



CHORUS

This is the accepted manuscript made available via CHORUS. The article has been published as:

Testing self-energy embedding theory in combination with GW

Tran Nguyen Lan, Avijit Shee, Jia Li, Emanuel Gull, and Dominika Zgid

Phys. Rev. B **96**, 155106 — Published 5 October 2017

DOI: [10.1103/PhysRevB.96.155106](https://doi.org/10.1103/PhysRevB.96.155106)

Testing self-energy embedding theory in combination with GW

Tran Nguyen Lan,^{1,2,*} Avijit Shee,¹ Jia Li,² Emanuel Gull,² and Dominika Zgid¹

¹*Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109, USA*

²*Department of Physics, University of Michigan, Ann Arbor, Michigan 48109, USA*

(Dated: September 19, 2017)

We present a theoretical framework and implementation details for self-energy embedding theory (SEET) with the GW approximation for the treatment of weakly correlated degrees of freedom and configuration interactions solver for handling the strongly correlated degrees. On a series of molecular examples, for which the exact results are known within a given basis, we demonstrate that SEET(CI/GW) is a systematically improvable and well controlled method capable of giving accurate results and well behaved causal self-energies and Green’s functions. We compare the theoretical framework of SEET(CI/GW) to that of GW+DMFT and comment on differences between these to approaches that aim to treat both strongly and weakly correlated simultaneously.

I. INTRODUCTION

The quantitative simulation of realistic correlated solids and molecules requires the accurate description of a large number of degrees of freedom. Only some of these are strongly correlated. A successful approach has therefore been the combination of weak correlation methods, such as the GW method or the density functional theory in the local density approximation (LDA), with non-perturbative methods for low-energy effective models, such as the dynamical mean field approximation.

The polynomial scaling of the weak correlation method then allows one to treat large systems, while the dynamical mean field approximation replaces the intractable lattice problem with an impurity problem coupled to a self-consistently adjusted bath, which is numerically solvable.

The combination of band structure methods with dynamical mean field theory (DMFT)^{1,2} has enjoyed great success, in particular when applied to materials with d and f shells. However, ambiguities at the interface of the weak- and strong-correlation methods, related in particular to the choice of “double counting” and the impurity interaction parameters in LDA+DMFT³, and to the proper numerical treatment of general four-fermion screened interaction terms in GW+DMFT⁴⁻⁶, have hampered progress. Moreover, the assumption that strong correlations are spatially localized, which is the premise of the dynamical mean field approximation, may not be valid in real compounds.

Recently we introduced an approximation scheme, the self-energy embedding theory (SEET)⁷⁻¹² that does not suffer from these limitations. First, due to the diagrammatic nature of the method, no double counting problem arises. Second, the absence of frequency-dependent interactions in the strongly correlated part means that standard multi-orbital impurity solvers can be employed and no ambiguity in how the Coulomb interactions are handled exist. Finally, the method does not rely on an a priori determination of the correlated subspace and in particular does not assume that strong correlations are local, but rather introduces a small control parameter that can be used to adaptively choose orbitals and sys-

tematically converge to the exact solution.

In a series of previous papers⁷⁻¹², we studied the behavior of SEET where the weak-correlation method employed was the second order perturbation theory (GF2). In the present paper, we show how SEET performs when GF2 is replaced by the GW approximation. QM/QM embedding methods for realistic systems are not well studied, and prior work on non-perturbative strong correlation methods for small systems based on diagrammatic theory have shown problems ranging from convergence to the “wrong” fixed point¹³ to causality violations¹⁴. We show here that no such problems are observed in our implementation of SEET, and that in fact the precision achieved with SEET along the entire range from weak to strong correlations is comparable to state-of-the-art quantum chemistry wave function methods, to which we carefully compare. As a testbed, we use small molecular examples, for which exact or nearly exact reference results within a given basis set are available. Energies are straightforward to compare to reference wave function methods which have entirely different strength and weaknesses; agreement therefore provides a stringent assessment of the quality of the result. Agreement for ground state energies in this context also asserts agreement for all higher temperatures, as SEET is formulated in a finite-temperature formalism that has temperature as a small parameter.

The remainder of this paper proceeds as follows. In section II, we briefly discuss the theoretical setup and the SEET functional. We then follow with benchmarking and in depth explanation of the numerical results in Sec. III. Finally, we present a theoretical discussion supporting our results and conclude in Sec. V.

II. METHODS

In this paper, we study small molecular systems. The specification of the interatomic distance and of a finite set of N Gaussian orbitals fully determines the Hamiltonian

in second quantized form,

$$H = \sum_{ij}^N t_{ij} a_i^\dagger a_j + \sum_{ijkl}^N v_{ijkl} a_i^\dagger a_j^\dagger a_l a_k, \quad (1)$$

where the operators a_i^\dagger (a_i) create (destroy) an electron in orbital i , t_{ij} denotes the single-particle contribution, and v_{ijkl} the Coulomb matrix element.

We express physical properties such as energies and single-particle response functions in a statistical mechanics approach based on an approximation to the grand partition function $Z = \text{Tr} e^{-\beta(H-\mu n)}$. Here, β denotes the inverse temperature, μ the chemical potential, and n the density operator. Our temperatures are chosen low enough that the system has converged to the ground state, and the chemical potential μ is adjusted to yield the correct particle density.

Within this framework, a functional $\Phi[G]$ of the Green's function G , which contains all linked closed skeleton diagrams,¹⁵ is used to express the grand potential as

$$\Omega = \Phi - \text{Tr}(\log G^{-1}) - \text{Tr}(\Sigma G), \quad (2)$$

where the self-energy Σ is defined with respect to a non-interacting Green's function G_0 via the Dyson equation

$$G = G_0 + G_0 \Sigma G. \quad (3)$$

The functional formalism is useful because approximations to Φ that can be formulated as a subset of the terms of the exact Φ functional can be shown to respect the conservation laws of electron number, energy, momentum, and angular momentum by construction.^{16,17} In addition, Φ -derivability ensures that quantities obtained by thermodynamic or coupling constant integration from non-interacting limits are consistent.¹⁷ Functional theory therefore provides a convenient framework for constructing diagrammatic approximations in situations where a direct solution of the Hamiltonian defined in Eq. 1 is not possible.

In order to discuss the formalism used in this paper, we first introduce GF2 and GW, then discuss SEET, and finally explain in detail the SEET+GW formalism.

A. GF2

The self-consistent second order perturbation theory, GF2, is a Φ -derivable diagrammatic approximation that includes all terms up to second order in the interaction.¹⁸⁻²⁶ In addition to the frequency independent Hartree-Fock terms, Σ_∞ , the second order self-energy contains

$$\begin{aligned} \Sigma_{ij}^{(2)}(\tau) = & - \sum_{klmnpq} G_{kl}(\tau) G_{mn}(\tau) G_{pq}(-\tau) \\ & \times v_{imqk} (2v_{lpnj} - v_{lpjn}), \end{aligned} \quad (4)$$

where $G(\tau)$ denotes the fully interacting Green's function in imaginary time that is obtained self-consistently from $\Sigma = \Sigma_\infty + \Sigma^{(2)}$ and the non-interacting Green's function G_0 using Eq. 3. The GF2 approach is performed iteratively, starting from a Hartree-Fock Green's function, until Σ or the total electronic energy is converged within a predefined tolerance. In contrast to the iterative perturbation theory often used in dynamical mean field theory,²⁷ which for the Hubbard lattice interpolates between perturbation theory and the exact solution in the limit of the separated Hubbard atoms, GF2 only relies on perturbation theory and does not recover the correct dissociated limit. Precise calculations of the second-order results require several numerical optimizations. Our implementation makes use of adaptive grids for both imaginary time²⁰ and imaginary frequency²¹ Green's functions.

B. GW

The self-consistent GW method^{28,29} is a diagrammatic approximation formulated in terms of renormalized propagators G and renormalized ("screened") interactions W . Similar to GF2, it can be written as an approximation to the Luttinger-Ward (LW) functional Φ ^{15,30} and is therefore thermodynamically consistent and conserving.^{17,30} However, it does not respect the crossing symmetries.¹² The method requires the self-consistent determination of propagators G , polarizations $P = GG$, self-energies $\Sigma = -GW$, and screened interactions W . The expressions for G and W are determined by the Dyson equations

$$G = G_0 - G_0 \Sigma G, \quad W = v + v P W, \quad (5)$$

where G_0 and v are the bare electronic propagator and Coulomb interactions defined in Eq. 1.

Self-consistent GW is only exact to first order in v , as the second-order exchange diagram is neglected. Our implementation of this approximation closely follows Refs. 12, 18, 29-31. The Green's function is initialized using the Hartree-Fock result. We then construct the polarization $P = GG$ and obtain W from Eq. (5). After computing the GW self-energy $\Sigma^{GW} = -GW$, we obtain the updated G by solving Dyson's equation, thus closing the self-consistency loop.

In order to reduce the size of W , we perform a Cholesky decomposition^{12,32} and truncation on v . This vastly reduces the numerical effort (for related decompositions see Ref. 33). Adaptive imaginary time²⁰ and frequency grids²¹ are essential to accurately represent the GW Green's function and self-energy.

C. SEET

The self-energy embedding theory⁷⁻¹¹ is a conserving approximation to the Luttinger-Ward functional Φ designed to treat strongly correlated degrees of freedom. It

consists of a two-step hierarchy in which the functional of the system is approximated by a solution of the entire system with a weak coupling method, which is then improved by the non-perturbative solution of correlated subsets of orbitals. In the most simple case, the strongly correlated subsets are disjoint (non-intersecting) and the SEET functional is defined as

$$\Phi_{\text{SEET}}^{\text{SPLIT}} = \Phi_{\text{weak}}^{\text{tot}} + \sum_{i=1}^M \left(\Phi_{\text{strong}}^{A_i} - \Phi_{\text{weak}}^{A_i} \right). \quad (6)$$

Here $\Phi_{\text{weak}}^{\text{tot}}$ denotes the approximation of the Φ functional of the entire system using a weak coupling technique. The index i enumerates the M non-intersecting, correlated subsets of orbitals A_i , and Φ^{A_i} is the functional evaluated within the orbital subset A_i , using a weak coupling or a non-perturbative method.

It is much more general to consider multiple intersecting orbital subspaces. In this case, as described in Ref. 10, the SEET functional generalizes to

$$\begin{aligned} \Phi_{\text{SEET}}^{\text{MIX}} = & \Phi_{\text{weak}}^{\text{tot}} + \sum_i^M (\Phi_{\text{strong}}^{A_i} - \Phi_{\text{weak}}^{A_i}) \\ & - \sum_{i<j}^M (\Phi_{\text{strong}}^{A_i \cap A_j} - \Phi_{\text{weak}}^{A_i \cap A_j}) \\ & + \sum_{i<j<k}^M (\Phi_{\text{strong}}^{A_i \cap A_j \cap A_k} - \Phi_{\text{weak}}^{A_i \cap A_j \cap A_k}) - \dots, \end{aligned} \quad (7)$$

where the additional terms account for the double counting of overlapping strongly correlated subspaces. The ellipse denotes summations over intersections of increasing numbers of subspaces, which enter with alternating signs.¹⁰ The indices $i, j, k, \dots = 1, \dots, M$ are used to enumerate intersecting orbital subspaces.

The solution of the SEET equations proceeds according to Refs. 9 and 11. After an initial self-energy of the system is obtained within the weak coupling method and correlated subspaces are identified, the algorithm alternates between solving impurity problems with hybridization functions determined by the weak coupling part, and solving the entire systems with propagators that contain impurity self-energies. All impurity problems can be solved in parallel, making this formalism ideally suited for high-performance computing environments.

In SEET, unlike in DMFT-type approximations, the strongly correlated orbitals are not determined a priori, and are typically not localized. Rather, they are determined after the solution of the system using the weak coupling method. We found that diagonalizing the one-body density matrix allows us to identify orbitals that have partial occupations significantly different from 0 or 2. These partial occupations can be used to guide the selection of strongly correlated orbitals. Since the selection is done based on occupation numbers obtained from weakly correlated method, it can happen that many orbitals can have similar occupations. Then visualizing

the orbitals and analyzing the contributions of parent atomic orbitals to each of the partially occupied natural orbitals can help with the final selection. By gradually adding more of the strongly correlated orbitals to the correlated subspace, convergence to the exact limit can be observed.^{7,9,10}

In the following, we compare SEET results based on the two weak coupling methods, GW and GF2. For the impurity problems containing strongly correlated orbitals, we use a full configuration interaction (CI) impurity solver and its truncated versions.^{34,35} The functional $\Phi_{\text{weak}}^{\text{tot}}$ then becomes $\Phi_{\text{GF2}}^{\text{tot}}$ and $\Phi_{\text{GW}}^{\text{tot}}$, respectively, and the functional $\Phi_{\text{strong}}^{A_i}$ is $\Phi_{\text{CI}}^{A_i}$.

Several self-consistent iterative procedures developed in other contexts have been shown to suffer from problems such as convergence to an unphysical fixed point¹³ or causality violations.¹⁴ While we cannot exclude that such problems could occur in principle, we have not observed any convergence to unphysical solutions or non-causal Green's functions or self-energies.

D. SEET with GW

The general algorithmic structure of SEET is described in Refs. 7 and 9. Here, we highlight some aspects of SEET in combination with GW.

1. Orbital choice

The self-energy embedding can be performed in any orthogonal basis, including in a localized basis [symmetrized atomic orbitals (SAOs) or localized molecular orbitals (LMOs)] or in an energy basis [natural orbitals (NOs) or molecular orbitals (MOs)]. Where we use NOs obtained from the one-body density matrix of the self-consistent GW or GF2 procedure, we select those orbitals to the strongly correlated space which have occupation numbers significantly different from 0 or 2. This space can then be split into multiple intersecting or non-intersecting smaller orbital subspaces with fewer correlated orbitals. After the selection of strongly correlated orbitals is done, quantities such as $\Sigma_{\text{weak}}^{\text{tot}}$, $[\Sigma_{\text{weak}}^{\text{tot}}]^{A_i}$, $[G_{\text{weak}}^{\text{tot}}]^{A_i}$, and $[t]^{A_s}$ are expressed in the new basis. Here $[X_{\text{weak}}^{\text{tot}}]^{A_i}$ denotes the quantity X first obtained for the whole system (denoted with superscript ‘‘tot’’) and then restricted to the orbital subspace A_i . Only a subset of the transformed Coulomb integrals $[\tilde{v}]^{A_s}$, namely those where all four orbital indices belong to the subset A_i , needs to be evaluated. Here, v stands for all two-body Coulomb integrals in the AO basis while \tilde{v} stands for all the two-body Coulomb integrals in the new basis.

2. Double counting

In SEET, the subtraction of doubly counted diagrams is rigorously defined and unique. No double counting ‘problem’ as in *e.g.* LDA+DMFT exists.³⁶ We subtract GF2 or GW self-energy evaluated exclusively within each orbital subset A_s . In case of GF2, the double counted self-energy is evaluated for $i, j \in A_s$ as,

$$[\Sigma_{\text{GF2}}^{DC}(\tau)]_{ij}^{A_s} = - \sum_{klmnpq \in A_s} [G_{\text{GF2}}^{\text{tot}}(\tau)]_{kl}^{A_s} [G_{\text{GF2}}^{\text{tot}}(\tau)]_{mn}^{A_s} \quad (8)$$

$$\times [G_{\text{GF2}}^{\text{tot}}(-\tau)]_{pq}^{A_s} [\tilde{v}]_{imqk}^{A_s} (2[\tilde{v}]_{lpnj}^{A_s} - [\tilde{v}]_{lpjn}^{A_s}).$$

The Coulomb integrals used in this evaluation are the transformed integrals $[\tilde{v}]^{A_s}$ belonging to the orbital subset A_s . The Green’s functions, similarly, are the transformed Green’s functions $[G_{\text{GF2}}^{\text{tot}}(\tau)]_{kl}^{A_s}$ belonging to the orbital subset A_s .

A similar procedure is used in GW. Here, the decomposed integrals $v_{ijkl} = \sum_{\mu} \tilde{v}_{ij}^{\mu} \tilde{v}_{kl}^{\mu}$ are transformed to the basis of choice (most frequently natural orbitals) in each orbital subset A_s , yielding $[\tilde{v}]_{mnpq}^{A_s} = \sum_{\mu} \tilde{v}_{mn}^{\mu} \tilde{v}_{pq}^{\mu}$, where m, n, p, q orbital indices belong to orbital subset A_s . The polarization is then evaluated in the orbital subspace A_s using these transformed integrals,

$$\Pi(\tau)_{\mu\nu} = -2 \sum_{ijlm \in A_s} \tilde{v}_{il}^{\mu} \tilde{v}_{jm}^{\nu} [G_{\text{GW}}^{\text{tot}}(\tau)]_{ij}^{A_s} [G_{\text{GW}}^{\text{tot}}(-\tau)]_{lm}^{A_s}. \quad (9)$$

Subsequently, this polarization diagram is used to evaluate $W(i\omega) = [1 - \Pi(i\omega)]^{-1}$ and, after its Fourier transform, the GW self-energy exclusively in the subset A_s is evaluated as

$$[\Sigma_{\text{GW}}^{DC}(\tau)]_{ij}^{A_s} = - \sum_{lm \in A_s} \sum_{\mu\nu} \tilde{v}_{il}^{\mu} \tilde{v}_{jm}^{\nu} [G_{\text{GW}}^{\text{tot}}(\tau)]_{lm}^{A_s} [W(\tau)]_{\mu\nu}. \quad (10)$$

While the Green’s functions, $[G_{\text{GW}}^{\text{tot}}(\tau)]_{lm}^{A_s}$, necessary for the evaluation of the GW self-energy are constructed as a truncation of the total Green’s function to the subset A_s , the polarization diagram $\Pi(\tau)$ and $W(\tau)$ necessary to construct the double counting corrections are evaluated using the truncated Green’s functions and transformed integrals for the subset A_s . We emphasize that they are not simply truncations of $\Pi_{\text{GW}}^{\text{tot}}(\tau)$ or $W_{\text{GW}}^{\text{tot}}(\tau)$ to the subset A_s . Such a definition of these quantities would be incorrect and would likely result in a causality violation of the total self-energy.

3. Total self-energy

The total self-energy (containing both frequency dependent and independent parts $\Sigma = \Sigma_{\infty} + \Sigma(i\omega)$) of strongly correlated orbitals in each of the subsets A_s has the form

$$[\Sigma]_{ij}^{A_s} = [\Sigma_{\text{weak}}^{\text{tot}}]_{ij}^{A_s} + ([\Sigma_{\text{strong}}]_{ij}^{A_s} - [\Sigma_{\text{weak}}^{DC}]_{ij}^{A_s}), \quad (11)$$

where the method describing “weak” correlation is either GF2 or GW. The double counting correction is subtracted in the term $[\Sigma_{\text{weak}}^{DC}]_{ij}^{A_s}$, which is given by Eq. 8 in the case of GF2 and by Eq. 10 in case of GW. The self-energy $[\Sigma_{\text{strong}}]_{ij}^{A_s}$ is obtained by solving a quantum impurity problem. The construction of the impurity model for SEET will be discussed in Sec. IID 4.

We can rewrite Eq. 11 as

$$[\Sigma]_{ij}^{A_s} = [\Sigma_{\text{weak}}^{\text{embedding}}]_{ij}^{A_s} + [\Sigma_{\text{strong}}]_{ij}^{A_s}. \quad (12)$$

Here $[\Sigma_{\text{weak}}^{\text{embedding}}]_{ij}^{A_s}$ provides an effective correction to $[\Sigma_{\text{strong}}]_{ij}^{A_s}$. While $[\Sigma_{\text{strong}}]_{ij}^{A_s}$ has all contributions from the subset of interactions $[\tilde{v}]_{ijkl}^{A_s}$ where $i, j, k, l \in A_s$, the correction $[\Sigma_{\text{weak}}^{\text{embedding}}]_{ij \in A_s}^{A_s}$ accounts for all those diagrammatic contributions obtained using \tilde{v}_{mnlk} where at least one of the indices is outside the orbital subset A_s but at the same time at least one of the indices belongs to A_s .

4. Impurity problem in SEET

In SEET, the Green’s function restricted to subspace A_s is written as

$$[G^{\text{tot}}]^{A_s} = \left([G_0^{-1}]^{A_s} - [\Sigma]^{A_s} - \Delta \right)^{-1}, \quad (13)$$

where Δ describes a hybridization between the orbital subset A_s and the environment. The self-energy $[\Sigma]^{A_s}$ is defined by Eq. 11 and $[G_0^{-1}]^{A_s}$ is the inverse of the bare Green’s function in the orbital subspace A_s , defined as

$$[G_0^{-1}]^{A_s} = (i\omega + \mu)\mathbf{1} - [t]^{A_s}. \quad (14)$$

Here $[t]^{A_s}$ denotes a kinetic energy operator truncated to the subspace A_s .

The embedding of correlated orbitals into a background of other orbitals requires the solution of a quantum impurity problem with an impurity solver algorithm, which provides the interacting self-energy. The impurity \mathcal{G}_0 Green’s function is defined as

$$\mathcal{G}_0^{-1} = [G_0^{-1}]^{A_s} - [\Sigma_{\infty}^{NDC}]^{A_s} - \Delta, \quad (15)$$

where the frequency independent part of the self-energy is constructed as

$$[\Sigma_{\infty}^{NDC}]^{A_s} = [\Sigma_{\infty}^{\text{tot}}]^{A_s} - [\Sigma_{\infty}^{DC}]^{A_s} \quad (16)$$

$$[\Sigma_{\infty}^{\text{tot}}]_{ij \in A_s}^{A_s} = \sum_{kl} \gamma_{kl} (\tilde{v}_{ijkl} - 0.5\tilde{v}_{ilkj}) \quad (17)$$

$$[\Sigma_{\infty}^{DC}]_{ij \in A_s}^{A_s} = \sum_{kl \in A_s} \gamma_{kl} (\tilde{v}_{ijkl} - 0.5\tilde{v}_{ilkj}), \quad (18)$$

with the density matrix γ obtained from GF2 or GW. Note that $[\Sigma_{\infty}^{DC}]^{A_s}$ is used to subtract double counting coming from the subspace A_s .

A quantum impurity solver will obtain an expression for a correlated $[G^{\text{imp}}]^{A_s}$ given Δ , $[\Sigma_\infty^{NDC}]^{A_s}$, and G_0 (Eq. 14) as well as a subset of interactions $[\tilde{v}]^{A_s}$. As we formulate SEET in this work for the Luttinger Ward Φ functional, rather than for the Almladh Ψ functional, the impurity interactions remain instantaneous.

Using the impurity problem Dyson equation, the self-energy for a strongly correlated orbital subset is obtained as

$$[\Sigma_{\text{strong}}]^{A_s} = \mathcal{G}_0^{-1} - ([G^{\text{imp}}]^{A_s})^{-1}. \quad (19)$$

Once this strongly correlated $[\Sigma_{\text{strong}}]^{A_s}$ is known, the total self-energy, $[\Sigma]^{A_s}$, in subspace s is evaluated according to Eq. 11.

In general, for a realistic system described by a Hamiltonian with non-local one- and two-body interactions, the hybridization Δ in the orbital subspace A_s contains off-diagonal terms (*i.e.* $\Delta_{ij} \neq 0$ for $i \neq j$). However, let us note that hybridizations in the MO or NO basis are often almost diagonal, so that these bases either completely eliminate the off-diagonal terms or greatly reduce them. Additionally, the magnitude of the diagonal hybridization elements is often substantially smaller than those obtained in the local basis of atomic orbitals. We will come back to these points in Sec. III A 2.

One of the few impurity solvers able to treat general hybridization functions and general interactions at low temperature is the configuration-interaction (CI) method or its truncated versions (RASCI)^{34,35}. (For later implementations of these methods see Refs. 37–39). When applied to impurity problems the CI method relies on a fit of the hybridization function to a discretized set of non-interacting bath levels, which we perform on the imaginary axis.

E. Interactions used for the impurity Hamiltonian

In both periodic and molecular systems, for an orbital subset A_s the self-energy as given by Eq. 12 has two contributions. The first arises due to the presence of non-local interactions and accounts for all those diagrammatic contributions obtained using \tilde{v}_{mnkl} where at least one of the indices is outside the orbital subset A_s but at the same time at least one of the indices belongs to A_s . The second one has all contributions from the subset of interactions $[\tilde{v}]_{ijkl}^{A_s}$ where $i, j, k, l \in A_s$.

For example, for the on-site frequency-independent part of the self-energy, we have $[\Sigma_\infty]_{ii} = \sum_{kl \neq i} \gamma_{kl} (v_{iikl} - 0.5v_{ilki}) + 2 \sum_{l \neq i} \gamma_{il} (v_{iil} - 0.5v_{ilii}) + 0.5\gamma_{ii}v_{iiii}$, where $\sum_{kl \neq i} \gamma_{kl} (v_{iikl} - 0.5v_{ilki}) + 2 \sum_{l \neq i} \gamma_{il} (v_{iil} - 0.5v_{ilii})$ arises due to the non-local interactions and $0.5\gamma_{ii}v_{iiii}$ is a contribution from on-site integrals. In this sense, the first contributions provides a “screening” to the on-site contribution of $0.5\gamma_{ii}v_{iiii}$, thus resulting in a modified overall $[\Sigma_\infty]_{ii}$ value. Exactly the same reasoning can be used when frequency dependent self-energy is considered.

As long as the contributions to the self-energy from non-local interactions can be evaluated and the “bare” self-energy within an orbital subset A_s (evaluated with diagrams having only $[\tilde{v}]_{ijkl}^{A_s}$ where $i, j, k, l \in A_s$) can be modified due to them, the overall self-energy has a proper magnitude. These effects are discussed in details in Ref. 11. For a detailed study showing the reduction of effective local interactions when non-local interactions are discarded, see Refs. 9 and 40.

Alternatively, if one chooses to neglect contributions from non-local interactions to the self-energy in the subset A_s , one can instead a priori modify the value of the interactions in the subset A_s such that the overall value of the self-energy within the subset remains proper. Using the example with $[\Sigma_\infty]_{ii}$ this would amount to $[\Sigma_\infty]_{ii} = 0.5\gamma_{ii}v_{iiii} + \sum_{kl \neq i} \gamma_{kl} (v_{iikl} - 0.5v_{ilki}) + 2 \sum_{l \neq i} \gamma_{il} (v_{iil} - 0.5v_{ilii}) = 0.5\gamma_{ii}U_{iiii}$, where U_{iiii} are effective or “screened” interactions that are evaluated/chosen such that the effective on-site contribution to the self-energy recovers $[\Sigma_\infty]_{ii} = 0.5\gamma_{ii}U_{iiii}$.

With this in mind, it becomes evident why the local interactions in GW+DMFT have to be screened. In GW+DMFT, at the beginning of calculations, one performs a mapping to a small orbital subspace and one evaluates screened interactions in this orbital subspace using a constrained RPA approach. This mapping from the whole space of all possible orbitals to a chosen, small subspace where both GW and a subsequent DMFT calculation are carried out means that one cannot explicitly calculate the contribution to the self-energy coming from those orbitals that are discarded in the initial mapping. Thus, to recover a proper value of the self-energy, screened interactions in the small orbital subspace have to be used.

In contrast, no orbitals are discarded in SEET at any point of the calculation. Thus, the self-energy contributions arising from non-local interactions (here $[\Sigma_{\text{weak}}^{\text{embedding}}]_{ij}$) are evaluated at the GF2 or GW level and the presence of this explicit contribution effectively “screens” or changes the value of the self-energy in the subset A_s . This means that the impurity problem in SEET can be evaluated using only bare subset interactions. We stress that the problem of modification of the value of the local self-energy in the presence of non-local interactions is not specific to periodic systems but also occurs in molecular systems, as shown in Ref. 40. The lack of an appropriate adjustment of these interactions is presumably the main cause for the discrepancy between molecular DMFT and exact diagonalization results in Ref. 41.

F. SEET self-consistency

The main algorithmic step in the SEET self-consistency are as follows (see appendix of Ref. 11 as well as Ref. 9 for a more extensive description).

1. Perform a Hartree-Fock (HF) or density functional theory (DFT) calculation. Then, using this starting point, perform a self-consistent GF2 or GW calculations yielding G_{GF2}^{tot} or G_{GW}^{tot} for the entire system.
2. If using a natural orbital basis, evaluate and diagonalize the GF2 or GW density matrix to obtain natural orbitals and occupation numbers. We define strongly correlated orbitals as those with occupations numbers significantly different from 0 or 2. If using a local basis (SAO), choose the correlated orbitals based on spatial criteria.
3. Transform all Green's function and self-energies as well as t and v integrals⁴² from the atomic orbital (AO) basis to an orthogonal orbital basis such as, MOs, NOs, LMOs, and SAOs. For details see Sec IID 1.
4. Construct using Eq. 8 or Eq. 10 the GF2 or GW self-energy which will be subtracted to eliminate the double counting.
5. Construct the impurity Hamiltonian, in which two-electron term is a subset of bare Coulomb interaction. Note that this interaction matrix was already transformed to the new orthogonal basis. For details see Sec. IID 4.
6. Perform the inner self-consistency loop:
 - (a) For every orbital subset A_s , use an impurity solver (RASCI/FCI) to obtain the impurity Green's function $[G^{imp}]^{A_s}$ and extract the impurity self-energy $[\Sigma_{strong}]^{A_s}$. At the first iteration the hybridization Δ is initialized using GF2 or GW quantities.
 - (b) Set up the total self-energy according to Eq. 11.
 - (c) Reconstruct the total Green's function via the Dyson equation and adjust the chemical potential to obtain a correct electron number for the whole system.
 - (d) Update the hybridization Δ using the new total Green's function and self-energy.
 - (e) Go back to step 6(a) and iterate until convergence is reached.
7. Pass the converged Green's function to a GF2 or GW calculation and perform only a single iteration.
8. Go back to step 2 and iterate until outer loop convergence is reached.

In practice, we found that closing the outer self-consistency loop and performing steps 7 and 8 has little effect when a natural orbital basis is used. In the SAO basis, the outer loop has a larger effect and is necessary⁹.

III. RESULTS

Unless otherwise noted, the ORCA program⁴³ was used for all calculations using standard methods such as second-order Møller-Plesset perturbation theory (MP2), FCI, complete active space configuration interaction (CASCI), and n -electron valence state second-order perturbation theory (NEVPT2)^{44,45}. A locally modified version of the DALTON code⁴⁶ was employed to generate a restricted HF reference prior to GF2 and GW calculations.

For all hydrogen chains examined, GF2 and GW electronic energies were converged to 10^{-5} Ha. All SEET(CI/GF2) and SEET(CI/GW) calculations were then executed with the convergence threshold of 10^{-4} Ha.

The inverse temperature β was set at 100 Ha^{-1} or 200 Ha^{-1} depending on the system and geometry. The Matsubara frequency grid was generated using the spline interpolation²¹ with the maximum number of points varying between 20,000 and 50,000.

We use the following shorthand notation, SEET(method strong/method weak)-[Ko] stands for SEET employing "method weak" as a low-level, weak correlation method for the entire system. The impurity problems composed of K strongly correlated orbitals and a number of bath orbitals necessary to fit the hybridization are solved using "method strong". Moreover, the two variants of SEET functional are denoted as SEET(method strong/method weak)-split[$L \times Ko$] and SEET(method strong/method weak)-mix[Ko]. The former employs "method strong" to solve impurity problems created from L non-intersecting orbital subspaces containing K strongly correlated orbitals while the latter is executed for intersecting orbital subspaces containing K strongly correlated orbitals.

A. H₂ molecule

1. Energetics

We first examine the implementation of SEET with GW as a weak correlation method for the simplest molecular system, H₂ in TZ(Dunning) basis⁴⁷. In this system, there are 2 electrons in 6 orbitals. Our results using SEET(FCI/HF)-[2o], SEET(FCI/GF2)-[2o], and SEET(FCI/GW)-[2o] are summarized in Fig. 1. The full configuration interacting (FCI) method is used as an exact solution within this basis set. The weakly correlated methods such as restricted HF, GF2 and GW that are based on restricted HF reference are provided for comparison.

In the weakly correlated limit, around the equilibrium geometry, both GF2 and GW agree qualitatively well with FCI. In contrast, in the strongly correlated limit, for large interatomic distances, the GF2 and GW are

qualitatively incorrect resulting in large errors. A significant improvement upon all of the weakly correlated methods is achieved when the SEET(FCI/method weak)-[2o] is executed. For GF2 and GW, the two strongly correlated impurity orbitals are identified as natural orbitals with partial occupancy at large interatomic distances. These two orbitals can be then systematically followed until short interatomic distances are reached. For HF, we defined strongly correlated orbitals as the highest occupied molecular orbital (HOMO) and the lowest unoccupied orbital (LUMO). Though the SEET(FCI/HF) curve is nearly parallel to the FCI curve for large interatomic distances, its errors are still considerable, with the maximum error of 0.019 Ha. These errors arise due to the lack of weak correlation between the strongly correlated impurity orbitals and the remaining (environment) orbitals. SEET(FCI/GF2) and SEET(FCI/GW), which can capture strong correlations within the impurity orbitals and weak correlations among the environment orbitals and between the environment and strongly correlated orbitals, yield curves closely following the FCI curve with maximum errors of 0.005 and 0.009 Ha for SEET(FCI/GF2)-[2o] and SEET(FCI/GW)-[2o], respectively.

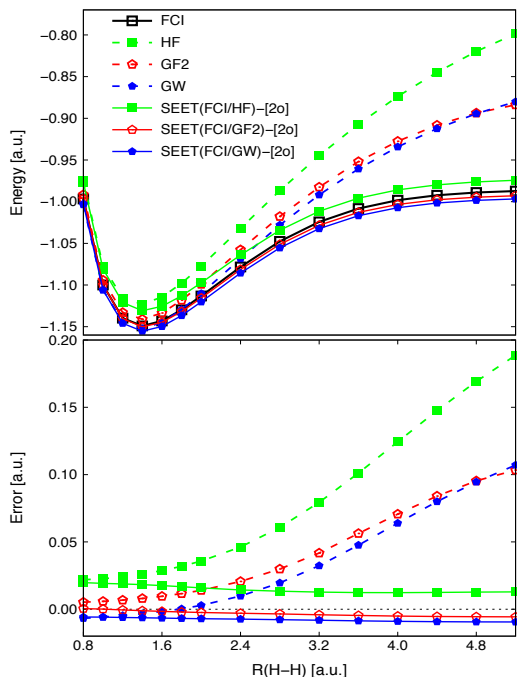


FIG. 1. Upper panel: Potential energy curves of H_2 in TZ(Dunning) basis. Lower panel: Errors with respect to the FCI reference.

Occupation numbers provide an additional insight into performance of SEET. In Fig. 2, we plot the occupation numbers as a function of the interatomic distance from low-level methods as well as different variants of SEET. FCI occupation numbers are listed for comparison. Both GF2 and GW occupation numbers differ from 2 and 0

when the interatomic distance is large but remain significantly different from the FCI reference. All SEET variants correctly describe the transition from the weakly to strongly correlated regime and SEET occupation numbers follow the trend displayed by FCI very well.

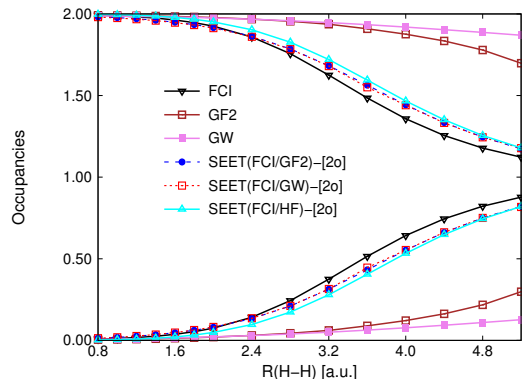


FIG. 2. Occupation numbers of H_2 in TZ(Dunning) as a function of interatomic distance.

2. Green's function, self-energy, and hybridization

In this section we will discuss both the hybridization and impurity self-energy to provide insight into the SEET performance.

SEET employs an accurate, non-perturbative method to solve the impurity problem containing strongly correlated orbitals and an approximate method to treat the remaining weakly correlated orbitals. Consequently, our major goal is to find an orbital basis that maximizes the magnitude of the impurity self-energy and minimizes the magnitude of the hybridization between the impurity orbitals and the environment. A small magnitude of the hybridization means that there is only a small electron exchange between the impurity and the environment and essentially only a weak entanglement of the strongly correlated impurity orbitals with the environment. The impurity self-energy, which illustrates many-body effects between the strongly correlated orbitals, can be large since an accurate, non-perturbative method is employed to calculate it.

Natural orbitals, which are eigenvectors of the one-body density matrix, are an orbital basis that fulfills the above criteria. It is evident from the top panel of Fig. 3 that when the orbitals are strongly correlated for $R \geq 3.6$ a.u., the imaginary part of hybridization is small while the imaginary part of the impurity self-energy is large. Only for $R = 1.4$ a.u. the hybridization is larger than the self-energy. However, in this limit, the system is weakly correlated and essentially the weakly correlated method is capable of describing it. From the bottom panel of Fig. 3, we see that in this weakly correlated limit ($R = 1.4$ a.u.), the magnitude of the frequency dependent impurity self-energy is similar to the total self-energy from Eq. 11.

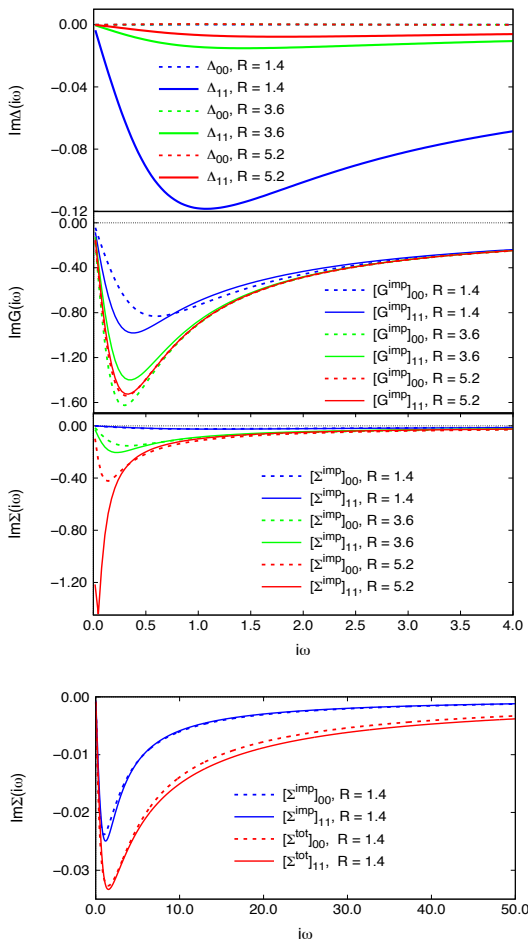


FIG. 3. Upper panels: Imaginary part of the diagonal elements of the hybridization, impurity Green’s function, and impurity self-energy evaluated at various bond lengths for H_2 in the TZ(Dunning) basis. Lower panel: Imaginary part of diagonal elements of the impurity self-energy and total self-energy (see Eq. 12) evaluated at a distance of 1.4 (a.u.) for H_2 in the TZ(Dunning) basis. Natural orbitals were used for all calculations.

Consequently, as expected, when hybridizations are large to achieve quantitative accuracy, it is essential for a well behaved quantitative theory to contain $\Sigma_{\text{weak}}^{\text{embedding}}$ as described in Eq. 12.

Recently, Ref. 14 reported non-causal hybridizations for the H_2 molecule for large interatomic distances when the GW+DMFT method was employed (for other recent GW+DMFT work see also Refs. 48–50). In SEET, as evident from Fig. 3, we do not observe any non-causal hybridizations, self-energies, or Green’s functions. In fact, at large interatomic distances, when natural orbital basis is employed, the hybridization of the two strongly correlated orbitals with the environment is close to zero. At

present it is unclear if the causality issue observed in Ref. 14 is a deficiency of the implementation or a more fundamental theory problem.

B. H_{10} chain

Here, we will further focus on a detailed comparison between SEET(CI/GF2) and SEET(CI/GW). There are two aspects in which SEET(CI/GF2) and SEET(CI/GW) can yield different results that we would like to examine:

- (i) the recovery of strong correlations when the strongly correlated orbitals are partitioned between multiple impurity problems. This happens when there are too many strongly correlated orbitals to contain them in a single impurity that is treated by a non-perturbative solver,
- (ii) the description of weak correlations (outside the strongly correlated space) among weakly correlated orbitals.

To access point (i) we employ a simple H_{10} chain in STO-6G^{51} basis. In this minimal basis, for large interatomic distances, all 10 orbitals are strongly correlated. This means that if we use a single impurity to contain all the strongly correlated orbitals, it would need to contain all 10 of them. Since most solvers, at low temperatures, cannot handle impurities with more than 5-6 strongly correlated orbitals, we choose to split these 10 orbitals into two smaller orbital subspaces containing 6 and 4 orbitals each. This means that strong correlations that are potentially arising between these two orbital subspaces are treated only at the perturbative level by GF2 or GW. The results of such a treatment employing SEET(FCI/GF2)-split[6o+4o]/SAO and SEET(FCI/GW)-split[6o+4o]/SAO are listed in the left panel of Fig. 4. Here, SAO denotes symmetrically orthogonal atomic orbital basis that results in almost localized orbitals.

As expected, both GF2 and GW capture weak correlation at the quantitative level. However, even though they can partially capture the strong correlations, they are not sufficient to accurately describe the dissociation limit (strongly correlated limit) quantitatively. Both SEET(FCI/GF2)-split[6o+4o] and SEET(FCI/GW)-split[6o+4o] improve dramatically upon GF2 and GW results at long distances when strong correlations are present. The differences between SEET(FCI/GF2) and SEET(FCI/GW) are more pronounced in the strongly correlated regime than in the weakly correlated one. While SEET(FCI/GF2)-split[6o+4o] energies are below the FCI reference past the distance of 2.8 a.u., such energy overestimation is much smaller for SEET(FCI/GW)-split[6o+4o]. Errors at the distance of 4.0 a.u. are -0.027 Ha and -0.002 Ha for SEET(FCI/GF2)-split[6o+4o] and SEET(FCI/GW)-split[6o+4o], respectively.

To further examine both (i) and (ii) points simultaneously, we employ an H_{10} chain in cc-pVDZ⁵². In the cc-pVDZ basis, there are 50 orbitals in total. Forty of them are weakly correlated while the remaining ten orbitals, for large internuclear distances, are strongly correlated. The DMRG reference was taken from Ref. 12. We start with the examination of SEET results in NOs which are obtained from the weak correlation method such as GF2 or GW, see the middle panel of Fig. 4. In the NO basis, orbitals remain delocalized.

In the strongly correlated regime, for large interatomic distances, both GF2 and GW that are based on restricted HF reference largely deviate from the DMRG results and yield underestimated correlation energies. In the weakly correlated regime, for short intermolecular distances, the behavior of GW differs from that of GF2. In this regime, GF2 potential energy curve is always above the DMRG reference, while the GW curve is below the DMRG reference, yielding overestimated correlation energies.

Performing SEET and splitting the strongly correlated orbitals into non-intersecting orbital subspaces containing 4 and 6 strongly correlated orbitals significantly improves the results. In particular, SEET(CI/GF2)-split[6o+4o]/NO noticeably reduces GF2 errors for the whole range of distances. SEET(CI/GW)-split[6o+4o]/NO, however, only improves GW energies at long distances where the strong correlations are important. At short interatomic distances, no visible improvement due to SEET(CI/GW)-split[6o+4o]/NO is observed. The overestimation of GW correlation energy at short distances cannot be improved by SEET. This is because in this regime, correlations are dominated by the weak correlations between the strongly correlated and weakly correlated orbitals, rather than only the strong correlations within the space of the strongly correlated orbitals.

Now, let us focus on longer interatomic distances for the example of H_{10} chain in cc-pVDZ⁵² basis using NOs. Here, in the SEET-split scheme strong correlations between both the orbitals subspaces containing 4 and 6 orbitals become very important allowing us to examine point (i) in detail. Simultaneously, in this larger basis there are of course correlations between the strongly correlated orbitals and the remaining weakly correlated ones giving rise to correlations described by point (ii). However, for long interatomic distances, the missing strong correlations appearing in the SEET-split scheme dominate any effects appearing due to the weak correlations.

For large interatomic distances, both SEET(CI/GF2)-split[6o+4o]/NO and SEET(CI/GW)-split[6o+4o]/NO are unable to sufficiently capture the strong correlations present between orbital subspaces containing 4 and 6 orbitals. Both SEET(CI/GF2)-split[6o+4o]/NO and SEET(CI/GW)-split[6o+4o]/NO result in large nonparallelity errors (NPEs)⁵³ of 0.068 Ha and 0.105 Ha, respectively. These significant NPEs are caused by the failure of SEET-split scheme in NO basis to recover important correlations between two non-intersecting strongly orbital

subspaces at large interatomic distances.

To investigate if this difficulty can be removed, we employed the generalized version of SEET (SEET-mix) allowing for the presence of intersecting strongly correlated orbital subspaces¹⁰. NPEs of SEET(CI/GF2)-mix[6o]/NO and SEET(CI/GW)-mix[6o]/NO are now 0.013 Ha and 0.015 Ha, respectively. These NPEs are one magnitude smaller than those of SEET-split scheme, indicating a proper description of the strong correlations between intersecting orbital subspaces.

While the NPEs are similarly small for both SEET-mix scheme with GF2 and GW, it is worth noting that the SEET(CI/GF2)-mix[6o]/NO curve is in a very good agreement with the DMRG reference with the maximum absolute error of 0.016 Ha. In contrast, due to the huge overestimation of GW correlation energy, reflected by its very low total energies at short distances, the SEET(CI/GW)-mix[6o]/NO curve remains parallel to however much below the DMRG reference. We also observed a similar behavior in the already considered H_2 example in TZ(Dunning) basis. Consequently, once the SEET-mix scheme is applied the weak correlations described by point (ii) are responsible for the overall quality of results.

When examining the example of H_{10} chain in the minimal STO-6G basis using local SAOs, we observed that in the strongly correlated regime, SEET(FCI/GW)-split gives better results than SEET(FCI/GF2)-split. It is therefore interesting to further examine the behavior of SEET(FCI/GW)-split versus SEET(FCI/GF2)-split when localized orbitals are employed for a larger basis set than the minimal one. To consider such a case, we continue to focus on the H_{10} chain in cc-pVDZ basis. Ten valence orbitals from the set of fifty NOs are localized using the Boys localization procedure⁵⁴. We will call these orbitals the localized natural orbitals (LNOs). Let us note here that both Boys and Wannier localizations are equivalent when finite systems are considered.

Here, yet again our major goal is to examine points (i) and (ii) simultaneously. First, we illustrate how strong correlations between non-intersecting subspaces are recovered when using LNO basis. The results are displayed on the right panel of Figure 4.

While there is no improvement of NPE for SEET(FCI/GF2)-split[6o+4o] (0.062 Ha in LNOs compared to 0.068 Ha in NOs), NPE for SEET(FCI/GW)-split[6o+4o] in LNOs (0.037 Ha) is considerably reduced from its counterpart in the delocalized NO basis (0.105 Ha).

Similarly to the case of the H_{10} chain in STO-6G, the SEET(FCI/GF2)-split[6o+4o]/LNO curve is below the DMRG one at long distances implying an overcorrelation in SEET with GF2. While the NPE of SEET(FCI/GW)-split[6o+4o]/LNO is small, the total energy is overestimated for all intermolecular distances.

In all cases considered, irrespective of the orbital basis or basis set, we have observed a systematic overcorrelation of total energies when any variant of SEET(CI/GW)

is employed. The overcorrelation present in GW becomes more pronounced when a larger basis set is used. Consequently, in the SEET(CI/GW) approach, the overcorrelation present in GW itself is similarly affecting the description of weakly correlated orbitals outside the strongly correlated orbital space (point (ii)). SEET(CI/GW)-mix scheme was not able to remove these overcorrelations.

For SEET(CI/GF2)-split, we observed the total energy overcorrelation only at large interatomic distances when the localized basis SAO or LNO was applied. In the NO basis, we observed underestimation of correlation energies. These problems in both localized (SAO, LNO) or energy (NO) basis can be partially relieved when SEET(CI/GF2)-mix scheme is employed¹⁰.

C. H₅₀ chain

We now apply the SEET(CI/GW) implementation to a much longer hydrogen chain, H₅₀ in the STO-6G basis. In quantum chemistry, this is a well-known benchmark for strongly correlated methods since for large interatomic distances the full strongly correlated space is large and contains 50 electrons in 50 orbitals. The reference solution is available from DMRG calculations.⁵⁵ For this system, traditional methods suitable for weakly correlated systems, such as CCSD(T), are unable to converge past the distance of 2.0 a.u.⁵⁵ We employ here the SAO basis, which yields almost localized orbitals. In SEET-split calculations, the full strongly correlated space of 50 orbitals is split into 25 non-intersecting subspaces composed of 2 orbitals each. Correlations arising between these non-intersecting subspaces are then treated either by GF2 or GW. The resulting potential energy curves and total errors with respect to the DMRG reference are summarized in Fig. 5.

In the weakly correlated regime ($R < 2.0$ a.u.), GF2 and GW give results of a similar quality to SEET(FCI/GF2)-split[25×2o] and SEET(FCI/GW)-split[25×2o]. Both SEET(FCI/HF)-split[25×2o] and SEET(FCI/HF)-mix[6o] significantly improve the HF result. In the strongly correlated regime ($R \geq 2.0$ a.u.), both GF2 and GW based on the RHF reference are unable to properly describe strong correlations and display large underestimation of correlation energies.

For these longer distances, SEET(FCI/HF)-split[25×2o], SEET(FCI/GF2)-split[25×2o], and SEET(FCI/GW)-split[25×2o] show a large improvement over the HF, GF2, and GW results, providing energies that are close to the DMRG reference. As opposed to SEET(FCI/GF2)-split[25×2o], SEET(FCI/GW)-split[25×2o] has a much smaller overcorrelation in the strongly correlated regime. The total error (-0.522 Ha) present in SEET(FCI/GF2)-split[25×2o] at the distance of 4.2 a.u. is one order of magnitude larger than that (-0.055 Ha) obtained with SEET(FCI/GW)-split[25×2o].

At large distances, in the minimal basis set when SAOs are used, these errors arise due to overcorrela-

tions present in SEET-split scheme with both GF2 and GW employed as weak correlation methods that are recovering correlations between non-intersecting subspaces of strongly correlated electrons. At least in the minimal basis, when SAOs are used, this overcorrelation can be removed when SEET(FCI/HF)-split[25×2o] or SEET(FCI/HF)-mix[6o] are performed, for details see Ref. 10. For SEET(FCI/HF)-mix[6o]/SAO that uses intersecting strongly correlated orbital subspaces, the strong correlation between orbitals is recovered well.

Results for SEET(FCI/HF)-split[25×2o] and SEET(FCI/HF)-mix[6o] of this quality are only possible in the minimal basis and when localized orbitals are employed. In a larger basis, where both weakly and strongly correlated orbitals are present and where there are fewer strongly than weakly correlated orbitals, SEET((FCI/GF2) or SEET((FCI/GW) are necessary to get quantitative results. We will discuss such a case in the next section.

D. N₂ molecule

In this section, we investigate the performance of SEET in combination with GW and GF2 for an N₂ molecule in the 6-31G basis. This system has 18 orbitals in total and six of them are strongly correlated for larger bond distances. Stretching the triple bond of N₂ is a difficult test case for many quantum chemistry methods. Potential energy curves evaluated with different variants of SEET and with standard wavefunction methods such as MP2, CCSD(T), CASCI, and NEVPT2 are shown in Fig. 6.

Standard single reference methods such as MP2 and CCSD(T) diverge once the bond is stretched. CASCI(6e,6o) and NEVPT2(6e,6o), which capture the strong correlations arising among 6 orbitals when the triple bond is stretched well, produce correct dissociation curves. The NEVPT2(6e,6o) energies are much lower than those of CASCI(6e,6o) due to the inclusion of weak correlations at the perturbative second order level between 6 strongly correlated orbitals and the remaining 12 weakly correlated orbitals. In CASCI(6e,6o), only correlations among the 6 strongly correlated orbitals are included, and the electronic effects among the remaining 12 orbitals are only illustrated at the HF level. It is quite surprising that MP2 gives lower energies than NEVPT2. This behavior was previously observed in Ref. 56.

To perform the SEET calculation, we use the following procedure. First, SEET(FCI/HF)-[6o] using HF MOs is performed. Strongly correlated orbitals are identified as those 6 valence molecular orbitals with orbital energies closest to the Fermi level. These strongly correlated orbitals are delocalized and we do not perform any localization procedure. On top of the SEET(FCI/HF)-[6o] reference, we carry out one iteration of GF2 or GW. The resulting natural orbitals as well as Green's functions and self-energies coming from this one iteration of GF2 or

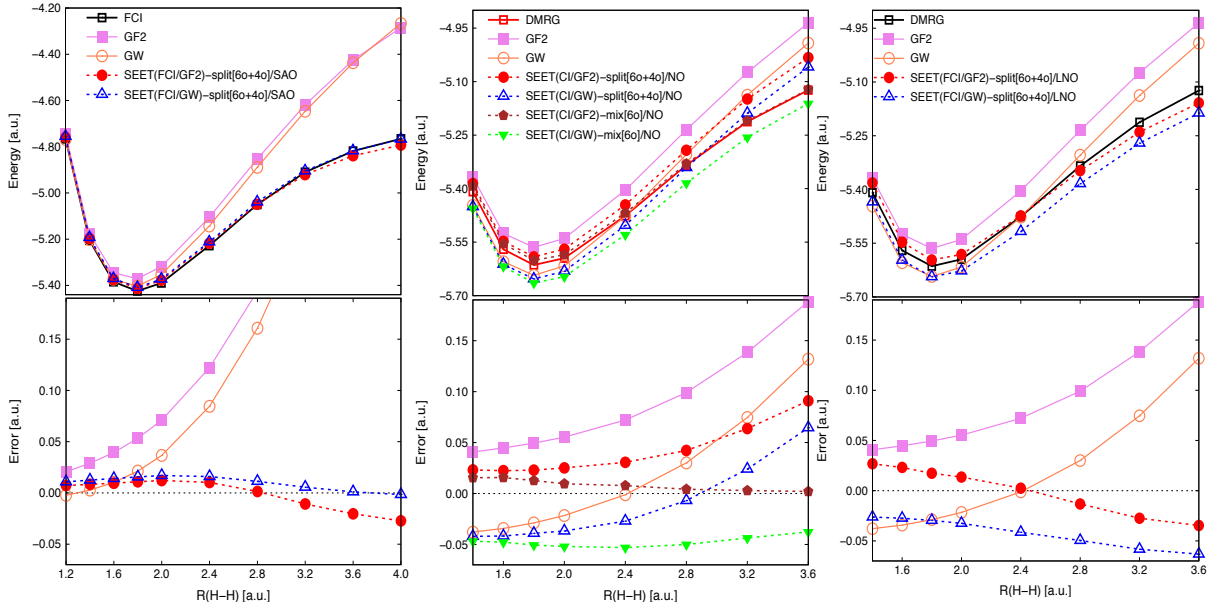


FIG. 4. Left panel: Potential energy curves (upper) and total errors (lower) with respect to the FCI reference for H₁₀ chain in STO-6G basis. Middle panel: Potential energy curves (upper) and total errors (lower) with respect to the DMRG reference for H₁₀ chain in cc-pVDZ basis. SEET calculations were performed using natural orbital (NO) basis. Right panel: Potential energy curves (upper) and total errors (lower) with respect to the DMRG reference for H₁₀ chain in cc-pVDZ basis. SEET calculations were performed using localized natural orbital (LNO) basis. The DMRG reference was taken from Ref. 12.

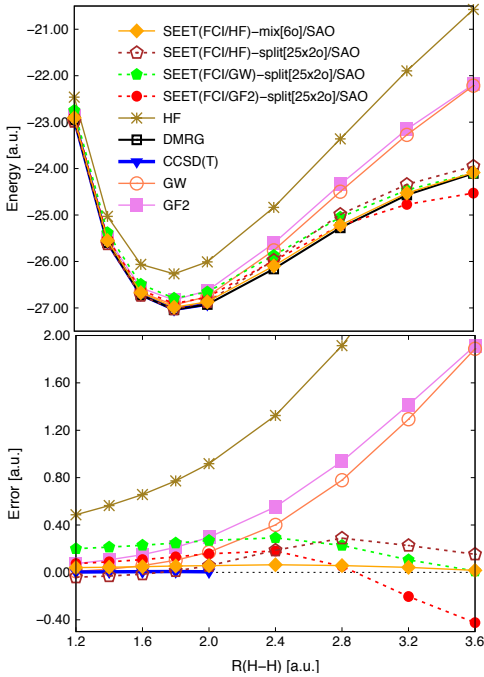


FIG. 5. Upper panel: Potential energy curves of H₅₀ chain in STO-6G basis. Lower panel: Errors from the DMRG reference. DMRG data were taken from Ref. 55.

GW are then used to initialize SEET(CI/GF2)-[6o] or SEET(CI/GW)-[6o] calculations.

The SEET(CI/HF)-[6o] curve is essentially identical

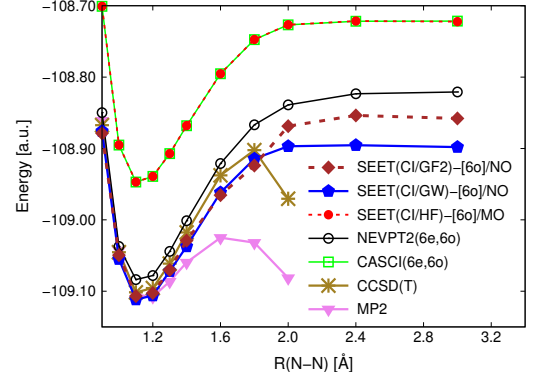


FIG. 6. Potential energy curves of N₂ in the 6-31G basis.

to the CASCI(6e,6o) curve. The weak correlations illustrated in GF2 and GW part of SEET(CI/GW)-[6o] and SEET(CI/GF2)-[6o] contribute to a significant lowering of the energy when compared to SEET(CI/HF)-[6o].

Around the equilibrium geometry, both SEET(CI/GW)-[6o] and SEET(CI/GF2)-[6o] yield energy curves close to MP2 and CCSD(T). Impressively, for stretched bond distances, unlike standard single reference methods such as MP2 and CCSD(T), SEET(CI/GF)-[6o] and SEET(CI/GW)-[6o] curves do not display any divergent behavior and remain parallel to those of CASCI(6e,6o) or NEVPT2(6e,6o). This means that both SEET(CI/GF2)-[6o] and SEET(CI/GW)-[6o] properly illustrate the strong correlations necessary to qualitatively describe the N₂ dissociation. Additionally,

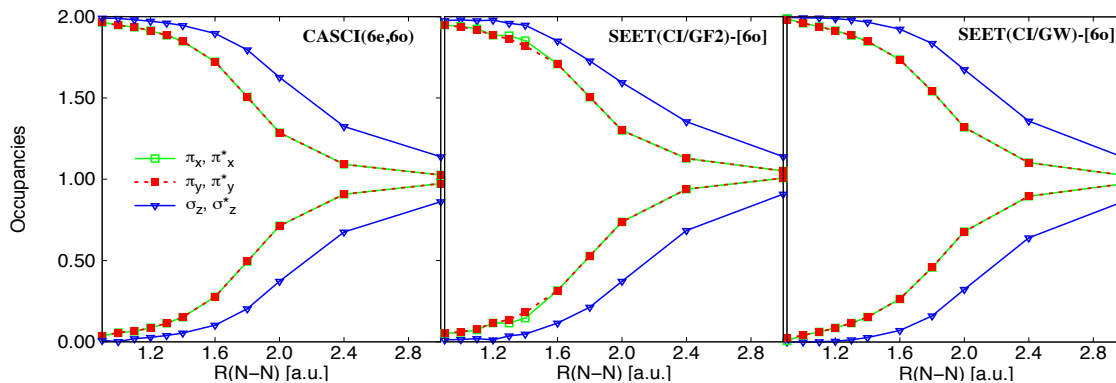


FIG. 7. Occupation numbers of 6 valence shell orbitals in N_2 ($\pi_x, \pi_x^*, \pi_y, \pi_y^*$, and σ_z, σ_z^*) obtained from CASCI(6e,6o) (left), SEET(CI/GF2)-[6o] (middle), and SEET(CI/GW)-[6o] (right) as a function of interatomic distance⁵⁷. Note that the π orbitals are spatially degenerate.

they also include the weak correlations necessary for quantitative results. Note that SEET(CI/GF2)-[6o] displays a small kink near the geometry point where CCSD(T) diverges.

To further highlight that the correct description of correlation in SEET(CI/GF2) and SEET(CI/GW) is not a fortuitous coincidence, we present a comparison of occupation numbers.

For N_2 in the 6-31G basis, CASCI(6e,6o) yields almost identical occupation numbers to CASCI(10e,16o). The latter method correlates 16 orbitals at the FCI level while keeping the remaining two orbitals that have very low energy frozen. A detailed list of all the occupation numbers in CASCI(6e,6o), CASCI(10e,16o), SEET(CI/GF2), and SEET(CI/GW) is provided in the supplement.

For clarity, in Fig 7 only focuses on a comparison of the CASCI(6e,6o) occupation numbers for the six strongly correlated orbitals $\pi_x, \pi_x^*, \pi_y, \pi_y^*$ and σ_z, σ_z^* to different variants of SEET. It is evident that SEET(CI/GF2)-[6o] and SEET(CI/GW)-[6o] all yield occupation numbers very similar to those from CASCI(6e,6o).

IV. RELATION OF SEET(CI/GW) TO GW+DMFT

In this paper we discussed the theoretical framework and implementation for SEET with GW as a weakly correlated method and CI as a strongly correlated method. A related method, GW+DMFT^{4,5,58} aims to describe the same class of physical problems where strongly and weakly correlated degrees of freedom are treated simultaneously. It is therefore instructive to summarize commonalities and differences.

The main difference between the two approaches is that SEET is built as an approximation to the Luttinger-Ward Φ functional, which is a functional of the Green's function and the bare interactions, whereas GW+DMFT is based on an approximation to the Ψ functional, which is a functional of the Green's function and the screened

interaction W . These distinct functional constructions lead to a fundamental difference in the type of impurity problem solved in SEET(CI/GW) and GW+DMFT. In SEET, the impurity problem is defined by the bare interactions (transformed to an orbital basis), the non-interacting impurity Hamiltonian, and the impurity hybridization. Consequently, impurity solvers such as configuration interaction^{34,35} or multi-orbital hybridization expansion quantum Monte Carlo⁵⁹ can be employed. In contrast, the impurity problem in GW+DMFT contains an additional dependence on the frequency-dependent interactions W . Impurity solvers for GW+DMFT therefore have to be able to handle the frequency dependent interactions and work in an action representation. Only few such solvers exist,⁶⁰⁻⁶³ and they perform much better when hybridizations are diagonal and interactions are of the density-density type.⁶² Consequently, current implementations of GW+DMFT either neglect the frequency dependence altogether or simplify the interaction structure of the impurity problem to simple density-density terms.

A second major difference is the choice of orbitals, which is always local in GW+DMFT, whereas SEET employs the freedom of choosing orbitals that minimize the hybridization to effectively decouple multiple impurity problems.

Both of these methods are iterative self-consistent methods and are started by running an initial GW simulation. This initial solution may be far from the 'true' self-consistent solution. How SEET(GW/CI) or GW+DMFT recover if the initial weak coupling answer is very different from the self-consistent fixed point is therefore an important question, especially since both SEET and GW+DMFT are highly non-linear, iterative procedures in which the final self-consistent result may depend strongly on the starting point.

SEET, when performed in a natural orbital basis (see the sections on H_2 and N_2), is capable of recovering from a qualitatively very different GW solution within the inner iteration of the self-consistency and the outer loop

has a lesser influence. When SEET is carried out using a localized orbital basis, the outer loop is very significant and is necessary to reach quantitative results and smooth potential energy surfaces.⁹

It is an interesting open question how the recovery from an inaccurate starting point is achieved in GW+DMFT. Here, the GW solution does not just enter the weakly correlated part of the system but also determines the impurity interactions via W . The self-consistent iteration therefore needs to adjust both the interactions and the hybridizations (rather than just the hybridizations). It may be for this reason that non-causal physics was reported for GW+DMFT,¹⁴ whereas no non-causal self-energies or Green's functions have so far been observed in SEET.

Finally, the possibility to use the generalized SEET framework for treating intersecting subsets of orbitals leads to a systematically improvable method which provides an internal assessment of its accuracy. While a similar framework can be performed at the level of the Ψ functional, this has not yet been implemented.

V. CONCLUSIONS

In summary, the major points observed during the numerical tests performed with SEET(CI/GW) are as follows.

When SEET(CI/GW) is performed and all strongly correlated orbitals are placed in the impurity, a good agreement with standard quantum chemistry methods is achieved. We observe this in particular for the H_2 molecule in the TZ (Dunning) basis and for the N_2 molecule in the 6-31G basis. For N_2 both types of correlations, weak and strong, are recovered during the SEET procedure irrespective of whether GW or GF2 is used to treat weakly correlated electrons. In the equilibrium geometry, SEET agrees well with CCSD(T) and MP2. However, it avoids the divergence at the long distances, yielding curves parallel to NEVPT(6e,6o).

In the case of the hydrogen chains, we observe that when SEET-split is executed in a localized basis (SAO or LNO), both SEET with GW and GF2 lead to an overestimation of correlation energies for large inter-

atomic distances. SEET(CI/GW) leads to potential energy curves more parallel to the FCI answer than SEET(CI/GF2). However, when LNOs are employed in a larger basis, SEET(CI/GW) yields overestimated electronic energies for all distances, not just for large distances. The generalized version of the SEET functional with GF2, SEET(CI/GF2)-mix, can partially remedy this situation. In contrast, a similar improvement of SEET-split(CI/GW) by performing SEET(CI/GW)-mix is unlikely, since the overestimation of SEET-split does not only come from using non-intersecting orbital subspaces but also from the GW approximation itself.

When the energy basis (NO) is used in conjunction with larger basis sets, we observe that for large interatomic distances SEET-split with both GF2 and GW leads to an underestimation of the total electronic energy. For small intermolecular distances SEET(CI/GW) and GW lead to an overestimation of the correlation energy. As for the case of a localized basis, we observe that SEET-mix(CI/GF2) can help with recovering the correlations among strongly correlated orbitals and improve the results for the energies. SEET-mix(CI/GW) results in potential energy curves that are more parallel to the exact answer. However, it cannot repair the energy overestimation due to GW itself.

For SEET(CI/GW) performed in a larger basis sets, additional studies are necessary not only to assess how correlations inside the strongly correlated space are recovered but also how they are captured in the space of weakly correlated orbitals. In particular, the errors arising due to the overestimation of correlation effects among the weakly correlated orbitals when a large orbital basis is used are presently unclear. These questions require studies involving molecular cases where the strong correlations do not arise due to bond stretching but are intrinsic due to the presence of d - and f -electrons.

VI. ACKNOWLEDGEMENTS

D.Z. and T.N.L. acknowledge funding by DOE Grant no. ER16391. A.S. was supported by the NSF grant no. CHE-1453894. E. G. and J. L. are funded by the Simons collaboration on the many-electron problem.

* On leave from: Ho Chi Minh City Institute of Physics, VAST, Ho Chi Minh City, Vietnam

¹ A. Georges and G. Kotliar, Phys. Rev. B **45**, 6479 (1992).

² A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996).

³ G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti, Rev. Mod. Phys. **78**, 865 (2006).

⁴ S. Biermann, F. Aryasetiawan, and A. Georges, Phys. Rev. Lett. **90**, 086402 (2003).

⁵ S. Biermann, F. Aryasetiawan, and A. Georges, "Electronic structure of strongly correlated materials: Towards a first principles scheme," in *Physics of Spin in Solids: Materials, Methods and Applications*, edited by S. Halilov (Springer Netherlands, Dordrecht, 2005) pp. 43–65.

⁶ P. Werner and M. Casula, Journal of Physics: Condensed Matter **28**, 383001 (2016).

⁷ A. A. Kananenka, E. Gull, and D. Zgid, Phys. Rev. B **91**, 121111 (2015).

⁸ T. N. Lan, A. A. Kananenka, and D. Zgid, The Journal of Chemical Physics **143**, 241102 (2015),

- <http://dx.doi.org/10.1063/1.4938562>.
- ⁹ T. Nguyen Lan, A. A. Kananenka, and D. Zgid, *Journal of Chemical Theory and Computation* **12**, 4856 (2016).
 - ¹⁰ T. N. Lan and D. Zgid, *The Journal of Physical Chemistry Letters* **8**, 2200 (2017), pMID: 28453934, <http://dx.doi.org/10.1021/acs.jpcllett.7b00689>.
 - ¹¹ D. Zgid and E. Gull, *New Journal of Physics* **19**, 023047 (2017).
 - ¹² M. Motta, D. M. Ceperley, G. Kin-Lic Chan, J. A. Gomez, E. Gull, S. Guo, C. Jimenez-Hoyos, T. Nguyen 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, *ArXiv e-prints* (2017), arXiv:1705.01608 [physics.comp-ph].
 - ¹³ E. Kozik, M. Ferrero, and A. Georges, *Phys. Rev. Lett.* **114**, 156402 (2015).
 - ¹⁴ J. Lee and K. Haule, *Phys. Rev. B* **95**, 155104 (2017).
 - ¹⁵ J. M. Luttinger and J. C. Ward, *Physical Review* **118**, 1417 (1960).
 - ¹⁶ G. Baym and L. P. Kadanoff, *Phys. Rev.* **124**, 287 (1961).
 - ¹⁷ G. Baym, *Phys. Rev.* **127**, 1391 (1962).
 - ¹⁸ J. J. Phillips and D. Zgid, *The Journal of Chemical Physics* **140**, 241101 (2014).
 - ¹⁹ J. J. Phillips, A. A. Kananenka, and D. Zgid, *The Journal of Chemical Physics* **142**, 194108 (2015), <http://dx.doi.org/10.1063/1.4921259>.
 - ²⁰ A. A. Kananenka, J. J. Phillips, and D. Zgid, *Journal of Chemical Theory and Computation* **12**, 564 (2016), <http://dx.doi.org/10.1021/acs.jctc.5b00884>.
 - ²¹ A. A. Kananenka, A. R. Welden, T. N. Lan, E. Gull, and D. Zgid, *Journal of Chemical Theory and Computation* **12**, 2250 (2016), <http://dx.doi.org/10.1021/acs.jctc.6b00178>.
 - ²² A. A. Rusakov and D. Zgid, *The Journal of Chemical Physics* **144**, 054106 (2016), <http://dx.doi.org/10.1063/1.4940900>.
 - ²³ A. R. Welden, A. A. Rusakov, and D. Zgid, *The Journal of Chemical Physics* **145**, 204106 (2016).
 - ²⁴ N. E. Dahlen and R. van Leeuwen, *The Journal of Chemical Physics* **122**, 164102 (2005), <http://dx.doi.org/10.1063/1.1884965>.
 - ²⁵ N. E. Dahlen and U. v. Barth, *Phys. Rev. B* **69**, 195102 (2004).
 - ²⁶ N. E. Dahlen and U. von Barth, *The Journal of Chemical Physics* **120**, 6826 (2004).
 - ²⁷ H. Kajueter and G. Kotliar, *Phys. Rev. Lett.* **77**, 131 (1996).
 - ²⁸ L. Hedin, *Phys. Rev.* **139**, A796 (1965).
 - ²⁹ A. Stan, N. E. Dahlen, and R. van Leeuwen, *EPL (Europhysics Letters)* **76**, 298 (2006).
 - ³⁰ C.-O. Almbladh, U. V. Barth, and R. V. Leeuwen, *International Journal of Modern Physics B* **13**, 535 (1999).
 - ³¹ A. Stan, N. E. Dahlen, and R. van Leeuwen, *The Journal of Chemical Physics* **130**, 114105 (2009).
 - ³² H. Koch, A. S. de Meras, and T. B. Pedersen, *J. Chem. Phys.* **118**, 9481 (2003).
 - ³³ F. Caruso, P. Rinke, X. Ren, A. Rubio, and M. Scheffler, *Physical Review B* **88**, 075105 (2013).
 - ³⁴ D. Zgid and G. K.-L. Chan, *The Journal of Chemical Physics* **134**, 094115 (2011), <http://dx.doi.org/10.1063/1.3556707>.
 - ³⁵ D. Zgid, E. Gull, and G. K.-L. Chan, *Phys. Rev. B* **86**, 165128 (2012).
 - ³⁶ M. Karolak, G. Ulm, T. Wehling, V. Mazurenko, A. Poteryaev, and A. Lichtenstein, *Journal of Electron Spectroscopy and Related Phenomena* **181**, 11 (2010), proceedings of International Workshop on Strong Correlations and Angle-Resolved Photoemission Spectroscopy 2009.
 - ³⁷ C. Lin and A. A. Demkov, *Phys. Rev. B* **88**, 035123 (2013).
 - ³⁸ Y. Lu, M. Hoppner, O. Gunnarsson, and M. W. Haverkort, *Phys. Rev. B* **90**, 085102 (2014).
 - ³⁹ A. Go and A. J. Millis, *ArXiv e-prints* (2017), arXiv:1703.04928 [cond-mat.str-el].
 - ⁴⁰ A. A. Rusakov, J. J. Phillips, and D. Zgid, *The Journal of Chemical Physics* **141**, 194105 (2014), <http://dx.doi.org/10.1063/1.4901432>.
 - ⁴¹ N. Lin, C. A. Marianetti, A. J. Millis, and D. R. Reichman, *Phys. Rev. Lett.* **106**, 096402 (2011).
 - ⁴² Only a subset of v integrals has to be transformed to the new orthogonal basis.
 - ⁴³ F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2**, 73 (2012).
 - ⁴⁴ C. Angeli, R. Cimiraglia, and J.-P. Malrieu, *J. Chem. Phys.* **117**, 9138 (2002).
 - ⁴⁵ C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, and J.-P. Malrieu, *J. Chem. Phys.* **114**, 10252 (2001).
 - ⁴⁶ K. Aidas, C. Angeli, K. L. Bak, V. Bakken, R. Bast, L. Boman, O. Christiansen, R. Cimiraglia, S. Coriani, P. Dahle, E. K. Dalskov, U. Ekstrom, T. Enevoldsen, J. J. Erikssen, P. Ettenhuber, B. Fernandez, L. Ferrighi, H. Fliegl, L. Frediani, K. Hald, A. Halkier, C. Hattig, H. Heiberg, T. Helgaker, A. C. Hennum, H. Hettema, E. Hjertenaes, S. Host, I.-M. Hoyvik, M. F. Iozzi, B. Jansik, H. J. A. Jensen, D. Jonsson, P. Jorgensen, J. Kauczor, S. Kirpekar, T. Kjargaard, W. Klopper, S. Knecht, R. Kobayashi, H. Koch, J. Kongsted, A. Krapp, K. Kristensen, A. Ligabue, O. B. Lutnas, J. I. Melo, K. V. Mikkelsen, R. H. Myhre, C. Neiss, C. B. Nielsen, P. Norman, J. Olsen, J. M. H. Olsen, A. Osted, M. J. Packer, F. Pawlowski, T. B. Pedersen, P. F. Provasi, S. Reine, Z. Rinkevicius, T. A. Ruden, K. Ruud, V. V. Rybkin, P. Salek, C. C. M. Samson, A. S. de Meras, T. Saue, S. P. A. Sauer, B. Schimmelpennig, K. Sneskov, A. H. Steindal, K. O. Sylvester-Hvid, P. R. Taylor, A. M. Teale, E. I. Tellgren, D. P. Tew, A. J. Thorvaldsen, L. Thogersen, O. Vahtras, M. A. Watson, D. J. D. Wilson, M. Ziolkowski, and H. gren, *Wiley Interdisciplinary Reviews: Computational Molecular Science*, n/a (2013).
 - ⁴⁷ T. H. Dunning, *J. Chem. Phys.* **55**, 716 (1971).
 - ⁴⁸ J. M. Tomczak, M. Casula, T. Miyake, and S. Biermann, *Phys. Rev. B* **90**, 165138 (2014).
 - ⁴⁹ T. Ayral, S. Biermann, and P. Werner, *Phys. Rev. B* **87**, 125149 (2013).
 - ⁵⁰ L. Boehnke, F. Nilsson, F. Aryasetiawan, and P. Werner, *Phys. Rev. B* **94**, 201106 (2016).
 - ⁵¹ W. J. Hehre, R. F. Stewart, and J. A. Pople, *J. Chem. Phys.* **51**, 2657 (1969).
 - ⁵² T. H. Dunning, *J. Chem. Phys.* **90**, 1007 (1989).
 - ⁵³ Nonparallelity errors (NPEs) are defined as the difference between the largest and smallest errors with respect to the reference data. Methods with small NPEs may not give an exact total energy but result in many useful data since the shift in total energy is then only a constant shift.
 - ⁵⁴ S. F. Boys, *Rev. Mod. Phys.* **32**, 296 (1960).
 - ⁵⁵ J. Hachmann, W. Cardoen, and G. K.-L. Chan, *J. Chem. Phys.* **125**, 144101 (2006).
 - ⁵⁶ A. Y. Sokolov and G. K.-L. Chan, *The Journal of Chemical Physics* **142**, 124107 (2015),

- <http://dx.doi.org/10.1063/1.4916315>.
- ⁵⁷ “See supplemental material at [url will be inserted by publisher] for occupation numbers and orbital energies at different interatomic distances of the nitrogen dimer in the 6-31g basis set.”
- ⁵⁸ S. Biermann, *Journal of Physics: Condensed Matter* **26**, 173202 (2014).
- ⁵⁹ P. Werner and A. J. Millis, *Phys. Rev. B* **74**, 155107 (2006).
- ⁶⁰ F. F. Assaad and T. C. Lang, *Phys. Rev. B* **76**, 035116 (2007).
- ⁶¹ P. Werner and A. J. Millis, *Phys. Rev. Lett.* **99**, 146404 (2007).
- ⁶² P. Werner and A. J. Millis, *Phys. Rev. Lett.* **104**, 146401 (2010).
- ⁶³ E. Gull, A. J. Millis, A. I. Lichtenstein, A. N. Rubtsov, M. Troyer, and P. Werner, *Rev. Mod. Phys.* **83**, 349 (2011).