

# On the Relation between Equation-of-Motion Coupled-Cluster Theory and the *GW* Approximation

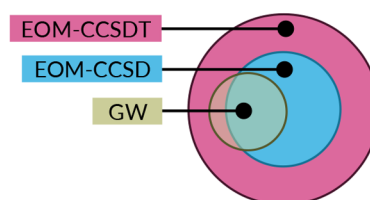
Malte F. Lange and Timothy C. Berkelbach\*<sup>✉</sup>

Department of Chemistry and James Franck Institute, University of Chicago, Chicago, Illinois 60637, United States

**S** Supporting Information

**ABSTRACT:** We discuss the analytic and diagrammatic structure of ionization potential (IP) and electron affinity (EA) equation-of-motion coupled-cluster (EOM-CC) theory, in order to put it on equal footing with the prevalent *GW* approximation. The comparison is most straightforward for the time-ordered one-particle Green's function, and we show that the Green's function calculated by EOM-CC with single and double excitations (EOM-CCSD) includes fewer ring diagrams at higher order than does the *GW* approximation, due to the former's unbalanced treatment of time-ordering. However, the EOM-CCSD Green's function contains a large number of vertex corrections, including ladder diagrams, mixed ring-ladder diagrams, and exchange diagrams. By including triple excitations, the EOM-CCSDT Green's function includes all diagrams contained in the *GW* approximation, along with many high-order vertex corrections. In the same language, we discuss a number of common approximations to the EOM-CCSD equations, many of which can be classified as elimination of diagrams. Finally, we present numerical results by calculating the principal charged excitations energies of the molecules contained in the so-called *GW100* test set [*J. Chem. Theory Comput.* 2015, 11, 5665–5687]. We argue that (in molecules) exchange is as important as screening, advocating for a Hartree–Fock reference and second-order exchange in the self-energy.

Number of diagrams included in the one-particle Green's function



## 1. INTRODUCTION

The accurate calculation of excited-state properties constitutes one of the major challenges in modern computational materials science. For charged excitations, namely, the ionization potentials and electron affinities as measured by photoelectron spectroscopy, the *GW* approximation has proven to be a powerful and successful tool in the condensed phase. Formally, the *GW* approximation (reviewed below) arises as the lowest-order self-energy diagram when the one-particle Green's function  $G$  is expanded in terms of the screened Coulomb interaction  $W$ ,<sup>1,2</sup> with screening treated in the random-phase approximation. Neglected diagrams can be assigned to vertex corrections (appearing in both the self-energy and the polarization propagator), which are a natural target for post-*GW* theories and an ongoing area of activity.<sup>3–7</sup>

In contrast to time-dependent Green's function-based diagrammatic theories, wave function-based theories and concomitant time-independent perturbation theory offer an alternative route toward systematically improvable excited-state calculations.<sup>8</sup> The great variety of wave function ansatzes, combined with the long history of development and benchmarking in the molecular quantum chemistry community, makes such approaches particularly promising. Unfortunately, the formal comparison between wave function-based and Green's function-based techniques is complicated by a difference in both the approach and the language. Here, we present such a comparison, by analyzing the one-particle Green's function calculated using equation-of-motion coupled-cluster

theory to that calculated using the *GW* approximation. A relation between the two can be anticipated on the basis of the known exact relation between total *ground-state* energies calculated using the ring coupled-cluster doubles approach and using the random-phase approximation,<sup>9–11</sup> the latter of which is at the heart of screening in the *GW* approximation. However, for charged excitation energies, the equivalence is not so straightforward.

Although this article focuses on the connections between the *GW* approximation and EOM-CC theory because of their shared connections to the random-phase approximation,<sup>9–11</sup> we emphasize that a number of other Green's function techniques have been developed and used successfully on molecular problems.<sup>12,13</sup> A nonexhaustive list of the most popular approaches include those that are essentially finite-order [such as the outer-valence Green's function approach,<sup>12</sup> the partial third-order (P3) approach,<sup>14</sup> and its renormalized variant (P3+)<sup>15</sup>] and those that are rigorously infinite-order [such as the two-particle-hole Tamm-Dancoff approximation,<sup>16</sup> its extended variant,<sup>17</sup> the algebraic diagrammatic construction (ADC),<sup>18</sup> and the nondiagonal renormalized approach exact to second order (NR2)].<sup>19</sup> Recently, there has been renewed interest in the second-order Green's function,<sup>20</sup> especially in its self-consistent<sup>21,22</sup> and finite-temperature<sup>23</sup> variations. We emphasize that many of the above methods exhibit an attractive

Received: May 10, 2018

Published: July 20, 2018

$N^5$  scaling, whereas the *GW* approximation and EOM-CCSD both exhibit  $N^6$  scaling (see below), which we identify as the cost needed for a rigorous treatment of RPA physics.

In wave function-based techniques, ionization potentials and electron affinities can be calculated either as a difference in ground-state energies (between the neutral and ionic systems) or via the equation-of-motion framework, which directly results in an eigensystem whose eigenvalues are the ionization potentials or electron affinities. Equation-of-motion coupled-cluster (EOM-CC) theory is one such framework, which typically achieves accurate excitation energies when performed with single and double excitations (EOM-CCSD).<sup>24,25</sup> At the intersection of these methods, Nooijen and Snijders derived a one-particle Green's function in the CC framework,<sup>26,27</sup> the poles and residues of which are precisely those of the conventional EOM-CC formalism (in the bivariational framework). The CC Green's function has seen a renewed interest in recent years.<sup>28–34</sup> One of the main goals of the present work is to relate the latter theory to the *GW* approximation, which is carried out in section 3.

A number of numerical comparisons between Green's function-based and wave function-based techniques for charged excitation energies have been carried out in recent years. In particular, comparisons between the *GW* approximation and wave function-based techniques have been performed for one-dimensional lattice models,<sup>35</sup> for a test set of 24 organic acceptor molecules,<sup>36</sup> for oligoacenes,<sup>37</sup> and for a test set of 100 molecules;<sup>38,39</sup> the latter test set is known as the *GW100*, introduced in ref 40, and forms the basis of our numerical study in section 4.2.

In light of recent efforts to bring the systematic improvability of wave function-based theories into the solid state,<sup>28,41–49</sup> we believe it timely to establish the relationship, both formally and numerically, between popular wave function approaches and Green's function approaches, the latter of which has dominated solid-state electronic structure. Future work in both Green's function-based and wave function-based approaches can benefit from the analysis and results of the present work.

The layout of this article is as follows. In section 2, we provide the requisite theoretical background associated with general features of the one-particle Green's function, the *GW* approximation to the self-energy, and equation-of-motion coupled-cluster theory. In section 3, we perform a detailed diagrammatic comparison of the two methods, comparing separately their Green's functions and self-energies. In section 4.2, we use equation-of-motion coupled-cluster theory to calculate ionizations potentials and electron affinities of the *GW100* test set, and evaluate a number of accurate but efficient approximations, which are straightforwardly analyzed with the previously introduced diagrammatic description. In section 5, we conclude with an outlook for future developments.

## 2. THEORY

**2.1. One-Particle Green's Function.** The one-particle time-ordered Green's function is defined by<sup>12,50</sup>

$$iG_{pq}(\omega) = \int d(t_1-t_2) e^{i\omega(t_1-t_2)} \langle \Psi_0 | T[\hat{a}_p(t_1) \hat{a}_q^\dagger(t_2)] | \Psi_0 \rangle \quad (1)$$

where  $T$  is the time-ordering operator,  $\Psi_0$  is the exact interacting ground state, and  $p, q$  index a complete set of single-particle spin-orbitals. The irreducible self-energy matrix  $\Sigma(\omega)$  satisfies the relation  $\mathbf{G}(\omega) = [\omega - \mathbf{f} - \Sigma(\omega)]^{-1}$ , where  $\mathbf{f}$  is the matrix

associated with some one-body (mean-field) operator such that  $\Sigma$  contains the remaining effects of the electronic interactions. Although in practice,  $\mathbf{f}$  is commonly the Kohn–Sham matrix of density functional theory, here we consider it to be the Fock matrix and will let  $p, q, r, s$  index the canonical Hartree–Fock (HF) orbitals, such that  $\mathbf{f}$  is diagonal:  $f_{pq} = \varepsilon_p \delta_{pq}$ . Following convention, indices  $i, j, k, l$  are used for the  $n_{\text{occ}}$  occupied orbitals in the HF determinant and  $a, b, c, d$  for the  $n_{\text{vir}}$  virtual (unoccupied) orbitals; in total, there are  $M = n_{\text{occ}} + n_{\text{vir}}$  orbitals.

**2.2. *GW* Approximation.** The charged excitation energies (ionization potentials and electron affinities) occur at the poles of the Green's function; i.e., they are the self-consistent eigenvalues of a frequency-dependent one-particle matrix  $\mathbf{H}(\omega)$ :

$$\sum_q H_{pq}(\omega=E_n) R_q^n = E_n R_p^n \quad (2)$$

with

$$H_{pq}(\omega) = \omega \delta_{pq} - [\mathbf{G}(\omega)]_{pq}^{-1} = \varepsilon_p \delta_{pq} + \Sigma_{pq}(\omega) \quad (3)$$

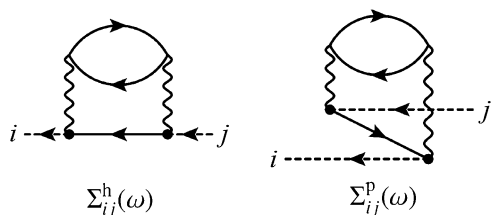
In the *GW* approximation,<sup>1</sup> the self-energy is given by

$$\Sigma(\mathbf{r}_1, \mathbf{r}_2; \omega) = \frac{i}{2\pi} \int d\omega' e^{i\eta\omega'} G(\mathbf{r}_1, \mathbf{r}_2; \omega + \omega') \times W_c(\mathbf{r}_2, \mathbf{r}_1; \omega') \quad (4)$$

where  $W_c = W - v$  is the correlation part of the screened interaction; recall that the bare exchange term has been included (self-consistently) in the Fock operator  $\mathbf{f}$ . The dielectric function that screens the Coulomb interaction is evaluated with the random-phase approximation (RPA), corresponding to a resummation of all ring diagrams contributing to the polarization propagator; furthermore, all vertex corrections are neglected. Henceforth, we limit the discussion to the nonself-consistent  $G_0W_0$  approximation, where  $G_0$  and  $W_0$  are evaluated in a “one-shot” manner using the orbitals and orbital energies of the mean-field problem (in this case, HF). In a finite single-particle basis set, the frequency integration can be done analytically to show that the self-energy has separate hole (h) and particle (p) contributions,<sup>51–53</sup>

$$\begin{aligned} \Sigma_{pq}(\omega) &= \Sigma_{pq}^h(\omega) + \Sigma_{pq}^p(\omega) \\ &= \sum_\nu \left\{ \sum_k \left[ \frac{[M_{kp}^\nu]^* M_{kq}^\nu}{\omega - (\varepsilon_k - \Omega_\nu) - i\eta} \right] \right. \\ &\quad \left. + \sum_c \left[ \frac{[M_{cp}^\nu]^* M_{cq}^\nu}{\omega - (\varepsilon_c + \Omega_\nu) + i\eta} \right] \right\} \quad (5) \end{aligned}$$

which are associated with the two possible time-orderings (Goldstone diagrams) of the corresponding Feynman diagram for the self-energy, i.e.,  $\Sigma_{pq}(t_2 - t_1)$  with  $t_2 > t_1$  or with  $t_2 < t_1$ . Expressed in terms of time-ordered Goldstone diagrams, the lowest-order ring diagrams appearing in the hole and particle contributions to the *GW* self-energy are shown in Figure 1, for the single-particle indices  $i, j$  in the occupied orbital subspace. The poles of the *GW* self-energy occur at  $\omega = \varepsilon_k - \Omega_\nu$  and  $\omega = \varepsilon_c + \Omega_\nu$ , i.e., at sums and differences of the orbital energies  $\varepsilon_p$  and neutral excitation energies  $\Omega_\nu$ . For future reference, we note that, in gapped molecules and materials, the particle contribution to the self-energy of hole states is only weakly dependent on frequency, because the quasiparticle energy  $\omega \approx \varepsilon_k$  is far from the poles at  $\varepsilon_c + \Omega_\nu$ ; the separation is roughly twice the gap.



**Figure 1.** Lowest-order time-ordered (Goldstone) ring diagrams appearing in the hole (h) and particle (p) contributions to the self-energy, for occupied orbitals  $ij$ . The dashed lines only serve to indicate the connectivity in a Goldstone diagram for the Green's function.

The transition amplitudes associated with the poles of the self-energy are given by

$$M_{pq}^\nu = \int d\mathbf{x}_1 d\mathbf{x}_2 \rho_\nu(\mathbf{x}_1) r_{12}^{-1} \phi_p^*(\mathbf{x}_2) \phi_q(\mathbf{x}_2) \quad (6)$$

where

$$\rho_\nu(\mathbf{x}) = \langle \Psi_0 | \hat{n}(\mathbf{x}) | \Psi_\nu \rangle = \sum_{pq} \phi_p^*(\mathbf{x}) \phi_q(\mathbf{x}) \langle \Psi_0 | \hat{a}_p^\dagger \hat{a}_q | \Psi_\nu \rangle \quad (7)$$

is the transition density of the neutral excited state  $\Psi_\nu$ .

The level of theory used to construct the polarizability determines the energies  $\Omega_\nu$  and wave functions  $\Psi_\nu$  entering in the above equations. Using any time-dependent mean-field response,  $|\Psi_\nu\rangle = \sum_{ai} [X_{ai}^\nu \hat{a}_a^\dagger \hat{a}_i - Y_{ai}^\nu \hat{a}_i^\dagger \hat{a}_a] |\Psi_0\rangle$ , leads to an eigensystem commonly associated with the RPA,

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{B}^* & -\mathbf{A}^* \end{pmatrix} \begin{pmatrix} \mathbf{X}^\nu \\ \mathbf{Y}^\nu \end{pmatrix} = \Omega_\nu \begin{pmatrix} \mathbf{X}^\nu \\ \mathbf{Y}^\nu \end{pmatrix} \quad (8)$$

and transition amplitudes

$$M_{pq}^\nu = \sum_{ai} [X_{ai}^\nu \langle ip|aq \rangle + Y_{ai}^\nu \langle apliq \rangle] \quad (9)$$

where the two-electron integrals are given by  $\langle pq|rs \rangle \equiv \int d\mathbf{x}_1 d\mathbf{x}_2 \phi_p^*(\mathbf{x}_1) \phi_q^*(\mathbf{x}_2) r_{12}^{-1} \phi_r(\mathbf{x}_1) \phi_s(\mathbf{x}_2)$  and  $x$  is a combined spin and spatial variable.

Specifically using time-dependent HF theory, the  $\mathbf{A}$  and  $\mathbf{B}$  matrices (each of dimension  $n_{\text{occ}} n_{\text{vir}} \times n_{\text{occ}} n_{\text{vir}}$ ) have elements

$$A_{ai,bj} = (\epsilon_a - \epsilon_i) \delta_{ab} \delta_{ij} + \langle aj||ib \rangle \quad (10a)$$

$$B_{ai,bj} = \langle ij||ab \rangle \quad (10b)$$

where the antisymmetrized two-electron integrals are  $\langle pq||rs \rangle \equiv \langle pq|rs \rangle - \langle pqls \rangle$ . Using the more conventional time-dependent Hartree dielectric function yields the same structure but neglects the exchange integrals in the  $\mathbf{A}$  and  $\mathbf{B}$  matrices; this corresponds to the common version of the RPA and the one used in the  $GW$  approximation. The RPA eigenvalues come in positive- and negative-energy pairs, comprising only  $n_{\text{occ}} n_{\text{vir}}$  distinct eigenvalues; thus there are  $M n_{\text{occ}} n_{\text{vir}}$  poles in the  $GW$  self-energy.

In the form given here, the solution of the RPA eigenvalue problem in eq 8 highlights the canonical  $N^6$  scaling of the  $GW$  approximation,<sup>52,54</sup> which is identical to that of EOM-CCSD. This  $GW$  scaling comes from the need to calculate *all* RPA eigenvalues in order to reliably calculate just one quasiparticle energy in eq 5. Alternative formulations can reduce this scaling.

### 2.3. Equation-of-Motion Coupled-Cluster Theory.

Equation-of-motion coupled-cluster theories start from the ground-state CC wave function,  $|\Psi\rangle = e^T |\Phi\rangle$ , where the cluster

operator  $\hat{T}$  creates neutral excitations with respect to the reference determinant  $|\Phi\rangle$ ,

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \dots = \sum_{ai} t_i^a \hat{a}_a^\dagger \hat{a}_i + \frac{1}{4} \sum_{abij} t_{ij}^{ab} \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i + \dots \quad (11)$$

The ground-state energy and cluster amplitudes are determined by the conditions

$$E_{\text{CC}} = \langle \Phi | \bar{H} | \Phi \rangle \quad (12a)$$

$$0 = \langle \Phi_i^a | \bar{H} | \Phi \rangle \quad (12b)$$

$$0 = \langle \Phi_{ij}^{ab} | \bar{H} | \Phi \rangle \quad (12c)$$

and so on, where  $|\Phi_i^a\rangle = \hat{a}_i^\dagger \hat{a}_i |\Phi\rangle$ , etc. and  $\bar{H} \equiv e^{-\hat{T}} \hat{H} e^{\hat{T}}$  is a similarity-transformed Hamiltonian. As seen above, the reference determinant is the right-hand eigenvector of  $\bar{H}$ . Because  $\bar{H}$  is non-Hermitian, it has distinct left-hand and right-hand eigenvectors for each eigenvalue; for the ground state, the left-hand eigenvector of  $\hat{H}$  is given by

$$\langle \tilde{\Psi}_0 | = \langle \Phi | (1 + \hat{\Lambda}) e^{-\hat{T}} \quad (13)$$

$$\hat{\Lambda} = \hat{\Lambda}_1 + \hat{\Lambda}_2 + \dots = \sum_{ai} \lambda_a^i \hat{a}_i^\dagger \hat{a}_a + \frac{1}{4} \sum_{abij} \lambda_{ab}^{ij} \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a + \dots \quad (14)$$

Charged excitation energies in EOM-CC are calculated as eigenvalues of  $\bar{H}$  in a finite basis of  $(N \pm 1)$ -electron Slater determinants;  $(N - 1)$ -electron excitation energies are calculated via the ionization potential (IP) framework and  $(N + 1)$ -electron excitation energies via the electron affinity (EA) framework.<sup>24–27</sup> For example, the IP-EOM-CC energies are determined by

$$(E_{\text{CC}} - \bar{H}) \hat{R}^{N-1}(n) |\Phi\rangle = \Omega_n^{N-1} \hat{R}^{N-1}(n) |\Phi\rangle \quad (15)$$

$$\begin{aligned} \hat{R}^{N-1}(n) &= \hat{R}_1^{N-1}(n) + \hat{R}_2^{N-1}(n) + \dots \\ &= \sum_i r_i(n) \hat{a}_i + \frac{1}{2} \sum_{aij} r_{ij}^a(n) \hat{a}_a^\dagger \hat{a}_j \hat{a}_i + \dots \end{aligned} \quad (16)$$

where  $\Omega_n^{N-1} = E_0^N - E_n^{N-1}$  is the negative of a many-body ionization potential and corresponds to an exact pole of the one-particle Green's function. Again,  $\bar{H}$  has distinct left-hand eigenvectors,

$$\langle \Phi | \hat{L}^{N-1}(n) (E_{\text{CC}} - \bar{H}) = \langle \Phi | \hat{L}^{N-1}(n) \Omega_n^{N-1} \quad (17)$$

$$\begin{aligned} \hat{L}^{N-1}(n) &= \hat{L}_1^{N-1}(n) + \hat{L}_2^{N-1}(n) + \dots \\ &= \sum_a l^a(n) \hat{a}_a^\dagger + \frac{1}{2} \sum_{abi} l_{abi}^{ab}(n) \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_i + \dots \end{aligned} \quad (18)$$

The left-hand and right-hand eigenstates of the untransformed  $\hat{H}$  are then given by

$$|\Psi_n^{N-1}\rangle = e^{\hat{T}} \hat{R}^{N-1}(n) |\Phi\rangle \quad (19)$$

$$\langle \tilde{\Psi}_n^{N-1} | = \langle \Phi | \hat{L}^{N-1}(n) e^{-\hat{T}} \quad (20)$$

and form a biorthogonal set. With appropriate normalization, the eigenstates yield a resolution-of-the-identity in the  $(N \pm 1)$ -electron space

$$\begin{aligned}
 1 &= \sum_n |\Psi_n^{N\pm 1}\rangle \langle \tilde{\Psi}_n^{N\pm 1}| \\
 &= \sum_n e^{\hat{T}\hat{R}^{N\pm 1}}(n) |\Phi\rangle \langle \Phi \hat{L}^{N\pm 1}(n) e^{-\hat{T}}
 \end{aligned}
 \quad (21)$$

As first done by Nooijen and Snijders,<sup>26,27</sup> this enables an algebraic Lehmann representation of the Green's function, which (as usual) separates into IP and EA contributions due to the time-ordering operator,  $G_{pq}(\omega) = G_{pq}^{\text{IP}}(\omega) + G_{pq}^{\text{EA}}(\omega)$ . For example, the IP part is given by

$$G_{pq}^{\text{IP}}(\omega) = \sum_n \frac{\tilde{\psi}_q(n) \psi_p(n)}{\omega - \Omega_n^{N-1} + i\eta}
 \quad (22a)$$

$$\tilde{\psi}_q(n) = \langle \Phi | (1 + \hat{\Lambda}) e^{-\hat{T}} \hat{a}_q^\dagger e^{\hat{T}} \hat{R}^{N-1}(n) | \Phi \rangle
 \quad (22b)$$

$$\psi_p(n) = \langle \Phi | \hat{L}^{N-1}(n) e^{-\hat{T}} \hat{a}_p e^{\hat{T}} | \Phi \rangle
 \quad (22c)$$

Using conventional many-body techniques for the  $\hat{T}$ ,  $\hat{R}$ , and  $\hat{\Lambda}$  operators enables separate diagrammatic expansions of the IP and EA contributions to the Green's function,<sup>26,27</sup> which is properly size extensive as a sum of connected diagrams. In particular, using IP-EOM-CCSD, the IP Green's function is given as the sum over all time-ordered (Goldstone) diagrams for which cutting the diagram after each end point or vertex always leaves a sum of disconnected diagrams at previous times, each of which has no more than two electron and two hole open propagator lines. On the basis of the outcome of this procedure, components of each diagram can be classified as belonging to the cluster operators  $\hat{T}_m$ , the EOM operators  $\hat{R}_m$ , or the de-excitation operators  $\hat{\Lambda}_m$  (in refs 26 and 27, these operators are designated more precisely as  $\hat{T}_m^{(p)}$ ,  $\hat{S}_m^{(p)}$  ( $\omega$ ), and  $\hat{R}_m^{(pq)}$  ( $\omega$ ), respectively). Importantly in this construction, at each order in perturbation theory, all time-orderings of a given Feynman diagram are not included.

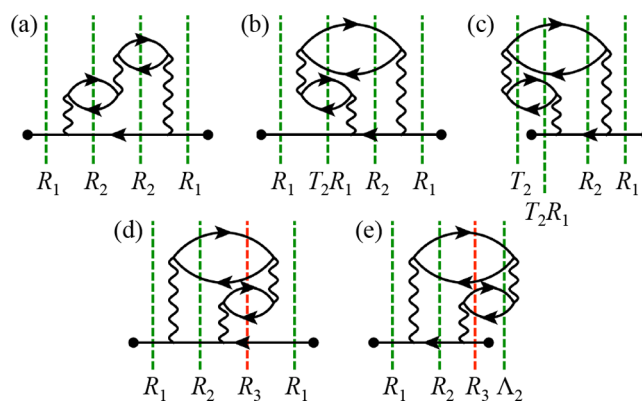
### 3. COMPARING THE GW APPROXIMATION AND EOM-CC THEORY

**3.1. Comparing the Green's Function.** We first compare the time-ordered Goldstone diagrams appearing in the Green's function of the GW approximation and EOM-CC theory. By construction, the first-order terms in the Green's function are vanishing. At second order, there are ten Feynman diagrams arising from six diagrams for the proper self-energy, only two of which are not accounted for by a self-consistent HF calculation. The GW Green's function includes only one of these two diagrams, with a single ring, which translates to  $4! = 24$  Goldstone diagrams, all of which are included in the EOM-CCSD Green's function. However, the EOM-CCSD Green's function also includes all Goldstone diagrams associated with the second-order exchange diagram (another 24 Goldstone diagrams). Therefore, as is well-known, the EOM-CCSD Green's function is correct through second order, and thus exact for two-electron problems; the GW Green's function is not.

At third order, the comparison is more complicated. Again, the GW Green's function contains one irreducible Feynman diagram, which includes two rings leading to  $5! = 120$  Goldstone diagrams. For simplicity of analysis, we focus on IP diagrams ( $t_1 < t_2$ ) in the occupied orbital subspace,  $G_{ij}^{\text{IP}}(t_1, t_2)$ , generated by the hole part of the self-energy, i.e.,  $\Sigma_{ij}(t_1, t_2)$  with  $t_1 < t_2$ . The GW approximation produces five such Goldstone diagrams (Figure

2), only three of which are included in the EOM-CCSD Green's function. When cut after the second interaction, diagrams (d) and (e) produce, at earlier times, a connected diagram in the 3-hole+2-particle (3h2p) space, which is included in the EOM-CCSDT Green's function, but not the EOM-CCSD one. In contrast, diagrams (b) and (c) also have a 3h2p configuration, but one that is generated by the disconnected product of the  $T_2$  [2-hole+2-particle (2h2p)] and  $R_1$  [1-hole (1h)] operators. Therefore, some of the non-TDA GW diagrams are included in the EOM-CCSD Green's function, but not all of them.

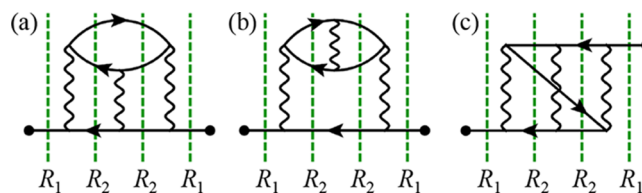
The above analysis is straightforward to generalize to higher order, and we find that the irreducible part of the EOM-CCSD



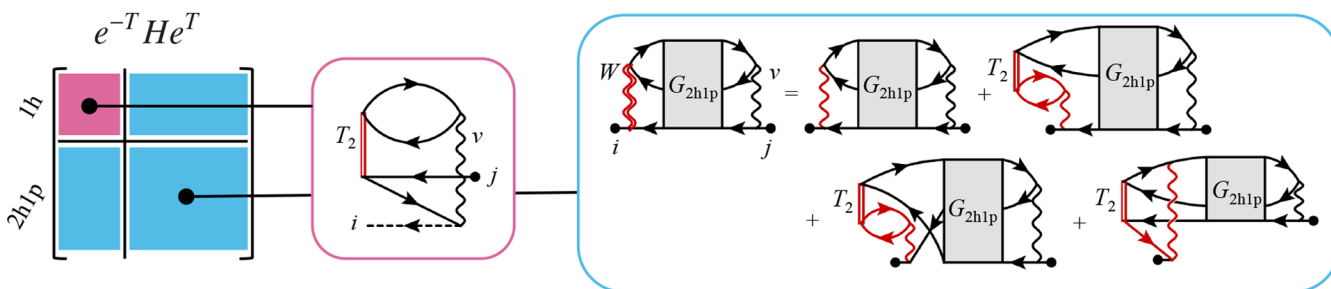
**Figure 2.** Only five third-order Goldstone diagrams contributing to  $G_{ij}^{\text{IP}}$  originating from  $\Sigma_{ij}(t_1, t_2)$  with  $t_1 < t_2$ , in the GW approximation. Only (a), (b), and (c) are included in the EOM-CCSD Green's function; all five are included in the EOM-CCSDT Green's function. Time increases from left to right.

Green's function at  $n$ th order contains only a vanishing fraction of the ring diagrams included in the GW Green's function; at  $n$ th order, the fraction of diagrams included is  $O(1/n)$ . (Of course, a large number of reducible Green's function diagrams are included at  $n$ th order, due to combinations of low-order diagrams.) In spite of this apparent flaw of EOM-CCSD theory, we emphasize that the EOM-CCSD Green's function contains many other nonring diagrams that are not contained in the GW approximation. For example, three third-order diagrams corresponding to various particle–particle and particle–hole ladders are shown in Figure 3. Diagrams (a) and (c) include vertex corrections to the self-energy and diagram (b) includes a vertex correction to the polarization propagator.

The behavior we have described should be compared to the enumeration of Goldstone diagrams for the correlation energy (the vacuum amplitude): in this context, ground-state CCSD includes ring diagrams with all possible time-orderings, completely



**Figure 3.** Three example third-order Goldstone diagrams contributing to the IP part of the EOM-CCSD Green's function with  $t_1 < t_2$ , which are not included in the GW Green's function. All diagrams shown are generated by the EOM (2h1p) formalism, independent of coupled-cluster theory. Time increases from left to right.



**Figure 4.** Goldstone self-energy diagrams arising from the different sectors of the IP-EOM-CCSD similarity-transformed Hamiltonian. The  $\hat{T}_2$ -transformed Fock operator in the 1h space generates a frequency-independent screened-exchange self-energy  $\tilde{\Sigma}_{ij}^p$ . The  $\hat{T}_2$ -transformed repulsion integral  $W_{iab}$  which couples the 1h and 2h1p spaces, generates non-Tamm–Dancoff screening on top of the Tamm–Dancoff ring and ladder diagrams generated in the 2h1p space; this leads to a frequency-dependent screened-exchange self-energy  $\tilde{\Sigma}_{ij}^h(\omega)$ .

encompassing those diagrams contained in the RPA.<sup>9,11</sup> Similarly, one of us (T.C.B.) has recently shown that the frequency-dependent polarizability calculated with neutral-excitation EOM-CCSD encompasses all diagrams contained in the (dynamical) RPA.<sup>55,56</sup> The difference observed here for the one-particle Green’s function can be traced to the need for the EOM-CC operators to simultaneously describe screening and free-particle propagation, as exemplified in diagrams (d) and (e) in Figure 2. For exactly this reason, the CCSD correlation energy is *not recovered* from the EOM-CCSD Green’s function, as discussed by Nooijen and Snijders,<sup>27</sup>

$$E_{\text{CCSD}} \neq \frac{1}{4\pi i} \int d\omega \text{Tr}\{(\mathbf{h} + \omega \mathbf{1}) \mathbf{G}_{\text{CCSD}}^{\text{IP}}(\omega)\} \quad (23)$$

Because of eq 23, there is no EOM-ring-CCD Green’s function that produces the RPA correlation energy. The above can roughly be viewed as a reminder that although the CCSD energy is exact to third order in perturbation theory, the EOM-CCSD energies are only exact to second order. However, the EOM-CCSD Green’s function does yield the CCSD reduced density matrix and is thus properly number-conserving,

$$\rho_{\text{CCSD}} = \frac{1}{2\pi i} \int d\omega [\mathbf{G}_{\text{CCSD}}^{\text{IP}}(\omega)] \quad (24)$$

$$N = \text{Tr}\{\rho_{\text{CCSD}}\} \quad (25)$$

Therefore, despite the error in the individual poles of the EOM-CCSD Green’s function, some “sum rules” are satisfied.

In this section, we have compared the Green’s functions generated by EOM-CC and the *GW* approximation. A more direct connection with the *GW* approximation and related time-dependent diagrammatic methods can be made by directly targeting an EOM-CC self-energy or polarization propagator; work along these lines is currently in progress in our group. However, an *approximate* algebraic self-energy can be worked out directly from the EOM-CC eigenvalue problem, which we turn to next.

**3.2. Comparing the Self-Energy.** For the remainder of the article, we will only consider IP-EOM-CCSD; the results for EA-EOM-CCSD are completely analogous. We introduce the normal-ordered Hamiltonian, with respect to the HF reference,  $\hat{H}_N \equiv \hat{H} - E_{\text{HF}}$  and its similarity-transformed variant  $\bar{H}_N \equiv \bar{H} - E_{\text{CC}}$ . The linear eigenvalue problem in eqs 15 and 16 clearly leads to a (schematic) matrix representation

$$\Delta \bar{H}_{\text{CC}} = - \begin{pmatrix} \langle \Phi_i | \bar{H}_N | \Phi_k \rangle & \langle \Phi_i | \bar{H}_N | \Phi_{kl}^b \rangle \\ \langle \Phi_{ij}^a | \bar{H}_N | \Phi_k \rangle & \langle \Phi_{ij}^a | \bar{H}_N | \Phi_{kl}^b \rangle \end{pmatrix} \quad (26)$$

In this section, we will show that the *GW* excitation energies are closely related to the eigenvalues of the approximated matrix

$$\Delta \bar{H}_{\text{CCGW}} = - \begin{pmatrix} \langle \Phi_i | e^{-\hat{T}_2} \hat{f}_N e^{\hat{T}_2} | \Phi_k \rangle & \langle \Phi_i | \bar{H}_N | \Phi_{kl}^b \rangle \\ \langle \Phi_{ij}^a | \bar{H}_N | \Phi_k \rangle & \langle \Phi_{ij}^a | \bar{H}_N | \Phi_{kl}^b \rangle \end{pmatrix} \quad (27)$$

where  $\hat{f}_N$  is the normal-ordered Fock operator,  $\hat{T}_1 = 0$  everywhere, the  $\hat{T}_2$  amplitudes satisfy an approximate version of eq 12c known as “ring-CCD”,<sup>9,11</sup> and the untransformed Hamiltonian ( $\hat{T}_2 = 0$ ) is used in the doubles–doubles block.

If desired, antisymmetrization can further be removed from most two-electron integrals leading to the use of “direct ring-CCD”,<sup>9,11</sup> in the one-hole space combined with a more conventional TDH treatment of screening in the two-hole +one-particle. However, we keep antisymmetrization throughout, which makes the theory manifestly self-interaction free, while retaining only the essential ingredients of the *GW* approximation.

Using a Löwdin partitioning,<sup>57</sup> the eigenvalues of the Hamiltonian in eq 27 can be found self-consistently for the frequency-dependent matrix

$$\begin{aligned} A_{ij}(\omega) &= -\langle \Phi_i | e^{-\hat{T}_2} \hat{f}_N e^{\hat{T}_2} | \Phi_j \rangle \\ &+ \frac{1}{4} \sum_{abklmn} \langle \Phi_i | \bar{H} | \Phi_{kl}^a \rangle [G_{2h1p}]_{klmn}^{ab}(\omega) \langle \Phi_{mn}^b | \bar{H} | \Phi_j \rangle \\ &= \varepsilon_i \delta_{ij} + \tilde{\Sigma}_{ij}^p + \tilde{\Sigma}_{ij}^h(\omega) \end{aligned} \quad (28)$$

where  $G_{2h1p}(\omega)$  is a specific time sequence of the three-particle Green’s function,

$$\begin{aligned} [G_{2h1p}]_{klmn}^{ab}(\omega) &= -i \int dt e^{i\omega t} \langle \Phi_0 | [\hat{a}_a^\dagger \hat{a}_k \hat{a}_l](0) [\hat{a}_b^\dagger \hat{a}_m \hat{a}_n](t) | \Phi_0 \rangle \\ &= \langle \Phi_0 | \hat{a}_a^\dagger \hat{a}_k \hat{a}_l \frac{1}{\omega - (-\mathcal{P}\hat{H}_N\mathcal{P})} \hat{a}_b^\dagger \hat{a}_m \hat{a}_n | \Phi_0 \rangle \end{aligned} \quad (29)$$

this Green’s function describes propagation in the 2h1p subspace generated by the projection operator  $\mathcal{P}$ . The self-consistent eigenvalue problem defined in eq 28 is analogous to that of the *GW* approximation defined in eqs 3 and 5. There are two significant differences that originate from the treatment of time-ordering in EOM-CC theory. First, the particle contribu-

tion to the self-energy of the IP part of the Green's function  $\tilde{\Sigma}_{ij}^p$  is frequency-independent; analogously, in EA-EOM-CC, the hole contribution to the self-energy of the EA part of the Green's function is frequency-independent. However, as discussed above, the frequency dependence of these terms, when the respective excitation energy is calculated, is typically very weak. Second, the effective self-energy in IP-EOM-CC (respectively EA-EOM-CC) only has matrix elements in the occupied (virtual) orbital space; this is in contrast to the self-energy in any proper diagrammatic theory, which has matrix elements in the entire orbital space. We emphasize that neither of these differences reflects an approximation, but only a difference in formalism; diagrammatically defined self-energy theories and EOM-CC can both be made exact in their appropriate limits, while retaining their respective (different) mathematical structures.

To summarize the structure of this effective “self-energy” from an EOM-CC-based theory, for the IPs, the forward time-ordered self-energy (the hole contribution) arises from coupling between 1h and 2h1p configurations, whereas the reverse time-ordered self-energy (the particle contribution) arises from the similarity transformation of the Fock operator in the 1h subspace; this behavior is shown schematically in Figure 4. In the GW language, both of these effects can be viewed as giving rise to screening of the quasiparticle excitations. From this point of view, we observe that the IP-EOM-CISD methodology,<sup>58</sup> obtained by setting  $\hat{T}_1 = \hat{T}_2 = 0$  in IP-EOM-CCSD, only includes one of the two time-orderings, each diagram of which is fully forward-time-ordered in the TDA sense. Finally, we note that the true CCSD self-energy can be straightforwardly obtained numerically, by calculating both the IP and EA Green's functions and using Dyson's equation,

$$\Sigma_{pq}(\omega) = (\omega - \epsilon_p)\delta_{pq} - [\mathbf{G}^{\text{IP}}(\omega) + \mathbf{G}^{\text{EA}}(\omega)]_{pq}^{-1} \quad (30)$$

Naturally, the frequency-dependent matrix  $\omega - \mathbf{f} - \Sigma(\omega)$  will have eigenvalues given exactly by the IP/EA-EOM-CCSD excitation energies, as well as the proper analytical (frequency-dependent) structure.

We now proceed to make the comparison between the approximate IP-EOM-CCSD of eq 27 and the GW approximation more explicit. First, we consider the frequency-dependent hole contribution  $\Sigma_{ij}^h(\omega)$ . The 2h1p Green's function can be expressed in two ways: first, as a perturbative series that can be translated into diagrams,

$$[G_{2\text{h}1\text{p}}]_{klmn}^{ab}(\omega) = \sum_{p=0}^{\infty} \langle \Phi_{kl}^a | \hat{R}_0(\omega) [\mathcal{P} \hat{V}_N \hat{R}_0(\omega)]^p | \Phi_{mn}^b \rangle \quad (31)$$

where  $\hat{R}_0(\omega) = [\omega - (-\hat{f}_N)]^{-1}$  is the resolvent of the Fock operator and  $\hat{V}_N = \hat{H}_N - \hat{f}_N$  is the normal-ordered fluctuation operator. Alternatively,  $G_{2\text{h}1\text{p}}$  can be expressed in terms of the solutions of an eigenproblem,

$$[G_{2\text{h}1\text{p}}]_{klmn}^{ab}(\omega) = \sum_{\nu} \frac{r_{kl}^a(\nu) [r_{mn}^b(\nu)]^*}{\omega - E_{\nu}^{2\text{h}1\text{p}}} \quad (32)$$

where  $E_{\nu}^{2\text{h}1\text{p}}$  is an eigenvalue of the 2h1p block of the Hamiltonian  $\langle \Phi_{kl}^a | (-\hat{H}_N) | \Phi_{mn}^b \rangle$  with eigenvector  $|\nu\rangle = \frac{1}{2} \sum_{akl} r_{kl}^a(\nu) | \Phi_{kl}^a \rangle$ . Notably, the set of diagrams contained in  $G_{2\text{h}1\text{p}}$  is identical to those included in the two-particle-hole TDA theory of the self-energy,<sup>16</sup> mentioned in the Introduction.

The CC self-energy goes beyond the TDA diagrams via the outer vertices, i.e., the matrix elements of the similarity-transformed Hamiltonian, which can be evaluated to give

$$\langle \Phi_i | \bar{H} | \Phi_{kl}^a \rangle = W_{iakl} = \langle ia || kl \rangle + \sum_{me} \langle im || ke \rangle t_{lm}^{ae} - \sum_{me} \langle im || le \rangle t_{km}^{ae} + \frac{1}{2} \sum_{ef} \langle ia || ef \rangle t_{kl}^{ef} \quad (33a)$$

$$\langle \Phi_{mn}^b | \bar{H} | \Phi_i \rangle = \langle mn || jb \rangle \quad (33b)$$

leading to the self-energy

$$\tilde{\Sigma}_{ij}^h(\omega) = \frac{1}{4} \sum_{abklmn} W_{iakl} [G_{2\text{h}1\text{p}}]_{klmn}^{ab}(\omega) \langle mn || jb \rangle \quad (34)$$

Viewing  $W$  as a screened Coulomb interaction leads to the set of diagrams shown in Figure 4. The construction of the intermediate  $W$  has a noniterative  $N^6$  cost, which is usually swamped by the iterative  $N^5$  cost of subsequent matrix-vector multiplications during Davidson diagonalization.

The use of exact CCSD amplitudes in eq 33a includes many beyond-GW insertions in the polarization propagator. However, as discussed, the closest comparison can be made when the  $\hat{T}_2$  amplitudes solve the approximate ring-CCD equations,

$$t_{ij}^{ab}(\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b) = \langle ab || ij \rangle + \sum_{ck} t_{ik}^{ac} \langle kb || cj \rangle + \sum_{ck} \langle ak || ic \rangle t_{kj}^{cb} + \sum_{cdkl} t_{ik}^{ac} \langle kl || cd \rangle t_{lj}^{db} \quad (35)$$

Iteration of these equations adds higher-order non-TDA ring diagrams in the self-energy, very much like in the GW approximation. However, consistent with the analysis presented in section 3.1, the non-TDA diagrams are generated in an asymmetric and incomplete manner.

Beyond this issue of non-TDA diagrams on the later-time side of the self-energy, the approximation described so far has three additional qualitative differences from the GW approximation. First, the presence of antisymmetrized vertices generates many exchange diagrams not included in the conventional GW approximation. In particular, the “exterior” antisymmetrization is responsible for some of the self-energy diagrams that are in the second-order screened exchange (SOSEX) approach<sup>59</sup> and “interior” antisymmetrization yields particle–hole ladders that improve the quality of the polarization propagator. Second, the 2h1p Green's function includes the interaction between two holes in the intermediate 2h1p state, leading to hole–hole ladder insertions, which are vertex corrections beyond the structure of the GW self-energy. Third, the final term in eq 33a can be shown to produce mixed ring-ladder diagrams that are not included at the GW or SOSEX levels of theory.

Finally, we consider the frequency-independent particle contribution to the self-energy. In IP-EOM-CCD, this term is given by

$$\tilde{\Sigma}_{ij}^p = \frac{1}{2} \sum_{kab} \langle ik || ab \rangle t_{kj}^{ba} \quad (36)$$

which can be represented by the single diagram shown in Figure 4. This diagram must be evaluated as a scalar without frequency dependence according to the usual diagrammatic rules of time-independent perturbation theory.<sup>8</sup> With this interpretation, the

iteration of the ring-CCSD equations again generates all TDA-screening diagrams plus an asymmetric subset of non-TDA-screening diagrams. Antisymmetrization is responsible for subsets of both *GW* and SOSEX diagrams.

#### 4. APPLICATION OF EOM-CCSD TO THE GW100 TEST SET

Having established the formal relation between EOM-CCSD and the *GW* approximation, we now present a numerical comparison. In particular, we will study the so-called GW100 test set,<sup>40</sup> comprising 100 small- to medium-sized molecules with up to 66 active electrons in 400 spatial orbitals. The GW100 test set was introduced by van Setten and co-authors<sup>40</sup> in order to provide a simple and controlled class of problems with which to compare theoretical and computational approximations of *GW*-based implementations. This important research agenda aims to enforce reproducibility within the community and highlight the successes and limitations of the aforementioned time-dependent diagrammatic techniques, thereby identifying avenues for future research. The GW100 has been studied by a number of different groups.<sup>38–40,60,61</sup>

In addition to providing results and analysis for conventional IP- and EA-EOM-CCSD excitation energies, we will also consider a number of approximations. These approximations make the computational cost more competitive with that of the *GW* approximation and, in light of the previous sections, many of them can be understood as selective inclusion of certain diagrams. These approximations are described in the next section. Some of these approximations have been investigated and compared for charged excitations<sup>62–64</sup> and for neutral electronic excitations.<sup>65,66</sup>

**4.1. Approximations to EOM-CCSD.** As mentioned previously, in their canonical forms, the *GW* approximation and the IP/EA-EOM-CCSD formalism both scale as  $N^6$ . For the latter class of methods, this scaling originates from the solution of the ground-state CCSD equations, while the subsequent ionized EOM eigenvalue problem exhibits only  $N^5$  scaling (with relatively cheap, noniterative  $N^6$  steps associated with construction of the intermediates). For this reason, a natural target for approximations leading to reduced cost is the ground-state calculation. Despite the distinction we draw between ground-state and excited-state approximations, we note that the results of the previous section have shown that the determination of the *T*-amplitudes via ground-state CCSD directly affects the diagrams contributing to the one-particle Green's function.

**MBPT2 Ground State.** The most severe approximation to the ground state is that of second-order many-body perturbation theory (MBPT2). For a canonical Hartree–Fock reference, which we use throughout this work, this is equivalent to second-order Møller–Plesset perturbation theory (MP2). In this approach,  $\hat{T}_1 = 0$  and the  $\hat{T}_2$  amplitudes are approximated by

$$t_{ij}^{ab} \approx \frac{\langle ab||ij \rangle}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b} \quad (37)$$

Due to the transformation from atomic orbitals to molecular orbitals, an MBPT2 calculation scales as  $N^5$  and so the use of MBPT2 amplitudes in an IP/EA-EOM calculation leads to overall  $N^5$  methods for ionization potentials and electron affinities. Following refs 65, 67, and 68, we call this method EOM-MBPT2; the same method has also been referred to as EOM-CCSD(2).<sup>69</sup>

**CC2 Ground State.** A popular approximation to reduce the cost of CCSD is the CC2 model.<sup>70</sup> In this technique, the  $\hat{T}_1$  amplitude equations are unchanged from those of CCSD, while the  $\hat{T}_2$  amplitude equations 12c are changed such that  $\hat{T}_2$  only connects to the Fock operator

$$0 = \langle \Phi_{ij}^{ab} | e^{-\hat{T}_1} \hat{H}_N e^{\hat{T}_1} + e^{-\hat{T}_2} \hat{f}_N e^{\hat{T}_2} | \Phi \rangle \quad (38)$$

This leads to approximate  $\hat{T}_2$  amplitudes that are very similar to those of MBPT2,

$$t_{ij}^{ab} \approx \frac{\langle ab||ij \rangle}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b} \quad (39)$$

where  $\langle ab||ij \rangle \equiv \langle \Phi_{ij}^{ab} | e^{-\hat{T}_1} \hat{H}_N e^{\hat{T}_1} | \Phi \rangle$  are  $\hat{T}_1$ -transformed two-electron integrals. Like MBPT2, the CC2 approximation removes the  $\hat{T}_2$  contractions responsible for  $N^6$  scaling and is thus an iterative  $N^5$  technique. While CC2 treats dynamical correlation at essentially the same level as MBPT2, the full treatment of single excitations generated by  $\hat{T}_1$  allows orbital relaxation, which should be beneficial in cases where the HF determinant is suboptimal.

**Linearized CCSD.** The final ground-state approximation that we consider is linearized CCSD (linCCSD),<sup>71</sup> which is the least severe approximation to CCSD. In this approach, all quadratic products of the CCSD amplitudes are neglected in the amplitude equations. In diagrammatic language, this approximation neglects many, *but not all*, of the non-TDA diagrams in the Green's function; for example, the third-order non-TDA diagrams shown in Figure 2b,c are included even when the amplitude equations are linearized. The non-TDA time-ordering is a result of the combination of  $\hat{R}_1$  and  $\hat{T}_2$ , rather than of nonlinear terms in the  $\hat{T}_2$  equations. Although linearized CCSD still scales as  $N^6$ , the method is more amenable to parallelization,<sup>71</sup> which may be desirable for large systems or solids.<sup>47–49,72</sup>

**Excited-State Approximation.** After the ground-state calculation, the most expensive contribution to an EOM-CCSD calculation comes from the large doubles–doubles block of the similarity-transformed Hamiltonian. A natural approximation then is to replace the doubles–doubles block by simple orbital-energy differences,

$$\langle \Phi_{ij}^a | \bar{H}_N | \Phi_{kl}^b \rangle \approx (\varepsilon_a - \varepsilon_i - \varepsilon_j) \delta_{ab} \delta_{ik} \delta_{jl} \quad (40)$$

leading to a diagonal structure and a straightforward Löwdin partitioning. Naturally, this approach is only reasonable for principle charged excitations with a large weight in the singles (one-hole or one-particle) sector.

This partitioned variant of EOM-CCSD theory still exhibits  $N^5$  scaling after the  $\hat{T}$ -amplitudes are determined but requires the construction and storage of far fewer integral intermediates. Formally, this approximate partitioning technique can be combined with any treatment of the ground-state CC equations, though it only makes practical sense for approximate ground-state calculations whose cost does not overwhelm that of the EOM calculation. We will combine the approximate partitioning technique with MBPT2 and CC2 ground states, denoting the results as P-EOM-MBPT2 and P-EOM-CC2, respectively.

As first discussed in ref 67, the P-EOM-MBPT2 method is formally very close to the nonself-consistent second-order Green's function technique (GF2), where the self-energy is composed of second-order ring and exchange diagrams.<sup>20,73</sup> When applied exactly as described, P-EOM-MBPT2 actually

includes a few third-order self-energy diagrams, as can be seen in Figure 4. These can be removed by also neglecting the  $\hat{T}$ -amplitudes in the screened Coulomb interaction  $W_{iakl}$  that couples the 1h and 2h1p space, given in eq 33a. However, this additional ground-state correlation is found to be responsible for a remarkable improvement in the accuracy of P-EOM-MBPT2 when compared to GF2.

With respect to the hierarchy of linear-response CC2 methods described in ref 74, the EOM-CC2 method described here is equivalent to IP-CCSD[f]<sub>CC2</sub> and the P-EOM-CC2 method is between IP-CCSD[0]<sub>CC2</sub> and IP-CCSD[1]<sub>CC2</sub>.

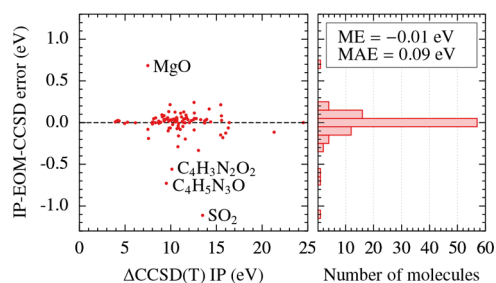
**4.2. Numerical Details.** We have applied the above methods to calculate the first few principle ionization potentials and electron affinities for the molecules in the GW100 test set. Following refs 38 and 39, we work in the localized-orbital def2-TZVPP basis set,<sup>75</sup> using corresponding pseudopotentials for elements in the fifth and sixth rows of the periodic table; core orbitals were frozen in all calculations. While this choice of basis is a good trade-off between cost and accuracy, our results are not converged with respect to the basis set and should not be compared directly to experiment or to calculations in other basis sets, such as plane-wave based GW calculations.<sup>60,61</sup> Instead, these calculations can be directly compared to preexisting ionization potentials in the same basis.<sup>38,39</sup> More importantly, our calculations are internally consistent; the main purpose of this section is to benchmark the accuracy of cost-saving approximations to EOM-CCSD and demonstrate the utility of EOM-CCSD techniques for excited-state properties of benchmark data sets. Extrapolation to the complete basis set limit and comparison with other non-CC techniques is reserved for future work.

An advantage of the IP/EA-EOM-CCSD approaches is an avoidance of open-shell calculations for charged molecules. As such, all of our calculations were performed using a spin-free implementation based on a closed-shell restricted HF reference, and free of spin contamination. All calculations were performed using the PySCF software package.<sup>76</sup>

Recent work<sup>60</sup> has identified two molecules from the original GW100 test set with incorrect geometries: vinyl bromide and phenol. For consistency with previously published results, we have performed calculations on the original geometries.

**4.3. Comparison to  $\Delta$ CCSD(T) Ionization Potentials.** We first aim to establish the accuracy of EOM-CCSD for the GW100 test set. As a ground-state theory, CCSD with perturbative triples [CCSD(T)] represents the “gold standard” for weakly correlated medium-sized molecules<sup>77</sup> and scales as  $N^7$ , which is more expensive than any method considered here. At this high level of theory, the first IP of each molecule has been calculated by Krause et al.,<sup>38</sup> as a difference in ground-state energies between neutral and charged molecules, the so-called  $\Delta$ CCSD(T) scheme. Higher-energy IPs and EAs, in particular those with the same symmetry as the first, cannot be calculated using this approach

In Figure 5, we show the comparison between IPs predicted by  $\Delta$ CCSD(T) and IP-EOM-CCSD. The IP-EOM-CCSD values exhibit a signed mean error (ME) of  $-0.01$  eV and a mean absolute error (MAE) of  $0.09$  eV. The small mean error indicates that the errors are not systematic. Only four molecules, identified in Figure 5 have errors larger than  $0.5$  eV, suggesting that IP-EOM-CCSD represents a good approximation to  $\Delta$ CCSD(T), at least for the molecules included in the GW100. The most fair comparison, based on diagrams generated, is to the  $G_0W_0$  approximation based on a HF



**Figure 5.** Comparison of ionization potentials predicted by IP-EOM-CCSD compared to those predicted by  $\Delta$ CCSD(T) from ref 38. Errors with respect to the latter are presented as a scatter plot (left) and histogram (right). Four molecules with errors exceeding  $0.5$  eV are explicitly labeled.

reference ( $G_0W_0$ @HF). As reported by Caruso et al.,<sup>39</sup> such an approach overestimates IPs, leading to a ME of  $0.26$  eV and a MAE of  $0.35$  eV. Interestingly, this is significantly better than the more popular (at least in the solid state)  $G_0W_0$  approximation based on a PBE starting point ( $G_0W_0$ @PBE), which severely and systematically underestimates IPs, leading to a ME of  $-0.69$  eV and a MAE of  $0.69$  eV.<sup>39</sup>

#### 4.4. Electron Affinities and Higher-Energy Excitations.

By construction, EOM-CCSD can straightforwardly predict higher-energy ionization potentials, corresponding to more deeply bound electrons, as well as the first and higher electron affinities. In order to provide a larger set of benchmark data for the GW community, in Table 1 we report the first three occupied and unoccupied quasiparticle energies (i.e., the negative of the first three IPs and EAs with large quasiparticle weights) for each molecule in the GW100 test set, as calculated by IP- and EA-EOM-CCSD, and accounting for their multiplicities.

**4.5. Accuracy of Approximate EOM-CCSD.** We next assess the accuracy of approximations to EOM-CCSD, using the GW100 test set. Henceforth, we compare all approximations to EOM-CCSD, and not to  $\Delta$ CCSD(T), for a number of reasons. First, the comparison is perhaps the most fair because all approximate techniques are derived from EOM-CCSD, and so the most we can expect is that they reproduce this parent method. Second, although EOM-CCSD was shown above to provide an accurate reproduction of the  $\Delta$ CCSD(T) values, the latter approach can be challenging for open-shell systems like those used in the  $(N \pm 1)$ -electron calculations. For example, while the unrestricted formalism used in ref 38 provides a better approximate treatment of multireference effects, it also suffers from spin contamination, which can affect the IPs and EAs by up to  $0.5$  eV, as discussed in ref 78. Third, it allows us to compare EAs, which are not available in the literature based on  $\Delta$ CCSD(T).

In Table 2, we present IP and EA error metrics for a variety of approximate techniques (all 100 molecules were studied by each approach except for CC2-based approaches, because the ground-state CC2 failed to converge for 12 molecules; the full data set for individual molecules is given as Supporting Information). Perhaps most remarkably, we find that all CC-based methods exhibit MEs of less than  $0.13$  eV and MAEs of less than  $0.16$  eV. These results can be compared to those based on the GW approximation, as given in Table 2. These latter results are taken from ref 39 and include only ionization potentials compared to CCSD(T) data; given the good agreement between CCSD(T) and IP-EOM-CCSD shown above, we expect similar error metrics if the GW results were



Table 1. Quasiparticle Energies (Negative of the Ionization Potentials and Electron Affinities) of Molecules in the GW100 Calculated with IP/EA-EOM-CCSD in the def2-TZVPP Basis Set

formula	name	HOMO-2	HOMO-1	HOMO	LUMO	LUMO+1	LUMO+2
He	helium			-24.51 (x1)	22.22 (x1)	39.82 (x3)	166.90 (x1)
Ne	neon		-48.33 (x1)	-21.21 (x3)	20.84 (x1)	21.87 (x2)	74.27 (x3)
Ar	argon		-29.58 (x1)	-15.63 (x2)	14.73 (x3)	17.19 (x4)	20.69 (x1)
Kr	krypton		-27.14 (x1)	-13.98 (x3)	10.41 (x2)	12.24 (x4)	18.78 (x1)
Xe	xenon		-23.67 (x1)	-12.23 (x3)	7.72 (x3)	8.91 (x3)	12.29 (x1)
H <sub>2</sub>	hydrogen			-16.40 (x1)	4.22 (x1)	8.05 (x1)	16.04 (x1)
Li <sub>2</sub>	lithium dimer	-63.39 (x1)	-63.38 (x1)	-5.27 (x1)	-0.12 (x1)	0.99 (x2)	1.00 (x1)
Na <sub>2</sub>	sodium dimer	-37.24 (x2)	-37.21 (x1)	-4.94 (x1)	-0.26 (x1)	0.60 (x2)	0.78 (x1)
Na <sub>4</sub>	sodium tetramer	-36.60 (x1)	-5.59 (x1)	-4.25 (x1)	-0.53 (x1)	-0.10 (x1)	0.10 (x1)
Na <sub>6</sub>	sodium hexamer	-36.52 (x1)	-5.68 (x1)	-4.37 (x2)	-0.49 (x1)	-0.34 (x2)	-0.07 (x1)
K <sub>2</sub>	potassium dimer	-23.83 (x2)	-23.77 (x1)	-4.08 (x1)	-0.32 (x1)	0.77 (x1)	0.83 (x2)
Rb <sub>2</sub>	rubidium dimer	-20.09 (x2)	-20.02 (x1)	-3.93 (x1)	-0.37 (x1)	0.18 (x1)	0.35 (x1)
N <sub>2</sub>	nitrogen	-18.84 (x1)	-17.21 (x1)	-15.60 (x1)	3.05 (x2)	8.97 (x1)	9.47 (x1)
P <sub>2</sub>	phosphorus dimer	-14.79 (x1)	-10.75 (x1)	-10.59 (x1)	-0.10 (x2)	3.47 (x1)	6.30 (x1)
As <sub>2</sub>	arsenic dimer	-14.64 (x1)	-10.14 (x1)	-9.91 (x2)	-0.26 (x1)	3.31 (x1)	6.88 (x1)
F <sub>2</sub>	fluorine	-21.09 (x1)	-18.85 (x1)	-15.53 (x1)	0.39 (x1)	15.34 (x1)	15.34 (x1)
Cl <sub>2</sub>	chlorine	-15.98 (x1)	-14.41 (x1)	-11.46 (x2)	-0.19 (x1)	9.10 (x1)	10.03 (x1)
Br <sub>2</sub>	bromine	-14.44 (x1)	-12.89 (x1)	-10.54 (x1)	-0.94 (x1)	7.05 (x1)	7.09 (x1)
I <sub>2</sub>	iodine	-12.76 (x1)	-11.38 (x2)	-9.55 (x2)	-1.50 (x1)	5.10 (x1)	5.15 (x1)
CH <sub>4</sub>	methane		-23.38 (x1)	-14.38 (x3)	3.45 (x1)	5.79 (x3)	7.88 (x3)
C <sub>2</sub> H <sub>6</sub>	ethane	-13.05 (x1)	-12.71 (x1)	-12.71 (x1)	3.11 (x1)	4.21 (x1)	5.24 (x2)
C <sub>3</sub> H <sub>8</sub>	propane	-12.25 (x1)	-12.12 (x1)	-12.05 (x1)	2.95 (x1)	4.13 (x1)	4.32 (x1)
C <sub>4</sub> H <sub>10</sub>	butane	-11.92 (x1)	-11.81 (x1)	-11.56 (x1)	2.88 (x1)	3.46 (x1)	4.12 (x1)
C <sub>2</sub> H <sub>4</sub>	ethylene	-14.91 (x1)	-13.11 (x1)	-10.69 (x1)	2.63 (x1)	3.94 (x1)	4.60 (x1)
C <sub>2</sub> H <sub>2</sub>	ethyne	-19.13 (x1)	-17.23 (x1)	-11.55 (x2)	3.50 (x2)	3.58 (x1)	4.53 (x1)
C <sub>4</sub>	tetracarbon	-14.65 (x1)	-11.46 (x1)	-11.27 (x1)	-2.36 (x1)	-0.01 (x1)	1.71 (x1)
C <sub>3</sub> H <sub>6</sub>	cyclopropane	-13.14 (x1)	-10.86 (x1)	-10.85 (x1)	3.46 (x1)	3.96 (x1)	4.10 (x1)
C <sub>6</sub> H <sub>6</sub>	benzene	-12.14 (x2)	-9.32 (x1)	-9.32 (x1)	1.78 (x2)	3.11 (x1)	4.00 (x2)
C <sub>8</sub> H <sub>8</sub>	cyclooctatetraene	-10.01 (x1)	-10.00 (x1)	-8.40 (x1)	0.79 (x1)	2.52 (x1)	2.54 (x1)
C <sub>5</sub> H <sub>6</sub>	cyclopentadiene	-12.52 (x1)	-11.03 (x1)	-8.69 (x1)	1.77 (x1)	3.27 (x1)	4.06 (x1)
C <sub>2</sub> H <sub>3</sub> F	vinyl fluoride	-14.79 (x1)	-13.86 (x1)	-10.60 (x1)	2.80 (x1)	3.93 (x1)	4.34 (x1)
C <sub>2</sub> H <sub>3</sub> Cl	vinyl chloride	-13.21 (x1)	-11.65 (x1)	-10.13 (x1)	2.12 (x1)	3.49 (x1)	3.84 (x1)
C <sub>2</sub> H <sub>3</sub> Br	vinyl bromide	-13.38 (x1)	-10.71 (x1)	-9.29 (x1)	2.02 (x1)	3.55 (x1)	4.27 (x1)
C <sub>2</sub> H <sub>3</sub> I	vinyl iodide	-11.71 (x1)	-9.92 (x1)	-9.36 (x1)	1.40 (x1)	1.75 (x1)	3.59 (x1)
CF <sub>4</sub>	tetrafluoromethane	-18.35 (x2)	-17.37 (x2)	-16.24 (x3)	4.89 (x1)	6.86 (x2)	9.20 (x2)
CCl <sub>4</sub>	tetrachloromethane	-13.40 (x2)	-12.46 (x3)	-11.60 (x3)	0.86 (x1)	2.20 (x3)	5.20 (x1)
CBr <sub>4</sub>	tetrabromomethane	-12.13 (x2)	-11.25 (x3)	-10.48 (x2)	-0.49 (x1)	1.21 (x3)	4.60 (x1)
CI <sub>4</sub>	tetraiodomethane		-9.99 (x2)	-9.30 (x2)	-1.62 (x1)	0.34 (x2)	4.73 (x2)
SiH <sub>4</sub>	silane		-18.46 (x1)	-12.84 (x3)	3.10 (x3)	3.71 (x1)	6.75 (x2)
GeH <sub>4</sub>	germane	-38.07 (x1)	-18.69 (x1)	-12.53 (x2)	3.16 (x1)	3.54 (x3)	7.04 (x2)
Si <sub>2</sub> H <sub>6</sub>	disilane	-12.25 (x1)	-12.25 (x1)	-10.71 (x1)	2.27 (x1)	2.28 (x2)	2.75 (x1)
Si <sub>5</sub> H <sub>12</sub>	pentasilane	-10.84 (x1)	-10.64 (x1)	-9.36 (x1)	0.79 (x1)	1.57 (x1)	1.61 (x1)
LiH	lithium hydride		-64.54 (x1)	-7.96 (x1)	0.09 (x1)	2.01 (x2)	3.41 (x1)
KH	potassium hydride	-24.59 (x2)	-24.38 (x1)	-6.13 (x1)	-0.04 (x1)	1.60 (x2)	1.90 (x1)
BH <sub>3</sub>	borane		-18.35 (x1)	-13.31 (x2)	0.33 (x1)	3.36 (x1)	4.32 (x2)
B <sub>2</sub> H <sub>6</sub>	diborane	-14.00 (x1)	-13.48 (x1)	-12.29 (x1)	1.20 (x1)	2.51 (x1)	3.46 (x1)
NH <sub>3</sub>	ammonia	-27.78 (x1)	-16.52 (x2)	-10.77 (x1)	2.84 (x1)	5.26 (x2)	11.18 (x2)
HN <sub>3</sub>	hydrazoic acid	-15.92 (x1)	-12.25 (x1)	-10.72 (x1)	2.02 (x1)	3.02 (x1)	3.17 (x1)
PH <sub>3</sub>	phosphine	-20.25 (x1)	-13.75 (x2)	-10.57 (x1)	2.95 (x1)	3.12 (x2)	7.11 (x2)
AsH <sub>3</sub>	arsine	-19.82 (x1)	-13.18 (x2)	-10.42 (x1)	2.86 (x1)	3.01 (x2)	7.44 (x2)
SH <sub>2</sub>	hydrogen sulfide	-15.65 (x1)	-13.39 (x1)	-10.35 (x1)	2.79 (x1)	3.20 (x1)	7.25 (x1)
FH	hydrogen fluoride	-39.30 (x1)	-19.84 (x1)	-15.90 (x1)	3.07 (x1)	14.20 (x1)	17.24 (x1)
ClH	hydrogen chloride	-25.44 (x1)	-16.65 (x1)	-12.64 (x1)	2.70 (x1)	7.91 (x1)	12.18 (x1)
LiF	lithium fluoride	-33.11 (x1)	-11.76 (x1)	-11.28 (x2)	-0.02 (x1)	2.74 (x2)	3.51 (x1)
F <sub>2</sub> Mg	magnesium fluoride	-14.15 (x1)	-13.76 (x2)	-13.71 (x1)	-0.04 (x1)	1.94 (x2)	4.13 (x1)
TiF <sub>4</sub>	titanium fluoride	-16.86 (x1)	-16.29 (x2)	-15.69 (x2)	-1.06 (x2)	0.07 (x3)	0.98 (x1)
AlF <sub>3</sub>	aluminum fluoride	-16.08 (x2)	-15.86 (x2)	-15.31 (x1)	0.67 (x1)	1.86 (x1)	3.90 (x2)
BF	boron monofluoride	-21.27 (x1)	-18.16 (x2)	-11.20 (x1)	1.51 (x2)	3.29 (x1)	4.70 (x1)
SF <sub>4</sub>	sulfur tetrafluoride	-15.26 (x1)	-15.04 (x1)	-12.70 (x1)	0.94 (x1)	3.79 (x1)	4.86 (x1)
BrK	potassium bromide	-19.37 (x1)	-8.41 (x1)	-8.17 (x2)	-0.45 (x1)	1.40 (x2)	1.73 (x1)

Table 1. continued

formula	name	HOMO-2	HOMO-1	HOMO	LUMO	LUMO+1	LUMO+2
GaCl	gallium monochloride	-14.09 (x1)	-11.46 (x2)	-9.79 (x1)	0.27 (x2)	2.49 (x1)	6.60 (x1)
NaCl	sodium chloride	-20.83 (x1)	-9.55 (x1)	-9.12 (x2)	-0.59 (x1)	1.18 (x2)	2.08 (x1)
MgCl <sub>2</sub>	magnesium chloride	-12.52 (x1)	-11.88 (x2)	-11.76 (x2)	-0.19 (x1)	1.40 (x2)	3.91 (x1)
Al <sub>3</sub>	aluminum chloride	-10.39 (x2)	-10.29 (x2)	-9.84 (x1)	-0.33 (x1)	-0.28 (x1)	2.25 (x2)
BN	boron nitride	-27.96 (x1)	-13.69 (x1)	-11.93 (x2)	-3.16 (x1)	2.65 (x2)	3.69 (x1)
NCH	hydrogen cyanide	-20.60 (x1)	-13.91 (x1)	-13.90 (x1)	3.21 (x2)	3.45 (x1)	4.55 (x1)
PN	phosphorus mononitride	-16.37 (x1)	-12.42 (x1)	-11.80 (x1)	0.46 (x1)	3.40 (x1)	8.42 (x1)
H <sub>2</sub> NNH <sub>2</sub>	hydrazine	-15.39 (x1)	-11.28 (x1)	-9.62 (x1)	2.51 (x1)	3.69 (x1)	4.55 (x1)
H <sub>2</sub> CO	formaldehyde	-16.04 (x1)	-14.56 (x1)	-10.78 (x1)	1.67 (x1)	3.68 (x1)	5.24 (x1)
CH <sub>4</sub> O	methanol	-14.32 (x1)	-13.33 (x1)	-10.18 (x1)	1.91 (x1)	3.14 (x1)	4.00 (x1)
C <sub>2</sub> H <sub>6</sub> O	ethanol	-13.43 (x1)	-12.27 (x1)	-10.61 (x1)	2.86 (x1)	3.73 (x1)	4.69 (x1)
C <sub>2</sub> H <sub>4</sub> O	acetaldehyde	-14.32 (x1)	-13.33 (x1)	-10.18 (x1)	1.91 (x1)	3.14 (x1)	4.00 (x1)
C <sub>4</sub> H <sub>10</sub> O	ethoxy ethane	-12.36 (x1)	-11.50 (x1)	-9.75 (x1)	2.96 (x1)	3.46 (x1)	3.93 (x1)
CH <sub>2</sub> O <sub>2</sub>	formic acid	-14.94 (x1)	-12.55 (x1)	-11.42 (x1)	2.70 (x1)	3.07 (x1)	4.29 (x1)
HOOH	hydrogen peroxide	-15.38 (x1)	-12.84 (x1)	-11.39 (x1)	3.01 (x1)	3.02 (x1)	4.87 (x1)
H <sub>2</sub> O	water	-18.90 (x1)	-14.70 (x1)	-12.48 (x1)	2.88 (x1)	4.91 (x1)	13.32 (x1)
CO <sub>2</sub>	carbon dioxide	-18.11 (x1)	-17.99 (x2)	-13.73 (x1)	2.80 (x1)	4.29 (x2)	6.59 (x1)
CS <sub>2</sub>	carbon disulfide	-14.57 (x1)	-13.27 (x2)	-10.01 (x2)	0.29 (x2)	3.34 (x1)	4.46 (x1)
OCS	carbon oxide sulfide	-16.12 (x1)	-16.12 (x2)	-11.24 (x2)	1.85 (x2)	3.13 (x1)	4.78 (x1)
OCS <sub>e</sub>	carbon oxide selenide	-15.89 (x2)	-15.63 (x1)	-10.50 (x2)	1.44 (x2)	2.60 (x1)	4.18 (x1)
CO	carbon monoxide	-19.42 (x1)	-15.49 (x1)	-14.37 (x1)	1.22 (x2)	5.30 (x1)	6.59 (x1)
O <sub>3</sub>	ozone	-13.48 (x1)	-12.93 (x1)	-12.79 (x1)	-1.52 (x1)	5.23 (x1)	7.28 (x1)
SO <sub>2</sub>	sulfur dioxide	-13.50 (x1)	-13.12 (x1)	-12.37 (x1)	-0.34 (x1)	3.98 (x1)	4.39 (x1)
BeO	beryllium monoxide	-26.77 (x1)	-10.97 (x1)	-9.88 (x2)	-2.01 (x1)	2.31 (x1)	2.48 (x1)
MgO	magnesium monoxide	-24.89 (x1)	-8.76 (x1)	-8.17 (x2)	-1.29 (x1)	1.16 (x2)	2.90 (x1)
C <sub>7</sub> H <sub>8</sub>	toluene	-11.74 (x1)	-9.19 (x1)	-8.90 (x1)	1.71 (x1)	1.85 (x1)	2.97 (x1)
C <sub>8</sub> H <sub>10</sub>	ethylbenzene	-11.57 (x1)	-9.15 (x1)	-8.85 (x1)	1.76 (x1)	1.76 (x1)	2.85 (x1)
C <sub>6</sub> F <sub>6</sub>	hexafluorobenzene	-14.09 (x1)	-13.12 (x1)	-10.15 (x2)	1.08 (x2)	1.15 (x1)	3.66 (x2)
C <sub>6</sub> H <sub>5</sub> OH	phenol	-11.99 (x1)	-9.42 (x1)	-8.69 (x1)	1.62 (x1)	2.35 (x1)	2.84 (x1)
C <sub>6</sub> H <sub>5</sub> NH <sub>2</sub>	aniline	-11.01 (x1)	-9.21 (x1)	-7.98 (x1)	1.83 (x1)	2.29 (x1)	2.82 (x1)
C <sub>5</sub> H <sub>5</sub> N	pyridine	-10.45 (x1)	-9.74 (x1)	-9.72 (x1)	1.24 (x1)	1.62 (x1)	3.21 (x1)
C <sub>5</sub> H <sub>5</sub> N <sub>5</sub> O	guanine	-10.05 (x1)	-9.81 (x1)	-8.04 (x1)	1.57 (x1)	1.87 (x1)	1.98 (x1)
C <sub>5</sub> H <sub>5</sub> N <sub>5</sub>	adenine	-9.59 (x1)	-9.39 (x1)	-8.33 (x1)	1.28 (x1)	2.06 (x1)	2.51 (x1)
C <sub>4</sub> H <sub>5</sub> N <sub>3</sub> O	cytosine	-9.66 (x1)	-9.54 (x1)	-8.78 (x1)	0.92 (x1)	2.29 (x1)	2.51 (x1)
C <sub>5</sub> H <sub>4</sub> N <sub>2</sub> O <sub>2</sub>	thymine	-10.55 (x1)	-10.19 (x1)	-9.15 (x1)	0.77 (x1)	2.15 (x1)	2.42 (x1)
C <sub>4</sub> H <sub>4</sub> N <sub>2</sub> O <sub>2</sub>	uracil	-10.65 (x1)	-10.29 (x1)	-9.57 (x1)	0.70 (x1)	2.10 (x1)	2.35 (x1)
CH <sub>4</sub> N <sub>2</sub> O	urea	-10.65 (x1)	-10.52 (x1)	-10.08 (x1)	2.33 (x1)	3.51 (x1)	4.09 (x1)
Ag <sub>2</sub>	silver dimer	-11.08 (x2)	-10.82 (x1)	-7.41 (x1)	-0.70 (x1)	1.00 (x2)	1.41 (x1)
Cu <sub>2</sub>	copper dimer	-9.37 (x2)	-9.22 (x1)	-7.38 (x1)	-0.34 (x1)	2.11 (x1)	2.16 (x2)
NCCu	copper cyanide	-12.17 (x1)	-11.21 (x2)	-10.69 (x1)	-0.98 (x1)	1.92 (x2)	3.08 (x1)

Table 2. Mean Error (ME) and Mean Absolute Error (MAE) in eV of Ionization Potentials (IPs) and Electron Affinities (EAs) for Molecules Contained in the GW100 Test Set<sup>a</sup>

method	IP ME (eV)	IP MAE (eV)	EA ME (eV)	EA MAE
EOM-linCCSD	0.13	0.14	0.05	0.11
EOM-CC2	0.00	0.14	0.08	0.15
EOM-MBPT2	0.03	0.13	0.08	0.16
P-EOM-CC2	-0.08	0.12	-0.04	0.08
P-EOM-MBPT2	-0.08	0.16	-0.03	0.08
GF2@HF	-0.38	0.42	-0.19	0.22
G <sub>0</sub> W <sub>0</sub> @HF <sup>39</sup>	0.26	0.35		
G <sub>0</sub> W <sub>0</sub> @PBE <sup>39</sup>	-0.69	0.69		
scGW <sup>39</sup>	-0.30	0.32		

<sup>a</sup>Error metrics are calculated with respect to IP/EA-EOM-CCSD without approximation, except for the GW results from ref 39, which are calculated with respect to  $\Delta$ CCSD(T) results from ref 38.

compared to IP-EOM-CCSD results instead. Depending on the choice of mean-field reference or with full self-consistency, GW-based approximations give errors on the order of 0.3–0.7 eV.

Overall, we see that approximations in the ground-state calculation lead to an average increase in the IP or EA and approximations in the EOM calculation lead to an average decrease in the IP or EA. This behavior can be understood because most perturbative approximations to CCSD lead to overcorrelation, which decreases the ground-state energy (increases the IP or EA) or decreases the excited-state energy (decreases the IP or EA).

Without any partitioning, the EOM-linCCSD, EOM-CC2, and EOM-MBPT2 all perform similarly. Although the error incurred by the most expensive linearized CCSD is slightly larger (compared to the error of other approximate treatments of the ground state), the error is extremely systematic with a very small spread; for example, over 50 molecules overestimate the IP by 0.1 eV and another 30 molecules overestimate the IP by 0.2 eV. All  $N^5$  approximate methods (based on CC2 or MBPT2, with or

without partitioning) perform impressively well. EOM-CC2 and EOM-MBPT2 exhibit very similar results (even on the level of individual molecules), suggesting that the orbital relaxation due to  $\hat{T}_1$  is not important in many of these cases. With partitioning, the ME becomes only slightly negative without any significant increase in the MAE.

The qualitatively similar performances of all approximate EOM-CC methods and their improvement compared to that for  $G_0W_0$ @HF suggests that the precise details of screening are not important in molecules, and perhaps second-order exchange is more important. To test this, we also show the results of IPs and EAs calculated using the second-order Green's function (GF2); the mean error is large and negative,  $-0.38$  eV for IPs and  $-0.19$  eV for EAs, which can be compared to the IP results of  $G_0W_0$ @HF,  $+0.26$  eV. Remarkably, the P-EOM-MBPT2 approach, which has a cost essentially identical to that for GF2, reduces the mean error of the latter to only  $-0.08$  eV. This result suggests that the combination of second-order exchange with a small amount of screening, beyond the second-order ring diagram, is important for quantitative accuracy. Given the extremely low cost, we identify P-EOM-MBPT2 as an attractive low-cost approach for IPs and EAs of larger molecules, with potential applications in the solid state. However, it must be kept in mind that the increased importance of screening in solids may preclude the success of perturbative approximations.

## 5. CONCLUSIONS AND OUTLOOK

To summarize, we have presented a diagrammatic, algebraic, and numerical evaluation of quasiparticle excitation energies predicted by EOM-CCSD, especially as compared to those of the  $GW$  approximation. Although the EOM-CCSD Green's function includes fewer ring diagrams than the  $GW$  approximation, we find that its inclusion of many more diagram (including ladders and exchange) produces excitation energies that are much more accurate than those from the  $GW$  approximation. To completely encompass all  $GW$  diagrams requires the use of nonperturbative EOM-CCSDT.

We also investigated the accuracy of a number of cost-saving approximations to EOM-CCSD, many of which reduce the canonical scaling to  $N^5$  (which could be further reduced through density-fitting<sup>79</sup> or tensor hypercontraction<sup>80</sup>). All CC-based approximations considered yield very small errors on average. For systems where screening is relatively unimportant, such as molecules or large band gap insulators, we identify P-EOM-MBPT2 as an accurate and inexpensive  $N^5$  approach. We attribute the success of P-EOM-MBPT2 to its exact treatment of screening and exchange through second order (as in GF2), combined with a small number of third-order diagrams.

We anticipate that the framework and connections laid out here will aid future work on the  $GW$  approximation, through the identification of the most important excluded diagrams. With respect to IP/EA-EOM-CCSD calculations of band structures in solids,<sup>49</sup> the present work motivates efforts to quantify the error induced by neglecting some of the non-TDA ring diagrams, which are conventionally thought to be crucial for screening in solids. In the same vein, the inclusion of triple excitations, perhaps even perturbatively, could be an important ingredient in recovering—and rigorously surpassing—RPA physics.

## ■ ASSOCIATED CONTENT

### 📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.jctc.8b00455](https://doi.org/10.1021/acs.jctc.8b00455).

Detailed tables of first ionization potentials and electron affinities for all molecules in the GW100 and all approximate CC methods described in the text (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

\*T. C. Berkelbach. E-mail: [berkelbach@uchicago.edu](mailto:berkelbach@uchicago.edu).

### ORCID

Timothy C. Berkelbach: [0000-0002-7445-2136](https://orcid.org/0000-0002-7445-2136)

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

T.C.B. thanks Garnet Chan for helpful conversations related to this work, Marco Govoni for discussions about the GW100, and Wim Klopper for sharing the data from ref 38. We thank Alan Lewis for a critical reading of the manuscript. Diagrams were drawn using the JaxoDraw program.<sup>81</sup> This work was supported by start-up funding from the University of Chicago and by the Midwest Integrated Center for Computational Materials (MICCoM), as part of the Computational Materials Sciences Program funded by the U.S. Department of Energy (DOE), Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division (SJ-30161-0010A).

## ■ REFERENCES

- (1) Hedin, L. New Method for Calculating the One-Particle Green's Function with Application to the Electron-Gas Problem. *Phys. Rev.* **1965**, *139*, A796–A823.
- (2) Hybertsen, M. S.; Louie, S. G. First-principles theory of quasiparticles: Calculation of band gaps in semiconductors and insulators. *Phys. Rev. Lett.* **1985**, *55*, 1418–1421.
- (3) Schindlmayr, A.; Godby, R. W. Systematic Vertex Corrections through Iterative Solution of Hedin's Equations Beyond the  $GW$  Approximation. *Phys. Rev. Lett.* **1998**, *80*, 1702–1705.
- (4) Shishkin, M.; Marsman, M.; Kresse, G. Accurate quasiparticle spectra from self-consistent  $GW$  calculations with vertex corrections. *Phys. Rev. Lett.* **2007**, *99*, 14–17.
- (5) Romaniello, P.; Guyot, S.; Reining, L. The self-energy beyond  $GW$ : Local and nonlocal vertex corrections. *J. Chem. Phys.* **2009**, *131*, 154111.
- (6) Grüneis, A.; Kresse, G.; Hinuma, Y.; Oba, F. Ionization potentials of solids: The importance of vertex corrections. *Phys. Rev. Lett.* **2014**, *112*, 096401.
- (7) Schmidt, P. S.; Patrick, C. E.; Thygesen, K. S. Simple vertex correction improves  $GW$  band energies of bulk and two-dimensional crystals. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2017**, *96*, 205206.
- (8) Shavitt, I.; Bartlett, R. J. *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*; Cambridge University Press: Cambridge; New York, 2009.
- (9) Freeman, D. L. Coupled-cluster expansion applied to the electron gas: Inclusion of ring and exchange effects. *Phys. Rev. B* **1977**, *15*, 5512–5521.
- (10) Ring, P.; Shuck, P. *The Nuclear Many-Body Problem*; Springer-Verlag: Berlin Heidelberg, 1980.
- (11) Scuseria, G. E.; Henderson, T. M.; Sorensen, D. C. The ground state correlation energy of the random phase approximation from a ring coupled cluster doubles approach. *J. Chem. Phys.* **2008**, *129*, 231101.
- (12) Cederbaum, L. S. One-body Green's function for atoms and molecules: theory and application. *J. Phys. B: At. Mol. Phys.* **1975**, *8*, 290.
- (13) von Niessen, W.; Schirmer, J.; Cederbaum, L. S. Computational methods for the one-particle green's function. *Comput. Phys. Rep.* **1984**, *1*, 57–125.

- (14) Ortiz, J. V. Partial third-order quasiparticle theory: Comparisons for closed-shell ionization energies and an application to the Borazine photoelectron spectrum. *J. Chem. Phys.* **1996**, *104*, 7599–7605.
- (15) Ortiz, J. V. An efficient, renormalized self-energy for calculating the electron binding energies of closed-shell molecules and anions. *Int. J. Quantum Chem.* **2005**, *105*, 803–808.
- (16) Schirmer, J.; Cederbaum, L. S. The two-particle-hole Tamm-Dancoff approximation (2ph-TDA) equations for closed-shell atoms and molecules. *J. Phys. B: At. Mol. Phys.* **1978**, *11*, 1889–1900.
- (17) Walter, O.; Schirmer, J. The two-particle-hole Tamm-Dancoff approximation (2ph-TDA) for atoms. *J. Phys. B: At. Mol. Phys.* **1981**, *14*, 3805.
- (18) Schirmer, J.; Cederbaum, L. S.; Walter, O. New approach to the one-particle Green's function for finite Fermi systems. *Phys. Rev. A: At, Mol., Opt. Phys.* **1983**, *28*, 1237–1259.
- (19) Ortiz, J. A nondiagonal, renormalized extension of partial third-order quasiparticle theory: Comparisons for closed-shell ionization energies. *J. Chem. Phys.* **1998**, *108*, 1008–1015.
- (20) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry*; Dover Publications, 1996.
- (21) Holleboom, L. J.; Snijders, J. G. A comparison between the Møller-Plesset and Green's function perturbative approaches to the calculation of the correlation energy in the many-electron problem. *J. Chem. Phys.* **1990**, *93*, 5826.
- (22) Phillips, J. J.; Zgid, D. Communication: The description of strong correlation within self-consistent Green's function second-order perturbation theory. *J. Chem. Phys.* **2014**, *140*, 241101.
- (23) Kananenka, A. A.; Phillips, J. J.; Zgid, D. Efficient Temperature-Dependent Green's Functions Methods for Realistic Systems: Compact Grids for Orthogonal Polynomial Transforms. *J. Chem. Theory Comput.* **2016**, *12*, 564–571.
- (24) Stanton, J. F.; Gauss, J. Analytic energy derivatives for ionized states described by the equation-of-motion coupled cluster method. *J. Chem. Phys.* **1994**, *101*, 8938.
- (25) Krylov, A. I. Equation-of-motion coupled-cluster methods for open-shell and electronically excited species: the Hitchhiker's guide to Fock space. *Annu. Rev. Phys. Chem.* **2008**, *59*, 433–462.
- (26) Nooijen, M.; Snijders, J. G. Coupled cluster approach to the single-particle Green's function. *Int. J. Quantum Chem.* **1992**, *44*, 55–83.
- (27) Nooijen, M.; Snijders, J. G. Coupled cluster Green's function method: Working equations and applications. *Int. J. Quantum Chem.* **1993**, *48*, 15–48.
- (28) McClain, J.; Lischner, J.; Watson, T.; Matthews, D. A.; Ronca, E.; Louie, S. G.; Berkelbach, T. C.; Chan, G. K.-L. Spectral functions of the uniform electron gas via coupled-cluster theory and comparison to the GW and related approximations. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2016**, *93*, 235139.
- (29) Bhaskaran-Nair, K.; Kowalski, K.; Shelton, W. A. Coupled cluster Green function: Model involving single and double excitations. *J. Chem. Phys.* **2016**, *144*, 144101.
- (30) Peng, B.; Kowalski, K. Coupled-cluster Green's function: Analysis of properties originating in the exponential parametrization of the ground-state wave function. *Phys. Rev. A: At, Mol., Opt. Phys.* **2016**, *94*, 062512.
- (31) Peng, B.; Kowalski, K. Properties of advanced coupled-cluster Green's function. *Mol. Phys.* **2018**, *116*, 561–569.
- (32) Nishi, H.; Kosugi, T.; Furukawa, Y.; Matsushita, Y.-i. Quasiparticle energy spectra of isolated atoms from coupled-cluster singles and doubles (CCSD): Comparison with exact CI calculations. *J. Chem. Phys.* **2018**, *149*, 034106.
- (33) Furukawa, Y.; Kosugi, T.; Nishi, H.; Matsushita, Y. I. Band structures in coupled-cluster singles-and-doubles Green's function (GFCCSD). *J. Chem. Phys.* **2018**, *148*, 204109.
- (34) Kosugi, T.; Nishi, H.; Furukawa, Y.; Matsushita, Y.-i. Comparison of Green's functions for transition metal atoms using self-energy functional theory and coupled-cluster singles and doubles (CCSD). *J. Chem. Phys.* **2018**, *148*, 224103.
- (35) Ou, Q.; Subotnik, J. E. Comparison between GW and Wave-Function-Based Approaches: Calculating the Ionization Potential and Electron Affinity for 1D Hubbard Chains. *J. Phys. Chem. A* **2016**, *120*, 4514.
- (36) Knight, J. W.; Wang, X.; Gallandi, L.; Dolgounitcheva, O.; Ren, X.; Ortiz, J. V.; Rinke, P.; Kördörfer, T.; Marom, N. Accurate Ionization Potentials and Electron Affinities of Acceptor Molecules III: A Benchmark of GW Methods. *J. Chem. Theory Comput.* **2016**, *12*, 615–626.
- (37) Rangel, T.; Hamed, S. M.; Bruneval, F.; Neaton, J. B. Evaluating the GW Approximation with CCSD(T) for Charged Excitations Across the Oligoacenes. *J. Chem. Theory Comput.* **2016**, *12*, 2834–2842.
- (38) Krause, K.; Harding, M. E.; Klopper, W. Coupled-cluster reference values for the GW27 and GW100 test sets for the assessment of GW methods. *Mol. Phys.* **2015**, *113*, 1952.
- (39) Caruso, F.; Dauth, M.; van Setten, M. J.; Rinke, P. Benchmark of GW Approaches for the GW100 Test Set. *J. Chem. Theory Comput.* **2016**, *12*, 5076.
- (40) van Setten, M. J.; Caruso, F.; Sharifzadeh, S.; Ren, X.; Scheffler, M.; Liu, F.; Lischner, J.; Lin, L.; Deslippe, J. R.; Louie, S. G.; Yang, C.; Weigend, F.; Neaton, J. B.; Evers, F.; Rinke, P. GW100: Benchmarking G0W0 for Molecular Systems. *J. Chem. Theory Comput.* **2015**, *11*, 5665–5687.
- (41) Marsman, M.; Grüneis, A.; Paier, J.; Kresse, G. Second-order Møller-Plesset perturbation theory applied to extended systems. I. Within the projector-augmented-wave formalism using a plane wave basis set. *J. Chem. Phys.* **2009**, *130*, 184103.
- (42) Grüneis, A.; Marsman, M.; Kresse, G. Second-order Møller-Plesset perturbation theory applied to extended systems. II. Structural and energetic properties. *J. Chem. Phys.* **2010**, *133*, 074107.
- (43) Grüneis, A.; Booth, G. H.; Marsman, M.; Spencer, J.; Alavi, A.; Kresse, G. Natural orbitals for wave function based correlated calculations using a plane wave basis set. *J. Chem. Theory Comput.* **2011**, *7*, 2780–2785.
- (44) Del Ben, M.; Hutter, J.; Vandevondele, J. Second-order moller-plesset perturbation theory in the condensed phase: An efficient and massively parallel gaussian and plane waves approach. *J. Chem. Theory Comput.* **2012**, *8*, 4177–4188.
- (45) Grüneis, A.; Shepherd, J. J.; Alavi, A.; Tew, D. P.; Booth, G. H. Explicitly correlated plane waves: Accelerating convergence in periodic wavefunction expansions. *J. Chem. Phys.* **2013**, *139*, 084112.
- (46) Grüneis, A. A coupled cluster and Møller-Plesset perturbation theory study of the pressure induced phase transition in the LiH crystal. *J. Chem. Phys.* **2015**, *143*, 102817.
- (47) Liao, K.; Grüneis, A. Communication: Finite size correction in periodic coupled cluster theory calculations. *J. Chem. Phys.* **2016**, *145*, 141102.
- (48) Hummel, F.; Tsatsoulis, T.; Grüneis, A. Low rank factorization of the Coulomb integrals for periodic coupled cluster theory. *J. Chem. Phys.* **2017**, *146*, 124105.
- (49) McClain, J.; Sun, Q.; Chan, G. K.-L.; Berkelbach, T. C. Gaussian-Based Coupled-Cluster Theory for the Ground State and Band Structure of Solids. *J. Chem. Theory Comput.* **2017**, *13*, 1209–1218.
- (50) Fetter, A. L.; Walecka, J. D. *Quantum Theory of Many-Particle Systems*; Dover Publications, 2003.
- (51) Pavlyukh, Y.; Hübner, W. Configuration interaction approach for the computation of the electronic self-energy. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2007**, *75*, 205129.
- (52) Bruneval, F. Ionization energy of atoms obtained from GW self-energy or from random phase approximation total energies. *J. Chem. Phys.* **2012**, *136*, 194107.
- (53) van Setten, M. J.; Weigend, F.; Evers, F. The GW-Method for Quantum Chemistry Applications: Theory and Implementation. *J. Chem. Theory Comput.* **2013**, *9*, 232–246.
- (54) Bruneval, F.; Rangel, T.; Hamed, S. M.; Shao, M.; Yang, C.; Neaton, J. B. MOLGW 1: Many-body perturbation theory software for atoms, molecules, and clusters. *Comput. Phys. Commun.* **2016**, *208*, 149–161.

- (55) Emrich, K. An Extension of the Coupled Cluster Formalism To Excited-States 0.2. Approximations and Tests. *Nucl. Phys. A* **1981**, *351*, 397–438.
- (56) Berkelbach, T. C. Communication: Random-phase approximation excitation energies from approximate equation-of-motion coupled-cluster doubles. *J. Chem. Phys.* **2018**, *149*, 041103.
- (57) Löwdin, P.-O. Studies in perturbation theory. *J. Mol. Spectrosc.* **1963**, *10*, 12–33.
- (58) Golubeva, A. A.; Pieniazek, P. A.; Krylov, A. I. A new electronic structure method for doublet states: Configuration interaction in the space of ionized 1h and 2h1p determinants. *J. Chem. Phys.* **2009**, *130*, 124113.
- (59) Ren, X.; Marom, N.; Caruso, F.; Scheffler, M.; Rinke, P. Beyond the GW approximation: A second-order screened exchange correction. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2015**, *92*, 081104.
- (60) Maggio, E.; Liu, P.; Van Setten, M. J.; Kresse, G. GW100: A Plane Wave Perspective for Small Molecules. *J. Chem. Theory Comput.* **2017**, *13*, 635–648.
- (61) Govoni, M.; Galli, G. GW100: Comparison of Methods and Accuracy of Results Obtained with the WEST Code. *J. Chem. Theory Comput.* **2018**, *14*, 1895–1909.
- (62) Dutta, A. K.; Vaval, N.; Pal, S. Performance of the EOMIP-CCSD (2) method for determining the structure and properties of doublet radicals: A benchmark investigation. *J. Chem. Theory Comput.* **2013**, *9*, 4313–4331.
- (63) Dutta, A. K.; Gupta, J.; Pathak, H.; Vaval, N.; Pal, S. Partitioned EOMEA-MBPT(2): An efficient  $N^5$  scaling method for calculation of electron affinities. *J. Chem. Theory Comput.* **2014**, *10*, 1923–1933.
- (64) Dutta, A. K.; Vaval, N.; Pal, S. Lower scaling approximation to EOM-CCSD: A critical assessment of the ionization problem. *Int. J. Quantum Chem.* **2018**, *118*, e25594.
- (65) Goings, J. J.; Caricato, M.; Frisch, M. J.; Li, X. Assessment of low-scaling approximations to the equation of motion coupled-cluster singles and doubles equations. *J. Chem. Phys.* **2014**, *141*, 164116.
- (66) Rishi, V.; Perera, A.; Nooijen, M.; Bartlett, R. J. Excited states from modified coupled cluster methods: Are they any better than EOM CCSD? *J. Chem. Phys.* **2017**, *146*, 144104.
- (67) Nooijen, M.; Snijders, J. G. Second order many-body perturbation approximations to the coupled cluster Green's function. *J. Chem. Phys.* **1995**, *102*, 1681.
- (68) Gwaltney, S. R.; Nooijen, M.; Bartlett, R. J. Simplified methods for equation-of-motion coupled-cluster excited state calculations. *Chem. Phys. Lett.* **1996**, *248*, 189–198.
- (69) Stanton, J. F.; Gauss, J. Perturbative treatment of the similarity transformed Hamiltonian in equation-of-motion coupled-cluster approximations. *J. Chem. Phys.* **1995**, *103*, 1064.
- (70) Christiansen, O.; Koch, H.; Jorgensen, P. The second-order approximate coupled cluster singles and doubles model CC2. *Chem. Phys. Lett.* **1995**, *243*, 409–418.
- (71) Byrd, J. N.; Rishi, V.; Perera, A.; Bartlett, R. J. Approximating electronically excited states with equation-of-motion linear coupled-cluster theory. *J. Chem. Phys.* **2015**, *143*, 164103.
- (72) Booth, G. H.; Grüneis, A.; Kresse, G.; Alavi, A. Towards an exact description of electronic wavefunctions in real solids. *Nature* **2013**, *493*, 365–370.
- (73) Negele, J. W.; Orland, H. *Quantum Many-Particle Systems*; Perseus Books, 1998.
- (74) Wälz, G.; Usvyat, D.; Korona, T.; Schütz, M. A hierarchy of local coupled cluster singles and doubles response methods for ionization potentials. *J. Chem. Phys.* **2016**, *144*, 084117.
- (75) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.
- (76) Sun, Q.; Berkelbach, T. C.; Blunt, N. S.; Booth, G. H.; Guo, S.; Li, Z.; Liu, J.; McClain, J. D.; Sayfutyarova, E. R.; Sharma, S.; Wouters, S.; Chan, G. K. L. PySCF: the Python-based simulations of chemistry framework. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*, e1340.
- (77) Bartlett, R. J.; Musial, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291–352.
- (78) Richard, R. M.; Marshall, M. S.; Dolgounitcheva, O.; Ortiz, J. V.; Brédas, J. L.; Marom, N.; Sherrill, C. D. Accurate Ionization Potentials and Electron Affinities of Acceptor Molecules I. Reference Data at the CCSD(T) Complete Basis Set Limit. *J. Chem. Theory Comput.* **2016**, *12*, 595–604.
- (79) Schütz, M.; Manby, F. R. Linear scaling local coupled cluster theory with density fitting. Part I: 4-external integrals. *Phys. Chem. Chem. Phys.* **2003**, *5*, 3349–3358.
- (80) Hohenstein, E. G.; Kokkila, S. I.; Parrish, R. M.; Martínez, T. J. Tensor hypercontraction equation-of-motion second-order approximate coupled cluster: Electronic excitation energies in  $O(N^4)$  time. *J. Phys. Chem. B* **2013**, *117*, 12972–12978.
- (81) Binosi, D.; Theußl, L. JaxoDraw: A graphical user interface for drawing Feynman diagrams. *Comput. Phys. Commun.* **2004**, *161*, 76–86.