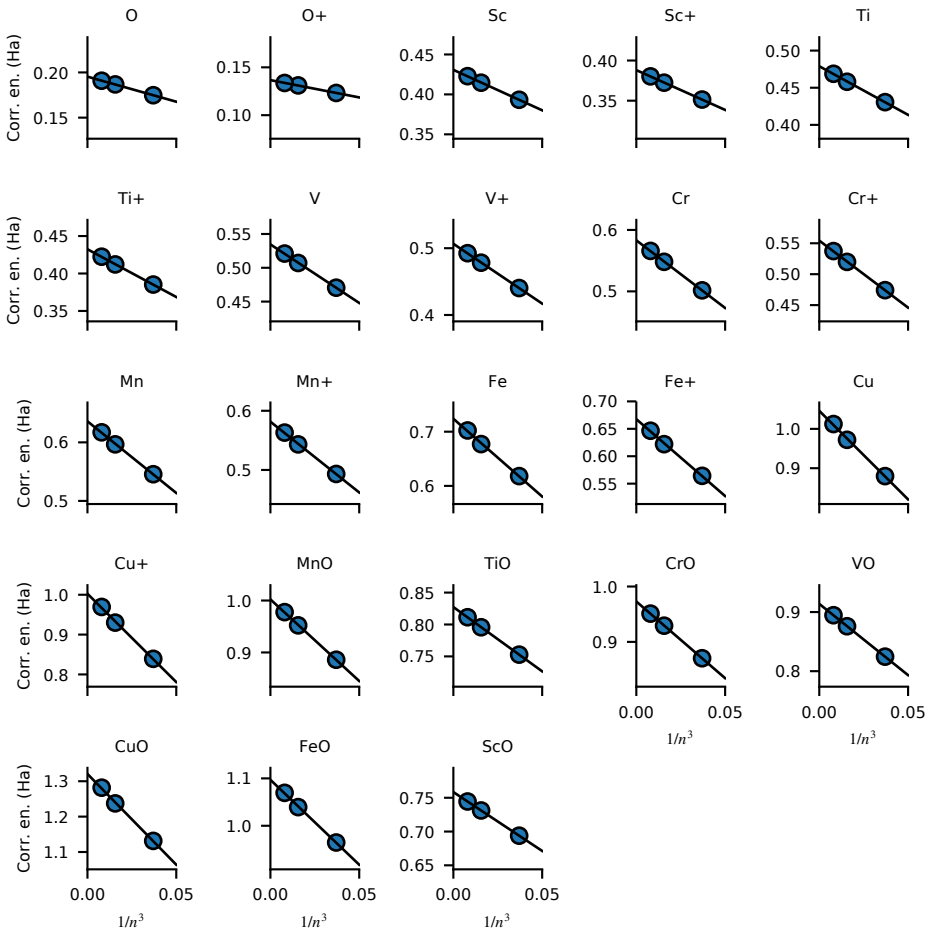

FIG. 1. Basis set extrapolation of the SHCI correlation energy versus the n in the basis vnz. Only n values from 3-5 are shown since the $n = 2$ points deviate significantly from the straight lines shown.

TABLE V. SHCI total energies for all basis sets and extrapolations. cbs23 indicates an extrapolation using only vdz and vtz bases, and cbs34/cbs45/cbs345 are labeled in the same way. The differences between the cbs345 and the cbs45 extrapolations increase with the atomic number of the transition metal and range from 1-8 mHa. The cbs45 extrapolation is the most accurate one.

	Total Energy (Ha)							
basis	O	Sc	Ti	V	Cr	Mn	Fe	Cu
vdz	-15.781250	-46.396850	-57.883050	-71.077670	-86.612030	-103.942260	-123.518120	-197.236900
vtz	-15.827020	-46.452300	-57.954630	-71.168980	-86.720920	-104.063660	-123.657200	-197.444640
vqz	-15.839090	-46.474000	-57.983000	-71.205780	-86.767610	-104.115010	-123.716750	-197.536420
v5z	-15.843170	-46.482130	-57.993860	-71.219830	-86.785320	-104.135420	-123.741800	-197.576080
cbs23	-15.846068	-46.475646	-57.985016	-71.207033	-86.766697	-104.114189	-123.715147	-197.531655
cbs34	-15.847828	-46.489701	-58.003293	-71.232489	-86.801544	-104.152337	-123.760024	-197.603262
cbs345	-15.847658	-46.490126	-58.004201	-71.233375	-86.802563	-104.154264	-123.763441	-197.609436
cbs45	-15.847432	-46.490691	-58.005411	-71.234555	-86.803920	-104.156833	-123.767993	-197.617661

	Total Energy (Ha)							
basis	O+	Sc+	Ti+	V+	Cr+	Mn+	Fe+	Cu+
vdz	-15.300670	-46.156830	-57.634020	-70.828400	-86.372410	-103.673360	-123.233540	-196.967230
vtz	-15.333580	-46.211490	-57.704240	-70.919550	-86.472980	-103.792140	-123.368570	-197.163040
vqz	-15.342100	-46.233110	-57.732560	-70.957590	-86.518670	-103.842720	-123.427090	-197.253430
v5z	-15.344790	-46.241150	-57.743160	-70.971820	-86.536270	-103.862840	-123.451830	-197.292510
cbs23	-15.348503	-46.234171	-57.733836	-70.957555	-86.515427	-103.840708	-123.424132	-197.245135
cbs34	-15.348554	-46.248705	-57.752710	-70.985120	-86.552131	-103.879380	-123.469470	-197.319302
cbs345	-15.348317	-46.248980	-57.753325	-70.985640	-86.553226	-103.881194	-123.472883	-197.325395
cbs45	-15.348001	-46.249346	-57.754146	-70.986333	-86.554685	-103.883611	-123.477431	-197.333514

	Total Energy (Ha)						
basis	ScO	TiO	VO	CrO	MnO	FeO	CuO
vdz	-62.420040	-73.903750	-87.085770	-102.558370	-119.850510	-139.435990	-213.123030
vtz	-62.530130	-74.030850	-87.235260	-102.718670	-120.029280	-139.635460	-213.377780
vqz	-62.568454	-74.075488	-87.288682	-102.779756	-120.097071	-139.711352	-213.484244
v5z	-62.582141	-74.091473	-87.307394	-102.802737	-120.122364	-139.741682	-213.529128
cbs23	-62.576098	-74.084170	-87.297330	-102.786429	-120.103370	-139.717990	-213.483952
cbs34	-62.595849	-74.107270	-87.326938	-102.823579	-120.145888	-139.766095	-213.561561
cbs345	-62.596114	-74.107694	-87.326918	-102.824937	-120.147133	-139.769181	-213.567787
cbs45	-62.596466	-74.108258	-87.326891	-102.826745	-120.148791	-139.773293	-213.576082

TABLE VI. SHCI ionization energies. cbs23 indicates an extrapolation using only vdz and vtz bases, and cbs34/cbs45/cbs345 are labeled in the same way. Experimental values are also shown.

	Ionization Potential (Ha)						
basis	Sc	Ti	V	Cr	Mn	Fe	Cu
vdz	0.240020	0.249030	0.249270	0.239620	0.268900	0.284580	0.269670
vtz	0.240810	0.250390	0.249430	0.247940	0.271520	0.288630	0.281600
vqz	0.240890	0.250440	0.248190	0.248940	0.272290	0.289660	0.282990
v5z	0.240980	0.250700	0.248010	0.249050	0.272580	0.289970	0.283570
cbs23	0.241474	0.251180	0.249479	0.251269	0.273481	0.291015	0.286520
cbs34	0.240996	0.250583	0.247369	0.249413	0.272957	0.290555	0.283961
cbs345	0.241146	0.250876	0.247735	0.249337	0.273071	0.290558	0.284041
cbs45	0.241346	0.251265	0.248222	0.249235	0.273222	0.290562	0.284147
exper	0.24113	0.25093	0.24792	0.24866	0.27320	0.29041	0.28394

TABLE VII. SHCI dissociation energies for the molecules considered in this work. cbs23 indicates an extrapolation using only vdz and vtz bases, and cbs34/cbs45/cbs345 are labeled in the same way.

	Dissociation Energy (Ha)						
basis	ScO	TiO	VO	CrO	MnO	FeO	CuO
vdz	0.241940	0.239450	0.226850	0.165090	0.127000	0.136620	0.104880
vtz	0.250810	0.249200	0.239260	0.170730	0.138600	0.151240	0.106120
vqz	0.255364	0.253398	0.243812	0.173056	0.142971	0.155512	0.108734
v5z	0.256841	0.254443	0.244394	0.174247	0.143774	0.156712	0.109878
cbs23	0.254384	0.253086	0.244229	0.173665	0.143113	0.156775	0.106230
cbs34	0.258321	0.256149	0.246622	0.174208	0.145724	0.158243	0.110471
cbs345	0.258330	0.255835	0.245885	0.174716	0.145210	0.158082	0.110693
cbs45	0.258343	0.255415	0.244903	0.175394	0.144526	0.157868	0.110989

TABLE VIII. RMS deviations of the SHCI total, ionization, and dissociation energies for various basis sets and extrapolations with respect to the cbs45 extrapolation.

	RMS Deviation (Ha)		
basis	Total energy	Ionization energy	Dissociation energy
vdz	0.169949	0.007215	0.015820
vtz	0.075498	0.001572	0.005997
vqz	0.034565	0.000756	0.002160
v5z	0.017665	0.000482	0.001063
cbs23	0.036296	0.001290	0.002683
cbs34	0.005110	0.000455	0.000981
cbs345	0.002919	0.000260	0.000561

[1] S. Zhang, J. Carlson, and J. E. Gubernatis, Phys. Rev. B **55**, 7464 (1997).
[2] S. Zhang and H. Krakauer, Phys. Rev. Lett. **90**, 136401 (2003).
[3] S. Zhang, *Auxiliary-Field Quantum Monte Carlo for Correlated Electron Systems*, Vol. 3 of *Emergent Phenomena in Correlated Matter: Modeling and Simulation*, Ed. E. Pavarini, E. Koch, and U. Schollwöck (Verlag des Forschungszentrum Jülich, 2013).
[4] W. A. Al-Saidi, S. Zhang, and H. Krakauer, The Journal of Chemical Physics **124**, 224101 (2006), https://doi.org/10.1063/1.2200885.
[5] M. Motta and S. Zhang, Wiley Interdisciplinary Reviews: Computational Molecular Science **8**, e1364 (2018), https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.1364.
[6] 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, A. J. Millis, N. V. Prokof'ev, U. Ray, G. E. Scuseria, S. Sorella, E. M. Stoudenmire, Q. Sun, I. S. Tupitsyn, S. R. White, D. Zgid, and S. Zhang (Simons Collaboration on the Many-Electron Problem), Phys. Rev. X **7**, 031059 (2017).
[7] W. Purwanto, S. Zhang, and H. Krakauer, The Journal of Chemical Physics **142**, 064302 (2015), https://doi.org/10.1063/1.4906829.
[8] H. Shi and S. Zhang, *To be published.*
[9] J. Shee, E. J. Arthur, S. Zhang, D. R. Reichman, and R. A. Friesner, Journal of Chemical Theory and Computation **14**, 4109 (2018), pMID: 29897748, https://doi.org/10.1021/acs.jctc.8b00342.
[10] W. A. Al-Saidi, S. Zhang, and H. Krakauer, The Journal of Chemical Physics **127**, 144101 (2007), https://doi.org/10.1063/1.2770707.
[11] W. Purwanto, S. Zhang, and H. Krakauer, The Journal of Chemical Physics **142**, 064302 (2015), https://doi.org/10.1063/1.4906829.
[12] W. Purwanto, S. Zhang, and H. Krakauer, The Journal of Chemical Physics **144**, 244306 (2016), https://doi.org/10.1063/1.4954245.
[13] M. Qin, H. Shi, and S. Zhang, Phys. Rev. B **94**, 085103 (2016).
[14] H. Shi and S. Zhang, Phys. Rev. B **88**, 125132 (2013).
[15] H. Shi, C. A. Jiménez-Hoyos, R. Rodríguez-Guzmán, G. E. Scuseria, and S. Zhang, Phys. Rev. B **89**, 125129 (2014).
[16] H. Shi, S. Chiesa, and S. Zhang, Phys. Rev. A **92**, 033603 (2015).
[17] H. Shi and S. Zhang, Phys. Rev. B **95**, 045144 (2017).
[18] M. Qin, H. Shi, and S. Zhang, Phys. Rev. B **94**, 235119 (2016).
[19] Y.-Y. He, M. Qin, H. Shi, Z.-Y. Lu, and S. Zhang, arXiv e-prints , arXiv:1811.07290 (2018), arXiv:1811.07290 [cond-mat.str-el].
[20] Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters, and G. K.-L. Chan, Wiley Interdisciplinary Reviews: Computational Molecular Science **8**, e1340.
[21] R. J. Bartlett and M. Musia, Reviews of Modern Physics **79**, 291 (2007).
[22] S. R. White, Physical review letters **69**, 2863 (1992).
[23] S. R. White and R. L. Martin, The Journal of chemical physics **110**, 4127 (1999).
[24] G. K.-L. Chan and M. Head-Gordon, J. Chem. Phys. **116**, 4462 (2002).
[25] S. Keller and M. Reiher, The Journal of chemical physics **144**, 134101 (2016).
[26] G. K.-L. Chan, A. Keselman, N. Nakatani, Z. Li, and S. R. White, The Journal of chemical physics **145**, 014102 (2016).
[27] J. Hachmann, W. Cardoen, and G. K.-L. Chan, The Journal of chemical physics **125**, 144101 (2006).
[28] S. Sharma and G. K.-L. Chan, The Journal of chemical physics **136**, 124121 (2012).
[29] R. Olivares-Amaya, W. Hu, N. Nakatani, S. Sharma, J. Yang, and G. K.-L. Chan, The Journal of chemical physics **142**, 034102 (2015).
[30] Ö. Legeza and G. Fáth, Physical Review B **53**, 14349 (1996).
[31] G. H. Booth, A. J. W. Thom, and A. Alavi, The Journal of Chemical Physics **131**, 054106 (2009), https://aip.scitation.org/doi/pdf/10.1063/1.3193710.
[32] G. H. Booth, A. Grüneis, G. Kresse, and A. Alavi, Nature **493**, 365 (2013).
[33] G. H. Booth, S. D. Smart, and A. Alavi, Molecular Physics **112**, 1855 (2014).
[34] D. Cleland, G. H. Booth, and A. Alavi, The Journal of Chemical Physics **132**, 041103 (2010), https://doi.org/10.1063/1.3302277.
[35] G. H. Booth, D. M. Cleland, A. J. W. Thom, and A. Alavi, J. Chem. Phys. **135**, 084104 (2011).
[36] N. S. Blunt, S. D. Smart, J. A. F. Kersten, J. S. Spencer, G. H. Booth, and A. Alavi, The Journal of Chemical Physics **142**, 184107 (2015), https://doi.org/10.1063/1.4920975.
[37] F. R. Petruzielo, A. A. Holmes, H. J. Changlani, M. P. Nightingale, and C. J. Umrigar, Phys. Rev. Lett. **109**, 230201 (2012).
[38] H. Flyvbjerg and H. G. Petersen, The Journal of Chemical Physics **91**, 461 (1989), https://doi.org/10.1063/1.457480.
[39] R. E. Thomas, G. H. Booth, and A. Alavi, Phys. Rev. Lett. **114**, 033001 (2015).
[40] L. K. Wagner, M. Bajdich, and L. Mitas, Journal of Computational Physics **228**, 3390 (2009).
[41] J. Toulouse and C. J. Umrigar, J. Chem. Phys. **126** (2007), 10.1063/1.2437215.
[42] J. Toulouse and C. J. Umrigar, The Journal of Chemical Physics **128**, 174101 (2008).

[43] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Reviews of Modern Physics **73**, 33 (2001).
[44] M. Casula, Physical Review B **74** (2006), 10.1103/PhysRevB.74.161102.
[45] J. J. Phillips and D. Zgid, The Journal of Chemical Physics **140**, 241101 (2014), https://doi.org/10.1063/1.4884951.
[46] A. A. Rusakov and D. Zgid, The Journal of Chemical Physics **144**, 054106 (2016), https://doi.org/10.1063/1.4940900.
[47] J. J. Phillips, A. A. Kananenka, and D. Zgid, J. Chem. Phys. **142**, 194108 (2015), http://dx.doi.org/10.1063/1.4921259.
[48] A. R. Welden, A. A. Rusakov, and D. Zgid, J. Chem. Phys. **145**, 204106 (2016).
[49] N. E. Dahlen and U. von Barth, The Journal of Chemical Physics **120**, 6826 (2004), https://doi.org/10.1063/1.1650307.
[50] J. M. Luttinger and J. C. Ward, Phys. Rev. **118**, 1417 (1960).
[51] G. Baym and L. P. Kadanoff, Phys. Rev. **124**, 287 (1961).
[52] G. Baym, Phys. Rev. **127**, 1391 (1962).
[53] A. A. Kananenka, A. R. Welden, T. N. Lan, E. Gull, and D. Zgid, Journal of Chemical Theory and Computation **12**, 2250 (2016), pMID: 27049642.
[54] J. Finley, P. ke Malmqvist, B. O. Roos, and L. Serrano-Andrs, Chemical Physics Letters **288**, 299 (1998).
[55] K. G. Dyall, The Journal of Chemical Physics **102**, 4909 (1995).
[56] C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, and J.-P. Malrieu, The Journal of Chemical Physics **114**, 10252 (2001).
[57] C. Angeli, R. Cimiraglia, and J.-P. Malrieu, The Journal of Chemical Physics **117**, 9138 (2002).
[58] S. Sharma and A. Alavi, The Journal of Chemical Physics **143**, 102815 (2015).
[59] S. Sharma, G. Jeanmairet, and A. Alavi, The Journal of Chemical Physics **144**, 034103 (2016).
[60] G. Jeanmairet, S. Sharma, and A. Alavi, The Journal of Chemical Physics **146**, 044107 (2017).
[61] S. Sharma, G. Knizia, S. Guo, and A. Alavi, Journal of Chemical Theory and Computation **13**, 488 (2017).
[62] R. F. Fink, Chemical Physics Letters **428**, 461 (2006).
[63] R. F. Fink, Chemical Physics **356**, 39 (2009).
[64] N. E. Dahlen, R. van Leeuwen, and U. von Barth, Phys. Rev. A **73**, 012511 (2006).
[65] E. L. Shirley, Phys. Rev. B **54**, 7758 (1996).
[66] B. Holm and U. von Barth, Phys. Rev. B **57**, 2108 (1998).
[67] M. Grumet, P. Liu, M. Kaltak, J. Klimeš, and G. Kresse, Phys. Rev. B **98**, 155143 (2018).
[68] T. Kotani, M. van Schilfgaarde, and S. V. Faleev, Phys. Rev. B **76**, 165106 (2007).
[69] K. D. Belashchenko, V. P. Antropov, and N. E. Zein, Phys. Rev. B **73**, 073105 (2006).
[70] S. V. Faleev, M. van Schilfgaarde, and T. Kotani, Phys. Rev. Lett. **93**, 126406 (2004).
[71] S. Ismail-Beigi, J. Phys.: Condens. Matter **29**, 385501 (2017).
[72] F. Bruneval, J. Chem. Phys. **136**, 194107 (2012).
[73] M. Fuchs, Y.-M. Niquet, X. Gonze, and K. Burke, The Journal of Chemical Physics **122**, 094116 (2005), https://doi.org/10.1063/1.1858371.
[74] T. Olsen and K. S. Thygesen, Phys. Rev. B **86**, 081103 (2012).
[75] X. Ren, N. Marom, F. Caruso, M. Scheffler, and P. Rinke, Phys. Rev. B **92**, 081104 (2015).
[76] E. Maggio and G. Kresse, Journal of Chemical Theory and Computation **13**, 4765 (2017), https://doi.org/10.1021/acs.jctc.7b00586.
[77] B. Cunningham, M. Grüning, P. Azarhoosh, D. Pashov, and M. van Schilfgaarde, Phys. Rev. Mater. **2**, 034603 (2018).
[78] P. Liu, B. Kim, X.-Q. Chen, D. D. Sarma, G. Kresse, and C. Franchini, Phys. Rev. Materials **2**, 075003 (2018).
[79] H. Eshuisa, J. Yarkony, and F. Furche, J. Chem. Phys. **132**, 234114 (2010).
[80] F. Bruneval, T. Rangel, S. Hamed, M. Shao, C. Yang, and J. Neaton, Comput. Phys. Commun. **208**, 149 (2016).
[81] L. Hedin, Phys. Rev. **139**, A796 (1965).
[82] E. Gull, S. Iskakov, I. Krivenko, A. A. Rusakov, and D. Zgid, Phys. Rev. B **98**, 075127 (2018).
[83] H. Shinaoka, J. Otsuki, M. Ohzeki, and K. Yoshimi, Physical Review B **96**, 035147 (2017).
[84] N. Chikano, K. Yoshimi, J. Otsuki, and H. Shinaoka, Computer Physics Communications **240**, 181 (2019).
[85] J. Li, M. Wallerberger, N. Chikano, C.-N. Yeh, E. Gull, and H. Shinaoka, arXiv preprint arXiv:1908.07575 (2019).
[86] A. Gaenko, A. Antipov, G. Carcassi, T. Chen, X. Chen, Q. Dong, L. Gamper, J. Gukelberger, R. Igarashi, S. Iskakov, M. Könz, J. LeBlanc, R. Levy, P. Ma, J. Paki, H. Shinaoka, S. Todo, M. Troyer, and E. Gull, Computer Physics Communications **213**, 235 (2017).
[87] M. Wallerberger, S. Iskakov, A. Gaenko, J. Kleinhenz, I. Krivenko, R. Levy, J. Li, H. Shinaoka, S. Todo, T. Chen, X. Chen, J. P. F. LeBlanc, J. E. Paki, H. Terletska, M. Troyer, and E. Gull, arXiv e-prints , arXiv:1811.08331 (2018), arXiv:1811.08331 [physics.comp-ph].
[88] C.-O. ALMBLADH, U. V. BARTH, and R. V. LEEUWEN, International Journal of Modern Physics B **13**, 535 (1999), https://doi.org/10.1142/S0217979299000436.
[89] D. Zgid and E. Gull, New J. Phys. **19**, 023047 (2017).
[90] A. A. Kananenka, E. Gull, and D. Zgid, Phys. Rev. B **91**, 121111 (2015).
[91] T. N. Lan, A. A. Kananenka, and D. Zgid, J. Chem. Phys. **143**, 241102 (2015), http://dx.doi.org/10.1063/1.4938562.
[92] T. Nguyen Lan, A. A. Kananenka, and D. Zgid, J. Chem. Theory Comput. **12**, 4856 (2016).
[93] T. N. Lan, A. Shee, J. Li, E. Gull, and D. Zgid, Phys. Rev. B **96**, 155106 (2017).
[94] T. N. Lan and D. Zgid, J. Phys. Chem. Lett. **8**, 2200 (2017), pMID: 28453934, http://dx.doi.org/10.1021/acs.jpclett.7b00689.
[95] L. N. Tran, S. Iskakov, and D. Zgid, J. Phys. Chem. Lett. **9**, 4444 (2018), pMID: 30024163.
[96] A. A. Rusakov, S. Iskakov, L. N. Tran, and D. Zgid, J. Chem. Theory Comput. **15**, 229 (2019),

https://doi.org/10.1021/acs.jctc.8b00927.
[97] J. J. Phillips and D. Zgid, J. Chem. Phys. **140**, 241101 (2014), http://dx.doi.org/10.1063/1.4884951.
[98] D. Neuhauser, R. Baer, and D. Zgid, ArXiv e-prints (2016), arXiv:1603.04141 [physics.chem-ph].
[99] A. A. Kananenka and D. Zgid, J. Chem. Theory Comput. **13**, 5317 (2017), https://doi.org/10.1021/acs.jctc.7b00701.
[100] D. Zgid and G. K.-L. Chan, J. Chem. Phys. **134**, 094115 (2011), http://dx.doi.org/10.1063/1.3556707.
[101] D. Zgid, E. Gull, and G. K.-L. Chan, Phys. Rev. B **86**, 165128 (2012).
[102] D. Medvedeva, S. Iskakov, F. Krien, V. V. Mazurenko, and A. I. Lichtenstein, Phys. Rev. B **96**, 235149 (2017).
[103] E. Gull, S. Iskakov, I. Krivenko, A. A. Rusakov, and D. Zgid, Phys. Rev. B **98**, 075127 (2018).
[104] A. A. Kananenka, E. Gull, and D. Zgid, Phys. Rev. B **91**, 121111 (2015).
[105] A. A. Holmes, N. M. Tubman, and C. J. Umrigar, J. Chem. Theory Comput. **12**, 3674 (2016).
[106] S. Sharma, A. A. Holmes, G. Jeanmairet, A. Alavi, and C. J. Umrigar, J. Chem. Theory Comput. **13**, 1595 (2017).
[107] A. A. Holmes, C. J. Umrigar, and S. Sharma, J. Chem. Phys. **147** (2017).
[108] J. Li, M. Otten, A. A. Holmes, S. Sharma, and C. J. Umrigar, J. Chem. Phys. **148**, 214110 (2018).
[109] C. F. Bender and E. R. Davidson, Phys. Rev. **183**, 23 (1969).
[110] B. Huron, J. Malrieu, and P. Rancurel, J. Chem. Phys. **58**, 5745 (1973).
[111] R. J. Buenker and S. D. Peyerimhoff, Theor. Chim. Acta **35**, 33 (1974).
[112] F. A. Evangelista, J. Chem. Phys. **140**, 124114 (2014).
[113] W. Liu and M. R. Hoffmann, J. Chem. Theory Comput. **12**, 1169 (2016).
[114] A. Scemama, T. Applencourt, E. Giner, and M. Caffarel, J. Comp. Chem. **37**, 1866 (2016).
[115] N. M. Tubman, J. Lee, T. Y. Takeshita, M. Head-Gordon, and K. B. Whaley, J. Chem. Phys. **145**, 044112 (2016).
[116] Y. Garniron, A. Scemama, P.-F. Loos, and M. Caffarel, J. Chem. Phys. **147**, 034101 (2017).
[117] M. Dash, S. Moroni, A. Scemama, and C. Filippi, J. Chem. Theory Comput. **14**, 4176 (2018).
[118] Y. Garniron, A. Scemama, E. Giner, M. Caffarel, and P.-F. Loos, J. Chem. Phys. **149** (2018).
[119] P.-F. Loos, A. Scemama, A. Blondel, Y. Garniron, M. Caffarel, and D. Jacquemin, J. Chem. Theory Comput. **14**, 43604379 (2018).
[120] D. Hait, N. M. Tubman, D. S. Levine, K. B. Whaley, and M. Head-Gordon, J. Chem. Theory Comput. **xx**, xx (2019).
[121] An alternative straightforward approach does not have a large memory requirement, but requires considerably larger computation time.
[122] J. E. Smith, B. Mussard, A. A. Holmes, and S. Sharma, J. Chem. Theory Comput. **13**, 5468 (2017).
[123] J. R. Trail and R. J. Needs, The Journal of Chemical Physics **139**, 014101 (2013).
[124] J. R. Trail and R. J. Needs, The Journal of Chemical Physics **142**, 064110 (2015).
[125] J. R. Trail and R. J. Needs, The Journal of Chemical Physics **146**, 204107 (2017).