Stochastic vertex corrections: $GW\Gamma_X$ approach for accurate quasiparticle energies

Vojtěch Vlček

Department of Chemistry and Biochemistry, University of California, Santa Barbara, 93106, U.S.A

E-mail: vlcek@ucsb.edu

Abstract

Quasiparticle energies for the highest occupied and lowest unoccupied states are investigated using the stochastic many-body perturbation theory with the inclusion of non-local vertex corrections. Approximate form of a vertex function, labeled $G_0W_0^{tc}\Gamma_X$, is developed and implemented within a real-time stochastic framework. This implementation is tested on acene molecules, C$_{60}$ and PC$_{60}$BM. $G_0W_0^{tc}\Gamma_X$ is compared to the one-shot stochastic $GW$ approach with and without random-phase approximation. Overall, the $G_0W_0^{tc}\Gamma_X$ method has the best agreement with CCSD(T) benchmark data. The vertex correction appears crucial for the description of unoccupied states. All three stochastic methods scale linearly; due to favorable self-averaging in large systems, stochastic $G_0W_0^{tc}$ is actually less expensive than stochastic $G_0W_0$.

1 Introduction

Efficient first principles methods allow calculations of the ground state electronic structure in large molecules and solids. However, quantitative prediction of electronic excitation is still computationally prohibitive. Traditional quantum chemistry methods, such as configurational interaction or coupled cluster approaches scale at least as $N_e^6$ (where $N_e$ is the number of electrons). As a result, these methods are applied only to small systems.

Many-body perturbation theory offers an alternative and becomes an increasingly popular tool for computation of quasiparticle (QPs) energies of molecules. The central quantity is the QP self-energy, i.e., a dynamical potential that embodies all many-body interactions. In principle, it is found in a self-consistent manner. The self-consistency relates the self-energy to the QP Green’s function, polarizability and screened Coulomb interaction.

In practice, the expression for the self-energy is often simplified by neglecting high-order terms (non-trivial part of vertex function $\Gamma$) leading to the $GW$ approximation. In addition, self-consistency in $GW$ calculations is either further approximated or completely avoided. The latter thus corresponds to a one-shot correction, labeled $G_0W_0$, on top a mean-field starting point (usually DFT). Such practical simplification of $GW$ has two purposes: (i) Conventional implementations scale as $N_e^3$ or $N_e^4$ and repeated evaluation of the self-energy is thus costly even for small systems. (ii) Self-consistent $GW$ may yield worse results than $G_0W_0$ due to the absence of the vertex term. The typical strategy is thus to use $G_0W_0$ on top of the “best” possible DFT starting point. Recent benchmark for acenes however revealed that GW suffers from substantial errors for QP energies of unoccupied states.

Beyond $GW$ techniques include approximate vertex functions ($\Gamma$), which are closely related to the electron-hole interaction kernel in the
Bethe-Salpeter equation (BSE). In practice, \( \Gamma \) is computed in various ways: Local vertex functions derived from the Kohn-Sham time-dependent density functional theory (TDDFT) are simple and relatively inexpensive, but they do not remedy failures of GW, such as spurious “self-screening error”. Furthermore, they do not outperform simple \( G_0 W_0 \).

Non-local vertex functions seem to improve the description of the QP energies; however, they are costly and suffer from steep scaling \( (N_e^6) \). Alternatively, the vertex term has been approximated up to the second order, but this is associated with only mild cost reduction \( (N_e^5 \) scaling). Consequently, the beyond GW calculations have been applied only to model or few-electron systems.

Here, numerical and theoretical developments are combined to overcome this limitation. A self-consistent expression for the self-energy with non-local \( \Gamma \) is constructed using derivatives of the inverse Green’s function. In practice, we apply only a one-shot correction, in which \( \Gamma \) is derived from the non-local DFT exchange term in the generalized Kohn-Sham (GKS) scheme. The approach is labeled GWT\(_X\).

To lower the computational cost, GWT\(_X\) is implemented using real-time stochastic numerical techniques. Up to now, stochastic calculations of QP energies were limited to \( G_0 W_0 \) with DFT based on the local density approximation (LDA) to exchange and correlation (xc). Here, we first extend the methodology to hybrid xc functionals. Next, the stochastic form of GWT\(_X\) is presented and tested on a set of large molecules.

The stochastic implementation scales (sub)linearly with the number of electrons. Furthermore, favorable self-averaging leads to low statistical noise. As a result, the GWT\(_X\) method is found to be computationally less expensive than the stochastic implementation of \( G_0 W_0 \) based on hybrid xc functionals. Results for ionization potentials and electron affinities of molecules show that the inclusion of a non-local vertex is necessary for accurate predictions of QP energies.

The manuscript is organized as follows: Derivation of the self-energy expressions is presented in Sec. 2. The stochastic formulation suitable for numerical implementation is shown in Sec. 3. Performance of the method and its implementation are demonstrated in the results section (Sec. 4) followed by conclusions and outlook (Sec. 5).

2 Theory

In this section, we first review the theoretical description of quasiparticles (QP), namely quasi-electron and quasi-holes. Propagation of a QP is described by a single-particle Green’s functions (GF), which is defined as

\[
iG(1,2) = \langle \psi | T \psi(1) \hat{\psi}^\dagger(2) | \psi \rangle ,
\]

where \( \psi \) is the ground state many-body wave-function of the \( N_e \)-electron system, \( T \) denotes time ordering operator, \( \hat{\psi} \) and \( \hat{\psi}^\dagger \) are the electron annihilation and creation operators. Here, we adopt a short-hand notation for the space-time coordinates: \( (r_1, t_1) \equiv 1 \).

The GF satisfies the equation of motion

\[
\left[ i \frac{\partial}{\partial t} - \hat{h} \right] G(1,2) - \Sigma_T(1,3)G(3,2) = \delta(r_1-r_2)
\]

where \( \hat{h} \) contains the kinetic energy and the electron-nuclear attraction terms, and \( \Sigma_T \) is the total self-energy, which contains both Hartree and exchange-correlation interactions. Further, we simplified the notation by omitting integration symbol and introducing a bar symbol above the space-time coordinates that should be integrated.

In this work, we focus on QP energies of quasi-electrons and quasi-holes (\( \varepsilon \)), which are obtained from the QP equation:

\[
\hat{h}\psi(r_1) + \int \Sigma_T(r_1, r_2, \omega = \varepsilon) \psi(r_2) dr_2 = \varepsilon \psi(r_1)
\]

where \( \psi \) is the QP state. Eq. 3 is a fixed point expression where \( \Sigma_T \) has to be computed at the frequency corresponding to \( \varepsilon \). In the rest of the paper, we use a real-time representation of the self-energy, \( \Sigma_T(r_1, r_2, t_1, t_2) \) which is merely a...
Fourier transformation of $\Sigma_T(r_1, r_2, \omega)$.

We will now review the expressions for the total self-energy in Sec. 2.1 and then we turn to the approximate forms in Sec. 2.2.

2.1 Self-energy

The total self-energy ($\Sigma_T$) in principle requires knowledge of the two-particle GF, leading to a hierarchy of coupled equations of motion. Alternatively, $\Sigma_T$ is written as a sum:

$$\Sigma_T(1, 2) = \Sigma_H(1)\delta(1, 2) + \Sigma_{xc}(1, 2), \quad (4)$$

where $\Sigma_H$ and $\Sigma_{xc}$ are the Hartree and exchange-correlation self-energies. Note that the former is local and instantaneous; hence it appears together with a delta function $\delta(1, 2) \equiv \delta(r_1 - r_2)\delta(t_1 - t_2)$.

The Hartree term represents the interaction with the electron density:

$$\Sigma_H(1) = -i\nu(1, 2)G(2, 2^+) \quad (5)$$

where $\nu$ is the instantaneous Coulomb interaction defined as

$$\nu(1, 2) = \frac{1}{|r_1 - r_2|}\delta(t_1 - t_2) \quad (6)$$

and the density is given by the equal-time GF, i.e., $n(r_1) \equiv G(1, 1^+)$. The $1^+$ argument represents $(r_1, t_1^+)$, where $t_1^+$ is only infinitesimally after $t_1$.

The exchange-correlation self-energy is:

$$\Sigma_{xc}(1, 2) = -i\nu(1, 4)G(1, 3)\frac{\delta G^{-1}(3, 2)}{\delta U(4)}, \quad (7)$$

where $U$ is an external potential introduced to remove the two-particle GF in the expression for $\Sigma_T$. The derivative of the inverse GF in Eq. 7 leads to two equivalent expressions: The first one is very compact and includes a three-point irreducible vertex function $\Gamma$; the second is slightly more involved, but it is more versatile expression leading to useful approximations to $\Gamma$. For completeness, we will now review both.

2.1.1 $\Sigma_{xc}$ with the vertex function

In the first route, we consider a chain rule of derivatives:

$$\frac{\delta G^{-1}(3, 2)}{\delta U(4)} = \frac{\delta G^{-1}(3, 2)}{\delta U_{cl}(5)} \frac{\delta U_{cl}(5)}{\delta U(4)}, \quad (8)$$

where $U_{cl}$ is a classical potential, consisting of the Coulomb and external potentials. From classical electrostatics, the change of $U_{cl}$ with the variation of the external potential $U$ corresponds to the inverse dielectric function:

$$\epsilon^{-1}(5, \bar{4}) \equiv \frac{\delta U_{cl}(5)}{\delta U(4)}. \quad (9)$$

The first derivative in the right side of Eq. 8 serves as the definition of the irreducible vertex function:

$$\Gamma(\bar{3}, 2, \bar{5}) := -\frac{\delta G^{-1}(3, 2)}{\delta U_{cl}(5)}. \quad (10)$$

Combination of Eqs. 7, 9 and 10 leads to the following compact expression for the exchange-correlation self-energy:

$$\Sigma_{xc}(1, 2) = iW(1, \bar{4})G(1, \bar{3})\Gamma(\bar{4}, \bar{3}, 2), \quad (11)$$

where $W$ is the screened Coulomb interaction obtained by convolution of Eqs. 6 and 9.

2.1.2 $\Sigma_{xc}$ with generalized polarizability

In the second route, we start again from Eq. 7 and make use of the functional derivatives. First, the change of the inverse GF with respect to $U$ is:

$$\Sigma_{xc}(1, 2) = i\nu(1, 4)G(1, 3) \left[ \delta(3, 2)\delta(\bar{3}, \bar{4}) + \frac{\delta \Sigma_T(\bar{3}, 2)}{\delta U(4)} \right]. \quad (12)$$

The two terms in the square brackets lead to a suitable definition of the exchange and polarization self-energies. The former is

$$\Sigma_x(1, 2) = i\nu(1, 2)G(1, 2). \quad (13)$$
Note that the Coulomb kernel is instantaneous (Eq. [6]), so \( G(1, 2) \) in Eq. [13] is the density matrix \( \rho(1, 2) \equiv G(1, 2)\delta(t_1 - t_2) \).

The polarization self-energy has a more complicated expression:

\[
\Sigma_P(1, 2) = i\nu(1, 4)G(1, 3)\frac{\delta\Sigma_T(3, 2)}{\delta U(4)}. \tag{14}
\]

It is convenient to recast the functional derivative of the total self-energy in Eq. [14] as

\[
\frac{\delta\Sigma_T(3, 2)}{\delta U(4)} = \frac{\delta\Sigma_T(3, 2)}{\delta G(6, 5)} \frac{\delta G(6, 5)}{\delta U(4)}. \tag{15}
\]

Further, we introduce a generalized three-point reducible polarizability:

\[
^3\chi(\bar{6}, 5, 4) := -\frac{\delta G(\bar{6}, 5)}{\delta U(4)}. \tag{16}
\]

The final expression of the polarization self-energy thus reads

\[
\Sigma_P(1, 2) = -i\nu(1, 4)G(1, 3)\frac{\delta\Sigma_T(3, 2)}{\delta G(6, 5)}^3\chi(\bar{6}, 5, 4). \tag{17}
\]

The compact expression with the vertex function (Eq. [11]) is equivalent to the sum of the exchange (Eq. [13]) and polarization (Eq. [17]) self-energies. Note that \( \Sigma_P \) depends on a functional derivative of \( \Sigma_T \) (Eq. [17]); hence, the total self-energy should be, in principle, found by a self-consistent cycle since.

### 2.2 Approximate Self-energy

Due to the quadruple integration, Eq. [17] is computationally difficult. In the following, we outline how to construct practical expressions for \( \Sigma_P \) based on successive approximations for the functional derivative \( \delta\Sigma_T/\delta G \).

#### 2.2.1 The GW approximation

In the GW approximation, the total-self-energy in Eq. [17] is substituted with the classical Hartree self-energy \( \Sigma_H \). Hence, the functional derivative of \( \Sigma_T \) with respect to the GF becomes

\[
\frac{\delta\Sigma_T(3, 2)}{\delta G(6, 5)} \approx -\nu(2, 5)\delta(6, 5)\delta(3, 2), \tag{18}
\]

where the delta functions are due to the locality of \( \Sigma_H \) (Eq. [5]).

The approximation introduced in Eq. [18] greatly simplifies the polarization self-energy: \( ^3\chi \) becomes two-point reducible polarizability, i.e.,

\[
\chi(5, 4) := -\frac{\delta G(\bar{5}, \bar{5}^+)}{\delta U(4)} = \frac{\delta n(\bar{5})}{\delta U(4)}. \tag{19}
\]

Note the \( \chi \) is a time-ordered quantity, but it is trivially related to the standard retarded response function \( \chi \). Consequently, the polarization self-energy has the following form:

\[
\Sigma_P^{GW}(1, 2) = i\nu(1, 4)G(1, 2)\nu(2, 5)\chi(\bar{5}, 4). \tag{20}
\]

The GW xc self-energy is a sum of Eqs. [13] and [20] which becomes

\[
\Sigma_{xc}^{GW}(1, 2) = iW(1, 2^+)G(1, 2), \tag{21}
\]

where we used an alternative definition of the screened Coulomb interaction:

\[
W(1, 2) = \nu(1, 2) + \nu(1, 3)\chi(\bar{3}, \bar{4})\nu(\bar{4}, \bar{2}). \tag{22}
\]

Note that the GW self-energy is often obtained from Eq. [11] by approximating the vertex function as \( \Gamma(1, 2, 3) \approx \delta(1, 2)\delta(1, 3) \). Such derivation of \( \Sigma_{xc}^{GW} \) is quite simple and compact, but it is not immediately clear how to construct a better approximation.

#### 2.2.2 GW\( \Gamma_X \) approximation

In this part, we will consider the next step in the construction of the self-energy. We use Eq. [17] and take \( \Sigma_T \approx \Sigma_H + \Sigma_x \). We denote this approximation \( GW\Gamma_X \). The derivative of the total self-energy becomes:

\[
\frac{\delta\Sigma_T(3, 2)}{\delta G(6, 5)} \approx
-\nu(2, 5)\delta(6, 5)\delta(3, 2) + \nu(\bar{3}, 2)\delta(\bar{5}, 2)\delta(\bar{6}, \bar{3}) \tag{23}
\]
The first term on the right is the classical Hartree interaction (as in Eq. 18), the second term is due to non-local exchange. While both terms include the Coulomb kernel (Eq. 6), they have a different structure, i.e., each contracts distinct space-time points. As a result, the polarization self-energy contains a contribution from the reducible three-point polarizability:

$$\Sigma_{P}^{GW}X(1, 2) = i\nu(1, 4)G(1, 2)\times$$

$$\left[\nu(2, 5)\chi(5, 4) - \nu(2, 3)\chi(3, 2, 4)\right].$$

(24)

The definition of $\nu$ in Eq. 9 contains a delta function which guarantees that the Coulomb interaction is instantaneous in time. Hence, the second term in the square brackets in Eq. 24 implicitly contains $\delta(t_2 - t_3)$. As a result, the generalized polarizability depends on three spatial coordinates but only two time points. In other words, $\chi(3, 2, 4)$ (defined by Eq. 16) yields induced density matrix $\delta\rho(\mathbf{r}_3, \mathbf{r}_2, t_3, t_2)\delta(t_2 - t_3)$ due to the variation of the external potential at $\mathbf{r}_4$. In contrast to the RPA-SOSEX approach, the GWT_X self-energy includes the generalized three-point polarizability and it is not based on RPA response function.

Note that the GWT_X self-energy is an approximation to the irreducible vertex correction (Eq. 10). If $\Sigma_{xc}$ is computed on top of a Hartree-Fock starting point, the sum of Eqs. 18 and 24 is equivalent to the GWT self-energy (Eq. 11) in the first iteration. Such one-shot GWT on top of HF was applied recently to finite systems. It was limited to small molecules. However, an affordable stochastic implementation of Eq. 24 is presented in the next section.

3 Computational methodology

In this section, we present practical steps which allow application of Eqs. 20 and 24 to large molecules. In practice, we employ two simplifications:

(I) We do not seek self-consistency in $\Sigma_T$. The $i^{th}$ QP energy is computed as:

$$\varepsilon_i = \varepsilon_i^{DFT} + \left< \phi_i \left| \tilde{\Sigma}(\omega = \varepsilon_i) - \delta_{xc} \right| \phi_i \right>,$$

(25)

where $\varepsilon_i^{DFT}$ is the DFT eigenvalue, $\phi_i$ is the corresponding DFT eigenstate and $\delta_{xc}$ is the mean-field DFT exchange-correlation potential. The one-shot correction means that the self-energy is constructed from the underlying mean-field Hamiltonian with non-local exchange (see below, Eq. 26). Hence, the GF and the screened Coulomb interaction are expressed using DFT eigenstates and they are denoted $G_0$ and $W_0$.

(II) We reformulate Eqs. 21 and 24 using the stochastic approach, i.e., the expectation values become statistical estimators over (many) stochastic samples. This method was applied to the $G_0W_0$ approximation as described in earlier publications. In contrast to the previous work, the underlying DFT is no longer constrained to local xc functionals. Further, the stochastic formulation of $\Sigma_{P}^{GW}X$ is a new development.

The details of the underlying DFT are given in the next subsection (3.1), followed by a short overview of the stochastic approach and the description of the new developments (3.2).

3.1 DFT starting point with the long-range exchange interaction

The ground state electronic structure is computed with optimally tuned range-separated xc functional, with Hamiltonian:

$$\hat{H}_0 = \hat{h} + \hat{\Sigma}_H + \hat{V}_c + \hat{V}_x + \hat{\Sigma}_x$$

(26)

where $\hat{h}$ contains the kinetic energy and the electron-nuclear attraction (as in Eq. 2), and $V_c$ is a correlation potential approximated by a semilocal functional of the density. The two last terms, $V_x$ and $\Sigma_x^\gamma$ represent semilocal and non-local exchange. These two terms are derived from a partitioned Coulomb kernel $1/r = \text{erf}(\gamma r)/r + \text{erfc}(\gamma r)/r$, where $\gamma$ is the range-separation parameter. The long-
range non-local exchange interaction is
\[\Sigma_x'(1, 2) = i\nu\gamma(1, 2)G_0(1, 2),\]  
(27)
where \(G_0\) is the GKS Green’s function (constructed from the \(H_0\) eigenstates) and
\[\nu\gamma(1, 2) = \frac{\text{erf}(\gamma r)}{|r_1 - r_2|}\delta(t_1 - t_2)\]  
(28)
The short-range part, \(V_s\), is derived from the complementary error-function term; it is given by a semilocal density functional which depends on the value of \(\gamma\).

For the calculations presented in this work, we chose optimally tuned LC-ωPBE functionals\(^{[62]}\) implemented within the LibXC library\(^{[53,54]}\). In practice, optimal tuning amounts to finding range-separation parameter \(\gamma\) which enforces the IP theorem, i.e., \(\gamma\) is varied such that the negative of the HOMO energy corresponds to the ionization potential (energy difference between a neutral system and a cation). Optimal tuning is associated with mitigation of spurious electron self-interaction and leads to good \(I\) and fundamental band gaps \(E_g\) in finite systems\(^{[65–67]}\). Further, TDDFT with optimally tuned functionals with long-range exchange treats attractive electron-hole interactions and efficiently mimics BSE\(^{[53,57,58,71]}\).

### 3.2 Stochastic approach

We will now introduce the basics of the stochastic approach and describe how \(\Sigma^{GWT}_x\) and \(\Sigma^{GW}\) are computed. In the initial part of the algorithm, random functions are prepared on a real space grid as:
\[\bar{\zeta}(r) = \pm \frac{1}{\sqrt{d\Omega}}\]  
(29)
where \(d\Omega\) is the volume element associated with each grid point. The ± in front of the fraction represents a random sign assigned to each space-point \(r\). This choice satisfies the stochastic resolution of identity \(\hat{I} \equiv \{|\bar{\zeta}\rangle\langle \bar{\zeta}|\}\), where \(\hat{I}\) is the identity operator and \(\{\ldots\}\) denotes an average over the entire (in principle infinite) set of random functions.

In the stochastic representation, the density matrix is given as an average:
\[\rho(r_1, r_2) = \{\eta(r_1)\eta^*(r_2)\}.\]  
(30)
Here, \(\eta\) are random vectors within the occupied subspace, i.e., \(|\eta\rangle = \bar{P}|\bar{\zeta}\rangle\) and \(\bar{P}\) is a projection operator. \(\bar{P}\) depends on the chemical potential and the Hamiltonian \(H_0\). In practice, \(\eta\) states can be constructed either by projecting on the occupied eigenstates or, e.g., by Chebyshev filtering\(^{[53,57,58,71]}\)\(^{53,57,58,72}\). In this paper, we follow the former approach.

The stochastic form of the GKS Green’s function \(G_0\) is
\[iG_0(r_1, r_2, t_1 - t_2) = \{\zeta(r_1, t_1)\zeta^*(r_2, t_2)\},\]  
(31)
where the \(\zeta\) vector is either in the occupied or unoccupied subspace, i.e., it is obtained by projection with \(\bar{P}\) or its complement \(\left(\hat{I} - \bar{P}\right)\). Since the equilibrium GF depends only on the difference between \(t_1\) and \(t_2\), we set \(t_1 = 0\) and let only the projected stochastic vectors to evolve in time. For negative/positive times, the GF represents a propagator of holes/electrons. The corresponding time-evolved random vectors are:
\[\zeta(r, t) = \begin{cases} \begin{align*} \langle r | e^{-iH_0t} \bar{P} | \bar{\zeta} \rangle & \quad t < 0 \\ \langle r | e^{-iH_0t} \left(\hat{I} - \bar{P}\right) | \bar{\zeta} \rangle & \quad t > 0. \end{align*} \]  
(32)
In practice, the time propagation is performed using Trotter (split operator) technique in which we introduce another set of stochastic orbitals, as described in Sec. 3.2.2 and in Refs. 53,57,58,72.

We will now focus on the two approximations introduced earlier (Eqs. 20 and 24) and combine them with the stochastic form of the GF (Eq. 31).

The GW approximation to \(\Sigma_P\) is:
\[\Sigma^{GW}_P(t) = \{\langle \phi_i | \zeta(t) | \hat{\nu} \zeta(t) | \hat{\nu} \phi_i \rangle\} = \{\langle \phi_i | \zeta(t) | \hat{\nu} \delta n(t) | \zeta \phi_i \rangle\}.\]  
(33)
In the first line, the operators \(\hat{\nu}\) and \(\hat{\chi}\) represent the Coulomb interaction (Eq. 6) and the two-
point polarizability (Eq. 19). In Eq. 34, the notation is simplified by introducing \( \delta \hat{n}(t) = \hat{\chi}(t) \hat{\nu} \), which corresponds to the induced time-dependent density at time \( t \) due to \( \hat{\chi}(r) \phi_i(r) \) (cf. Refs. 53, 57, 58). This step provides an intuitive interpretation of \( \Sigma_{GW}^{T_X} \): it is a time-dependent induced Coulomb potential due to the addition of an electron/hole to the state \( \phi \).

The polarization self-energy in the \( GW_T \) approximation is based on Eqs. 23 and 24. Here, only long-range exchange self-energy is considered, i.e., \( \Sigma_T \approx \Sigma_H + \Sigma_x^\lambda \). As a result

\[
\Sigma_{GW}^{T_X}(t) = \{ (\phi_i \hat{\zeta}(t)(\hat{\nu} \delta \hat{n}(t) - \hat{\nu}^\dagger \delta \hat{\rho}(t)) | \hat{\zeta} \phi_i) \}
\]

(35)

where \( \nu^\dagger \) (Eq. 28) is explicitly included. We also introduce an operator \( \delta \hat{\rho}(t) = \hat{\chi}(t) \hat{\nu} \), which corresponds to the induced time-dependent density at time \( t \) due to \( \hat{\chi}(r) \phi_i(r) \). For details see text after Eq. 24. Hence, \( \Sigma_{GW}^{T_X} \) is a time-dependent induced Coulomb and exchange potentials due to the addition of an electron/hole to the state \( \phi \).

In both approximations to the self-energy (\( \Sigma_{GW}^{T} \) and \( \Sigma_{GW}^{T_X} \)), the induced potentials are time-ordered quantities. In practice, we calculate them through retarded propagations of stochastic states (described below) and the time-ordering is performed in the frequency domain using sparse stochastic compression technique (see details in [1]).

### 3.2.1 Stochastic induced time-dependent density and density matrix

In this part, we describe how we evaluate both \( \delta \hat{\rho} \) and \( \delta \hat{n} \) using stochastic states. Since the density matrix is not constructed from eigenstates of \( \hat{H}_0 \), it will naturally fluctuate in time unless an infinite number of stochastic states is used. A finite number of random sampling vectors leads to an artificial dynamics that has to be disregarded when computing \( \delta \hat{\rho} \) and \( \delta \hat{n} \).

In practice, the density matrix induced by the addition/removal of an electron is computed as

\[
\delta \rho(r_1, r_2, t_1) = \rho_{\lambda}(r_1, r_2, t_1) - \rho_0(r_1, r_2, t_1)
\]

(36)

Here, \( \rho_{\lambda} \) represents the perturbed density matrix and \( \lambda \) denotes the strength of the perturbing potential due to charge addition. \( \rho_0 \) is the unperturbed density matrix which exhibits time dependence due to its stochastic nature.

The time-dependent density matrices are constructed from random vectors in the occupied subspace

\[
\rho_{\lambda}(r_1, r_2, t_1) = \{ \eta_{\lambda}(r_1, t) \eta_{\lambda}^\ast(r_2, t) \}
\]

(37)

The stochastic states \( \eta_{\lambda}(t) \) are found for each \( t \) by time evolution according to:

\[
\eta_{\lambda}(r_1, t_1) = \langle r_1 | e^{-i\hat{H}_0(t_1)t_1} | \eta_{\lambda} \rangle
\]

(38)

where \( \hat{H}_0 \) is the GKS Hamiltonian from Eq. 26 which adiabatically depends on \( t \) since \( \Sigma_H, V_c \) and \( \hat{V}_x \) are functionals of the time-dependent density and \( \hat{\Sigma}_x^\lambda \) is a functional of the density matrix. The time-dependent density is, of course, \( n(r_1, t_1) = \rho(r_1, r_1, t_1) \). Numerically, Eq. 38 is solved using Trotter propagation technique (see below).

The states \( \eta_{\lambda} \) are perturbed at \( t = 0 \): \n
\[
| \eta_{\lambda} \rangle = e^{i\hat{v}_{\lambda}} | \eta \rangle
\]

(39)

where \( \hat{v}_{\lambda} \) is a perturbing potential:

\[
v_{\lambda}(r_1) = \lambda \nu(r_1, r_2) \hat{\chi}(r_2) \phi_i(r_2)
\]

(40)

Here, \( \lambda \) is the strength of the perturbation. In practice, we take \( \lambda = 10^{-4} E_h^{-1} \), but the value of \( \lambda \) between \( 10^{-5} E_h^{-1} \) and \( 10^{-3} E_h^{-1} \) does not affect the results for molecules in Sec 4.

In practice, stochastic computation of the induced charge density requires only a few random states \( \eta \) (typically between 4 and 20), i.e. the number of states that are propagated by Eq. 38 is much smaller than the number of occupied states. Further, the induced density matrix is damped by a factor \( \exp[-(\alpha t)^2/2] \), where the damping factor is related to the maximum

---

1 Details of the implementation are given in Vlcek et al. Phys. Rev. B, 98(7):075107, 2018. Here, we use 20,000 stochastic vectors with size of 10% of the total real-space grid. This was found to be sufficient and agrees with our previous finding both molecules and periodic systems.
propagation time $\alpha = 3/t_{\text{max}}$.

3.2.2 Time propagation and stochastic decomposition of the long-range exchange

The self-energy requires two distinct time-propagations to be computed: for the GKS Green’s function (Eq. 32), and for the density matrix (Eq. 38). In both cases, the time-evolution operator is split into the local and non-local part of $\hat{H}_0$ and it is calculated in discrete time steps $\Delta t$ as:

$$e^{-i\hat{H}_0 \Delta t} = e^{-i\hat{\Sigma}^r \Delta t} e^{-i\hat{\Sigma}^\gamma \Delta t},$$  \hspace{1cm} (41)

where the local part of the Hamiltonian is

$$\hat{h}_L \equiv \hat{h} + \hat{\Sigma}_H + \hat{V}_c + \hat{V}_2.$$  \hspace{1cm} (42)

The time-evolution due to the non-local part is computed simply as:

$$e^{-i\hat{\Sigma}^\gamma \Delta t} \approx \left( \hat{T} - i\hat{\Sigma}^\gamma \frac{\Delta t}{2} \right).$$  \hspace{1cm} (43)

Here, the time-step is a parameter subject to convergence tests; in typical calculations, $\Delta t$ ranges between 0.02 and 0.05 a.u.

There are only a few vectors $\zeta$ and $\eta$. Nevertheless, the time evolution is costly due to the non-locality of $\Sigma^\gamma$. To make Eq. 43 less expensive, we use two additional sets of stochastic vectors to represent $\Sigma^\gamma_x$ (Eq. 27):

(I) The first set is used for the long-range Coulomb interaction $\nu^\gamma$:

$$\nu^\gamma(1, 2) = \{\chi^\gamma(r_1), \chi^\gamma(r_2)\} \delta(t_1, t_2).$$  \hspace{1cm} (44)

This form was applied previously to the ground state calculations \cite{53, 22} Here it is applied only to time-evolution of stochastic states.

(II) The second set, $\{\vartheta\}$, is used to decompose $G_0$. Since the exchange interaction is instantaneous, the GF is merely a density matrix. In practice, it is sufficient to use one or two the stochastic states $\vartheta$ (see the next section), which are obtained by a linear combination:

$$\vartheta(1) = \frac{1}{N_\eta} \sum_{j=1}^{N_\eta} e^{i\theta_j(t_1)} \eta_j(1),$$  \hspace{1cm} (45)

where $\theta \in [0, 2\pi]$ is a random phase and $N_\eta$ is the number of $\eta$ vectors used for decomposition of the density matrix (Eq. 30).

Together, the action of $\Sigma^\gamma_x$ on an arbitrary vector $\psi$ is

$$\langle r_1 | \hat{\Sigma}^\gamma_x | \psi \rangle = \{\vartheta(r_1) \chi^\gamma(r_1) \chi^\gamma(r_2) \vartheta(r_2) \psi(r_2)\}.$$  \hspace{1cm} (46)

Time arguments are omitted here since the exchange interaction is instantaneous (note the delta function in Eq. 44). Also, note that the $r_2$ coordinates are integrated out, i.e., $\chi^\gamma(r_2) \vartheta(r_2) \psi(r_2)$ is a complex number.

In practice, the numbers of $\vartheta$ and $\chi$ vectors are finite and hence the stochastic noise is, in principle, increased. However, at each time step, new random phases $\{\theta_j\}$ are selected. Frequent resampling of $\vartheta$ helps to reduce the stochastic error. As a result, only a few states are needed in actual calculations (see Sec. 4).

Finally, it is necessary to point out possible differences between the stochastic (real-time) and conventional $GW$ implementations. In one-shot $GW$ calculations, it is common to employ the random phase approximation (RPA), in which only the Hartree self-energy evolves in time (i.e., it is constructed at each time step). In the real-time implementation of RPA, the exchange self-energy does not evolve in time and $\vartheta$ is computed by repeated sampling of the static unperturbed vectors $\eta$ at $t = 0$. The corresponding results are labeled as $G_0W_0$.

It is relatively simple to go beyond-RPA in the real-time implementation. This amounts to the inclusion of the time-dependent xc terms and the screened Coulomb interaction is based on a test charge-test charge response function. \cite{23}

In this case, $\vartheta$ vectors that are constructed at each time step by a linear combination of time-propagated $\eta_\lambda$ states. To distinguish the level of theory applied, we label the beyond-RPA approaches as $G_0W_0^{tc}$ and $G_0W_0^{te}T_X$. All the methods are tested in the next section.
4 Results

4.1 Molecules, reference energies and the DFT starting point

Table 1: Real-space grids (characterized by the number of points $N_g$) and the range-separation parameters ($\gamma$) used in the calculations. Structures were taken from Refs. 38, 74, 75

| System       | $N_g$             | $\gamma$ [a$_0^{-1}$] |
|--------------|-------------------|------------------------|
| anthracene   | $80 \times 60 \times 50$ | 0.23                   |
| tetracene    | $88 \times 60 \times 50$ | 0.21                   |
| pentacene    | $108 \times 60 \times 50$ | 0.19                   |
| hexacene     | $112 \times 60 \times 50$ | 0.17                   |
| C$_{60}$     | $88 \times 88 \times 88$ | 0.18                   |
| PC$_{60}$BM  | $88 \times 88 \times 88$ | 0.15                   |

In this section, we use the stochastic approach described in Sec. 3 and calculate HOMO and LUMO QP energies of the molecules listed in Table 1. The set contains four acenes that were recently used to benchmark GW methods against high-level CCSD(T) results. In contrast to the original study, we omit the two smallest systems (benzene and naphtalene). We focus only on larger molecules that form stable anions since the goal is to test how different approximations treat both ionization potential and electron affinities. In addition, we selected two larger molecules ($C_{60}$ and PC$_{60}$BM), for which we use experimentally measured ionization potentials and affinities. Comparison to experiments should be taken as a rather qualitative indicator since our calculations do not consider vibrational effects and possible other contributions to the QP energies.

The starting point is based on calculations with the optimally-tuned LC-$\omega$PBE functional (see. Sec. 3). The ground state electronic structure was computed on a real-space grid (for each system specified in Table 1) with the grid-point spacing of 0.35 a$_0$ in combination with Trouiller-Martins pseudopotentials and kinetic energy cutoff of 28 $E_h$. With this set of parameters, the HOMO/LUMO eigenvalues are converged to 10 meV. These grids also yield converged QP energies, which are discussed in the next subsection. The values of the range-separation parameters are listed in Table 1, the DFT HOMO and LUMO eigenvalues are given in Table. 2.

4.2 QP energies

We now turn to the stochastic calculations of the QP energies using three approximations derived in Sec. 2, namely $G_0W_0$, $G_0W_0^tc$ and $G_0W_0^tcT_x$. The latter two methods do not employ RPA in the calculations of the screened Coulomb interaction. We first review convergence of the results with respect to the parameters of the computation and the number of stochastic samples; the results of the three methods are compared afterwards.

4.2.1 Convergence of stochastic errors

As discussed in the Sec. 3.2, the time propagations of $G_0$ and $\delta \rho$ are performed in discrete time steps $\Delta t$ for a limited total simulation time $t_{\text{max}}$. In all three approximations, the QP energies are converged to better than 0.02 eV with $\Delta t = 0.05$ a.u. and $t_{\text{max}} = 50$ a.u. $^2$ The same values were used in the previous calculations for $G_0W_0$ based on LDA starting point.

For a given set of time and real-space grid parameters, the QP energies exhibit stochastic fluctuations stemming from the random sampling vectors. In all three approximations, the following sets of stochastic orbitals are employed: (i) $\zeta$ for decomposition of the GKS Green’s function (Eq. 31); (ii) $\eta$ for decomposition of the induced density matrix (Eq. 37); (iii) $\chi$ for the decomposition of the exchange kernel $\nu^\gamma$ (Eq. 44); (iv) $\vartheta$ for decomposition of the density matrix in $\Sigma^\gamma_x$ (Eq. 45).

Three types of stochastic vectors are part of a “nested sampling”: There are $N_\eta \times N_\chi \times N_\vartheta$ states per each $\zeta$ vector. The overall error is thus governed mainly by the number of $N_\zeta$ samples, each having a stochastic fluctuation determined by $N_\eta \times N_\chi \times N_\vartheta$. In practice, $N_\zeta$ is

$^2$The values of $\Delta t$ and $t_{\text{max}}$ are parameters of the calculation similar to grid size and spacing, i.e., they are not directly related to the stochastic error.
increased until the statistical error is below a predetermined threshold.

For all systems investigated, we take $N_\eta = 15$, which is consistent with previous calculations for molecules. Nevertheless, additional tests for anthracene and C$_{60}$ show that the same QP energies are obtained with $N_\eta = 12$ (albeit with higher stochastic fluctuation per single $\zeta$). For the other two nested sampling vectors, it is sufficient to take $N_\theta = 1$ and $N_\chi = 20$. Such low values are due to a small magnitude of the $\Sigma_2^\chi$ term, which stems from a weak long-range exchange ($\gamma \leq 0.23 a_0^{-1},$ see Table.~[1]).

Fig.~[1] shows a graphical solution to the QP equation (Eq. 25) with different $N_\theta$ and $N_\chi$. The self-energy curves are almost identical; they differ by $\sim 0.02$ eV at the QP energy, which is about half of the statistical error due to to finite $N_\xi$ and $N_\eta$. Fig.~[1] further demonstrates that the self-energies are converged with $N_\chi$ and $N_\theta$ for frequencies several eV away from $\varepsilon$. Small values of $N_\theta$ and $N_\chi$ are sufficient in all three approximations tested here.

Finally, we compare the total stochastic error in the different approximations to $\Sigma_P$. In this analysis, the target fluctuation, $\sigma(\varepsilon)$, is 0.05 eV. The results are shown in Fig.~[2]. Note that the $G_0W_0$ calculations converge much slower compared to the other approximations, yet their total cost remains linear scaling (with a steep slope of $\sim 15$ core hours per electron). Further, $N_\xi$ is relatively high and rises with system size (for the two largest molecules). In contrast, the costs of $G_0W_0^{tc}$ and $G_0W_0^{tc}\Gamma_X$ show only a weak dependence on system size and the computational time remains practically identical.

The distinct behavior of the stochastic $G_0W_0$ calculations is due to RPA applied in the time propagation. As discussed in Sec.~[3.2.2] RPA assumes that $\Sigma_2^\chi$ is time-independent, i.e., it is not constructed from time propagated states. Although the $\Sigma_2^\chi$ term is sampled by distinct stochastic vectors $\vartheta$ at each time step, it leads to a strong stochastic noise. This random fluctuation is amplified with time (similar to the breakdown of stochastic BSE). Tests for hexacene showed that increase the value of $N_\theta$ to 2 leads to only a $\sim 1\%$ reduction of the fluctuation. Fig.~[3] clearly illustrates amplification of the stochastic error in $G_0W_0$ with time. In contrast, $\sigma(\varepsilon)$ in $G_0W_0^{tc}$ and $G_0W_0^{tc}\Gamma_X$ calculations remains almost constant regardless of $t_{\text{max}}$.

In general, the stochastic approach is well suited for large systems. The calculation for tetracene is the most expensive irrespective of the method chosen (Fig. 2). This behavior can be understood as follows: For small systems, $N_\eta = 15$ is relatively high compared to the number of occupied states (for instance, there are 33 occupied valence states in anthracene). As a result, the occupied subspace is sufficiently well sampled. For large systems, the stochastic approach exhibits strong self-averaging, which leads to a decrease of $N_\chi$ required for target $\sigma(\varepsilon)$, i.e., the computational cost is smaller. Tetracene is found to be the “worst-case” scenario in which the stochastic sampling introduces relatively large errors and there is only limited self-averaging.

Overall, the stochastic implementation of $G_0W_0^{tc}$ and $G_0W_0^{tc}\Gamma_X$ is efficient and numerically stable, while $G_0W_0$ suffers from large stochastic fluctuations. The total computational time of the beyond-RPA-methods is almost independent of the system size (Fig. 2). In the stochastic implementation, the more involved expressions for the self-energy are less expensive than their $G_0W_0$ counterpart. Low cost of stochastic beyond-GW calculations is in striking contrast to their conventional (deterministic) implementations.

4.2.2 Performance of the approximations to $\Sigma_P$

We will now address how distinct approximations to $\Sigma_P$ affect predictions of the HOMO and LUMO energies. Here, we will report results of stochastic calculations with $N_\xi = 1500$, $N_\theta = 1$, $N_\chi = 20$, and $N_\eta = 15$. The DFT starting point (optimally tuned LC-\omega PBE) is already in a good agreement with the reference values as shown in Table 2 yet, the HOMO/LUMO energies are consistently over/underestimated. The mean absolute error (MAE) of the DFT reference point is 0.16 eV for HOMO energies and 0.25 eV for LUMO energies.

The $G_0W_0$ approximation makes the HOMO
energies more negative (Table 2). As a result, the ionization potentials (negative of the HOMO energy) are significantly improved, and the corresponding MAE is only 0.07 eV. The $G_0W_0$ performance is, however, very different for LUMO. Here, the correction is too large and the QP energies are thus much higher than the reference values leading to MAE of 0.28 eV. If we exclude the experimental reference data for C$_{60}$ and PC$_{60}$BM, we get MAE of 0.04 eV and 0.39 eV for HOMO and LUMO, which are in agreement with an earlier benchmark for acenes.\textsuperscript{38}

This is a very disappointing result, because a more advanced computational technique ($G_0W_0$), which is aimed to improve upon DFT, yields worse results than the DFT itself. Further, the stochastic implementation of $G_0W_0$ on top of hybrid functionals is numerically expensive for large systems due to numerical instabilities discussed earlier. Note that such instabilities were not observed in previous stochastic $G_0W_0$ calculations based on LDA starting point.

The $G_0W_0^{tc}$ approximation is computationally stable and, similar to $G_0W_0$, it provides good ionization potentials; MAE for HOMO is 0.08 eV. However, the method amplifies the problems for unoccupied states. The affinities (negative of the LUMO energy) are predicted to be significantly larger than the reference values, leading to MAE of 0.44 eV which is worse than in $G_0W_0$. If experimental data for C$_{60}$ and PC$_{60}$BM are excluded from the analysis, $G_0W_0^{tc}$ leads to even larger errors for LUMO (0.55 eV). This failure for unoccupied states indicates that abandoning RPA has detrimental effects on unoccupied QP states.

Finally, we turn to the analysis of $G_0W_0^{tc}\Gamma_X$ predictions. The presence of the non-local exchange interaction in Eq. 23 has a significant impact on the QP energies. The HOMO states are shifted down by a smaller amount and lead to MAE of 0.09 eV. This error is only slightly worse than $G_0W_0$ but much better than the DFT results. In contrast to the other approximations tested here, $G_0W_0^{tc}\Gamma_X$ also improves QP energies of the unoccupied states, for which we find MAE of 0.19 eV. In this case, the perturbation theory improves upon the DFT starting point.

Based on the results for acenes, $G_0W_0^{tc}\Gamma_X$ ap-
Figure 2: The top panel shows the number of $\zeta$ states required to converge the statistical error of the HOMO and LUMO QP energies below 0.05 eV as a function of the total number of valence electrons $N_e$. The bottom graph shows the total number of core hours required for calculations of the QP energies. The calculations were performed on Bridges computer equipped with Intel Haswell (E5-2695 v3) CPUs.

Figure 3: Statistical error, $\sigma(\varepsilon)$, of the self-energy is shown as a function of the maximum propagation time, $t_{\text{max}}$, used for the calculation of the polarization self-energy. $\sigma(\varepsilon)$ is evaluated at the frequency corresponding to the QP energy $\varepsilon$. The data are for the LUMO QP state of the $C_{60}$ molecule. The number of $\zeta$ states is 1000 and the remaining parameters are described in the text.

pears to be more successful than self-consistent GW methods benchmarked in Ref. 38. If eigenvalue self-consistent GW is employed and both the GF and the screened Coulomb interactions are updated, the predictions yield MAE of 0.41 eV for the LUMO energies of the acene molecules. For the same subset of systems, $G_0W_0^{\text{tc}}\Gamma_X$ yields MAE of 0.25 eV. In an alternative scheme, the self-consistency is applied only to the GF, but for LUMO states the MAE increases to 0.44 eV, i.e., the deviation is still higher than for $G_0W_0^{\text{tc}}\Gamma_X$ results.

Together, these results indicate that better self-energy expressions, such as $G_0W_0^{\text{tc}}\Gamma_X$, are needed for quantitative predictions of QP energies, especially for unoccupied molecular states.

5 Conclusions and Outlook

In this work, we outlined an efficient way for improving predictions of QP energies beyond the popular GW approach using stochastic paradigm. In practice, this improvement amounts to the inclusion of nontrivial parts of the vertex function in the polarization self-energy ($\Sigma_P$). Here, an approximate non-local
While deterministic implementations scale as \( t^2 \) and \( N^6 \), where \( N \) is the number of electrons, the stochastic formulation scales (sub)linearly. In fact, we found that RPA is numerically unstable and its statistical error worsens with the system size and the simulation time. In contrast, more difficult \( G_0W_0^{tc} \) and \( G_0W_0^{tc}\Gamma_X \) calculations are stable and, paradoxically, computationally less expensive than \( G_0W_0 \).

The three approximations were tested on a set of acene molecules, \( C_{60} \) and \( PC_{60}BM \). The computational costs of \( G_0W_0^{tc} \) and \( G_0W_0^{tc}\Gamma_X \) were practically similar; the overall computational time required to converge QP energies was almost identical in all the systems irrespective of their size. This sublinear scaling is due to the rapid convergence of the statistical errors, i.e., the total number of stochastic samples of \( \Sigma_P \) decreases with the system size. Hence, stochastic algorithms will be a method of choice for demanding beyond-\( GW \) calculations or, at least, for efficient implementation of non-local vertex functions.

While DFT with optimally-tuned range-separated hybrid functionals (LC-\( \omega \)PBE) provides a good starting point, some deviation from reference data is observed. One-shot \( G_0W_0 \) and \( G_0W_0^{tc} \) improve the description of the ionization potentials compared to LC-\( \omega \)PBE, but severely increase errors for electron affinities. On average, both methods perform worse than DFT and the worst performance is observed for \( G_0W_0^{tc} \) approach. However, the

---

**Table 2: Quasiparticle energies computed with DFT (LC-\( \omega \)PBE) and distinct approximations to the self-energy.** Statistical errors of the stochastic methods are given in the parentheses. The reference values are taken from Refs. 38, 77–80. The number of \( \zeta \) states is 1,000; the remaining parameters are described in the text.

| system   | DFT  | \( G_0W_0 \) | \( G_0W_0^{tc} \) | \( G_0W_0^{tc}\Gamma_X \) | Ref. |
|----------|------|-------------|------------------|---------------------------|------|
| anthracene | 7.33 | 7.42 (±0.04) | 7.31 (±0.04) | 7.34 (±0.04) | 7.48 |
| tetracene | -6.70 | -7.00 (±0.06) | -6.89 (±0.05) | -6.98 (±0.05) | -6.96 |
| pentacene | -6.47 | -6.65 (±0.04) | -6.55 (±0.05) | -6.63 (±0.04) | -6.58 |
| hexacene  | -6.15 | -6.32 (±0.05) | -6.22 (±0.04) | -6.28 (±0.04) | -6.32 |
| \( C_{60} \) | -7.90 | -7.69 (±0.06) | -7.68 (±0.04) | -7.77 (±0.04) | -7.69* |
| PC\( _{60}BM \) | -7.27 | -7.42 (±0.06) | -7.26 (±0.03) | -7.38 (±0.03) | -7.17* |

| system   | DFT  | \( G_0W_0 \) | \( G_0W_0^{tc} \) | \( G_0W_0^{tc}\Gamma_X \) | Ref. |
|----------|------|-------------|------------------|---------------------------|------|
| anthracene | -0.49 | -0.54 (±0.04) | -0.71 (±0.04) | -0.36 (±0.04) | -0.28 |
| tetracene | -1.04 | -1.22 (±0.06) | -1.39 (±0.05) | -1.11 (±0.05) | -0.82 |
| pentacene | -1.53 | -1.65 (±0.05) | -1.80 (±0.05) | -1.50 (±0.04) | -1.21 |
| hexacene  | -1.84 | -1.94 (±0.05) | -2.07 (±0.04) | -1.84 (±0.04) | -1.47 |
| \( C_{60} \) | -2.47 | -2.77 (±0.07) | -2.91 (±0.04) | -2.74 (±0.04) | -2.68* |
| PC\( _{60}BM \) | -2.46 | -2.61 (±0.07) | -2.85 (±0.04) | -2.67 (±0.03) | -2.63* |
$G_0W_0^{te}\Gamma_X$ method performs best and outperforms DFT for both occupied and unoccupied states.

Previous calculations which included approximate vertex corrections were only applied to ionization potentials of small molecules that do not form stable anions. The accuracy of predicted electron affinities is the major difference among distinct approximations. Hence, comparison based solely on ionization potentials would in our case (incorrectly) suggests similar performance of all the methods. Electron affinities thus appear as a more sensitive indicator for performance assessment.

Together, these findings indicate that beyond GW schemes are crucial for improved description of QP energies. As shown here, stochastic techniques make such calculations affordable even for large systems. Future steps are directed toward the formulation of better self-energy expressions that include higher-order interactions (beyond $G_0W_0^{te}\Gamma_X$). Investigations of charge transfer systems with strong electron hole-interactions are underway.

Acknowledgement The calculations were performed as part of the XSEDE Project No. TG-CHE180051.

References

(1) Goedecker, S. Linear scaling electronic structure methods. Rev. Mod. Phys. 1999, 71, 1085.

(2) Goedecker, S.; Scuseria, G. Linear scaling electronic structure methods in chemistry and physics. Comput. Sci. Eng. 2003, 5, 14–21.

(3) Skylaris, C.-K.; Haynes, P. D.; Mostofi, A. A.; Payne, M. C. Introducing ONETEP: Linear-scaling density functional simulations on parallel computers. J. Chem. Phys. 2005, 122, 084119.

(4) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. Comput. Phys. Commun. 2005, 167, 103–128.

(5) Zhou, Y.; Saad, Y.; Tiago, M. L.; Chelikowsky, J. R. Self-consistent-field calculations using Chebyshev-filtered subspace iteration. J. Comput. Phys. 2006, 219, 172–184.

(6) Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. Efficient, approximate and parallel Hartree–Fock and hybrid DFT calculations. A ‘chain-of-spheres’ algorithm for the Hartree–Fock exchange. Chem. Phys. 2009, 356, 98–109.

(7) Saad, Y.; Chelikowsky, J. R.; Shontz, S. M. Numerical methods for electronic structure calculations of materials. SIAM Rev. 2010, 52, 3–54.

(8) Baer, R.; Neuhauser, D.; Rabani, E. Self-averaging stochastic Kohn–Sham density-functional theory. Phys. Rev. Lett. 2013, 111, 106402.

(9) Neuhauser, D.; Baer, R.; Rabani, E. Communication: Embedded fragment stochastic density functional theory. J. Chem. Phys. 2014, 141, 041102.

(10) Ostlund, N. S.; Szabo, A. Modern Quantum Chemistry: Introduction to advanced electronic structure theory; Dover Publications, 1996.

(11) Sherrill, C. D.; Schaefer III, H. F. Advances in quantum chemistry; Elsevier, 1999; Vol. 34; pp 143–269.

(12) Čížek, J. On the correlation problem in atomic and molecular systems. Calculation of wavefunction components in Ursell-type expansion using quantum-field theoretical methods. J. Chem. Phys. 1966, 45, 4256–4266.
(13) Paldus, J.; Čížek, J.; Shavitt, I. Correlation Problems in Atomic and Molecular Systems. IV. Extended Coupled-Pair Many-Electron Theory and Its Application to the BH 3 Molecule. Phys. Rev. A 1972, 5, 50.

(14) Bartlett, R. J.; Musial, M. Coupled-cluster theory in quantum chemistry. Rev. Mod. Phys. 2007, 79, 291.

(15) 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.

(16) Kaliman, I. A.; Krylov, A. I. New algorithm for tensor contractions on multicore CPUs, GPUs, and accelerators enables CCSD and EOM-CCSD calculations with over 1000 basis functions on a single compute node. J. Chem. Phys. 2017, 38, 842–853.

(17) Fetter, A. L.; Walecka, J. D. Quantum theory of many-particle systems; Dover Publications, 2003.

(18) Martin, R. M.; Reining, L.; Ceperley, D. M. Interacting electrons; Cambridge University Press, 2016.

(19) Hedin, L. New method for calculating the one-particle Green's function with application to the electron-gas problem. Phys. Rev. 1965, 139, A796.

(20) Aryasetiawan, F.; Gunnarsson, O. The GW method. Rep. Prog. Phys. 1998, 61, 237.

(21) Onida, G.; Reining, L.; Rubio, A. Electronic excitations: density-functional versus many-body Green’s-function approaches. Rev. Mod. Phys. 2002, 74, 601.

(22) 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.

(23) Hybertsen, M. S.; Louie, S. G. Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies. Phys. Rev. B 1986, 34, 5390.

(24) Faleev, S. V.; Van Schilfgaarde, M.; Kotani, T. All-Electron Self-Consistent GW Approximation: Application to Si, MnO, and NiO. Phys. Rev. Lett. 2004, 93, 126406.

(25) van Schilfgaarde, M.; Kotani, T.; Faleev, S. Quasiparticle Self-Consistent GW Theory. Phys. Rev. Lett. 2006, 96, 226402.

(26) Stan, A.; Dahlen, N. E.; Van Leeuwen, R. Levels of self-consistency in the GW approximation. J. Chem. Phys. 2009, 130, 114105.

(27) Bruneval, F.; Gatti, M. First Principles Approaches to Spectroscopic Properties of Complex Materials; Springer, 2014; pp 99–135.

(28) Deslippe, J.; Samsonidze, G.; Strubbe, D. A.; Jain, M.; Cohen, M. L.; Louie, S. G. BerkeleyGW: A massively parallel computer package for the calculation of the quasiparticle and optical properties of materials and nanostructures. Comput. Phys. Commun. 2012, 183, 1269–1289.

(29) Nguyen, H.-V.; Pham, T. A.; Rocca, D.; Galli, G. Improving accuracy and efficiency of calculations of photoemission spectra within the many-body perturbation theory. Phys. Rev. B 2012, 85, 081101.

(30) Pham, T. A.; Nguyen, H.-V.; Rocca, D.; Galli, G. GW calculations using the spectral decomposition of the dielectric matrix: Verification, validation, and comparison of methods. Phys. Rev. B 2013, 87, 155148.

(31) Liu, P.; Kaltak, M.; Klimeš, J.; Kresse, G. Cubic scaling GW: Towards fast quasi-
particle calculations. Phys. Rev. B 2016, 94, 165109.

(32) Holm, B.; von Barth, U. Fully self-consistent GW self-energy of the electron gas. Phys. Rev. B 1998, 57, 2108.

(33) Cao, H.; Yu, Z.; Lu, P.; Wang, L.-W. Fully converged plane-wave-based self-consistent GW calculations of periodic solids. Phys. Rev. B 2017, 95, 035139.

(34) Bruneval, F.; Marques, M. A. Benchmarking the starting points of the GW approximation for molecules. J. Chem. Theory Comput. 2012, 9, 324–329.

(35) Marom, N.; Caruso, F.; Ren, X.; Hofmann, O. T.; Körzdörfer, T.; Chelíkowsky, J. R.; Rubio, A.; Scheffler, M.; Rinke, P. Benchmark of GW methods for azabenzenes. Phys. Rev. B 2012, 86, 245127.

(36) Knight, J. W.; Wang, X.; Gallandi, L.; Dolgounitcheva, O.; Ren, X.; Ortiz, J. V.; Rinke, P.; Körzdörfer, 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) Caruso, F.; Dauth, M.; van Setten, M. J.; Rinke, P. Benchmark of GW Approaches for the GW 100 Test Set. J. Chem. Theory Comput. 2016, 12, 5076–5087.

(38) 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.

(39) Del Sole, R.; Reining, L.; Godby, R. GWT approximation for electron self-energies in semiconductors and insulators. Phys. Rev. B 1994, 49, 8024.

(40) Hung, L.; Felipe, H.; Souto-Casares, J.; Chelíkowsky, J. R.; Louie, S. G.; Öğüt, S. Excitation spectra of aromatic molecules within a real-space GW-BSE formalism: Role of self-consistency and vertex corrections. Phys. Rev. B 2016, 94, 085125.

(41) Romaniello, P.; Guyot, S.; Reining, L. The self-energy beyond GW: Local and non-local vertex corrections. J. Chem. Phys. 2009, 131, 154111.

(42) Morris, A. J.; Stankovski, M.; Delaney, K. T.; Rinke, P.; García-González, P.; Godby, R. W. Vertex corrections in localized and extended systems. Phys. Rev. B 2007, 76, 155106.

(43) Shishkin, M.; Marsman, M.; Kresse, G. Accurate quasiparticle spectra from self-consistent GW calculations with vertex corrections. Phys. Rev. Lett. 2007, 99, 246403.

(44) Ayral, T.; Werner, P.; Biermann, S. Spectral Properties of Correlated Materials: Local Vertex and Nonlocal Two-Particle Correlations from Combined GW and Dynamical Mean Field Theory. Phys. Rev. Lett. 2012, 109, 226401.

(45) Kuwahara, R.; Noguchi, Y.; Ohno, K. CWT+ Bethe-Salpeter equation approach for photoabsorption spectra: Importance of self-consistent GW calculations in small atomic systems. Phys. Rev. B 2016, 94, 121116.

(46) Maggio, E.; Kresse, G. GW Vertex Corrected Calculations for Molecular Systems. J. Chem. Theory Comput. 2017, 13, 4765–4778.

(47) Freeman, D. L. Coupled-cluster expansion applied to the electron gas: Inclusion of ring and exchange effects. Phys. Rev. B 1977, 15, 5512.

(48) Ren, X.; Marom, N.; Caruso, F.; Scheffler, M.; Rinke, P. Beyond the GW approximation: A second-order screened exchange correction. Phys. Rev. B 2015, 92, 081104.
(49) Grüneis, A.; Marsman, M.; Harl, J.; Schimka, L.; Kresse, G. Making the random phase approximation to electronic correlation accurate. *J. Chem. Phys.* 2009, *131*, 154115.

(50) Romaniello, P.; Bechstedt, F.; Reining, L. Beyond the GW approximation: Combining correlation channels. *Phys. Rev. B* 2012, *85*, 155131.

(51) Starke, R.; Kresse, G. Self-consistent Green function equations and the hierarchy of approximations for the four-point propagator. *Phys. Rev. B* 2012, *85*, 075119.

(52) Seidl, A.; Görling, A.; Vogl, P.; Majewski, J.; Levy, M. Generalized Kohn-Sham schemes and the band-gap problem. *Phys. Rev. B* 1996, *53*, 3764.

(53) Neuhauser, D.; Gao, Y.; Arntsen, C.; Karshenas, C.; Rabani, E.; Baer, R. Breaking the Theoretical Scaling Limit for Predicting Quasiparticle Energies: The Stochastic GW Approach. *Phys. Rev. Lett.* 2014, *113*, 076402.

(54) Gao, Y.; Neuhauser, D.; Baer, R.; Rabani, E. Sublinear scaling for time-dependent stochastic density functional theory. *J. Chem. Phys.* 2015, *142*, 034106.

(55) Neuhauser, D.; Rabani, E.; Cytter, Y.; Baer, R. Stochastic optimally tuned range-separated hybrid density functional theory. *J. Phys. Chem. A* 2015, *120*, 3071–3078.

(56) Rabani, E.; Baer, R.; Neuhauser, D. Time-dependent stochastic Bethe-Salpeter approach. *Phys. Rev. B* 2015, *91*, 235302.

(57) Vlcek, V.; Rabani, E.; Neuhauser, D.; Baer, R. Stochastic GW calculations for molecules. *J. Chem. Theory Comput.* 2017, *13*, 4997–5003.

(58) Vlcek, V.; Li, W.; Baer, R.; Rabani, E.; Neuhauser, D. Swift GW beyond 10,000 electrons using sparse stochastic compression. *Phys. Rev. B* 2018, *98*, 075107.

(59) Savin, A.; Flad, H.-J. Density functionals for the Yukawa electron-electron interaction. *Int. J. Quantum Chem.* 1995, *56*, 327–332.

(60) Leininger, T.; Stoll, H.; Werner, H.-J.; Savin, A. Combining long-range configuration interaction with short-range density functionals. *Chem. Phys. Lett.* 1997, *275*, 151–160.

(61) Baer, R.; Neuhauser, D. Density functional theory with correct long-range asymptotic behavior. *Phys. Rev. Lett.* 2005, *94*, 043002.

(62) Vydrov, O. A.; Scuseria, G. E. Assessment of a long-range corrected hybrid functional. *J. Chem. Phys.* 2006, *125*, 234109.

(63) Marques, M. A.; Oliveira, M. J.; Burdus, T. Libxc: A library of exchange and correlation functionals for density functional theory. *Comput. Phys. Commun.* 2012, *183*, 2272–2281.

(64) Lehtola, S.; Steigemann, C.; Oliveira, M. J.; Marques, M. A. Recent developments in libxc—a comprehensive library of functionals for density functional theory. *SoftwareX* 2018, *7*, 1–5.

(65) Stein, T.; Eisenberg, H.; Kronik, L.; Baer, R. Fundamental gaps in finite systems from eigenvalues of a generalized Kohn-Sham method. *Phys. Rev. Lett.* 2010, *105*, 266802.

(66) Stein, T.; Autschbach, J.; Govind, N.; Kronik, L.; Baer, R. Curvature and frontier orbital energies in density functional theory. *J. Phys. Chem. Lett.* 2012, *3*, 3740–3744.

(67) Kronik, L.; Stein, T.; Refaely-Abramson, S.; Baer, R. Excitation gaps of finite-sized systems from optimally tuned range-separated hybrid functionals. *J. Chem. Theory Comput.* 2012, *8*, 1515–1531.
(68) Yang, Z.-h.; Sottile, F.; Ullrich, C. A. Simple screened exact-exchange approach for excitonic properties in solids. Phys. Rev. B 2015, 92, 035202.

(69) Refaely-Abramson, S.; Jain, M.; Sharifzadeh, S.; Neaton, J. B.; Kronik, L. Solid-state optical absorption from optimally tuned time-dependent range-separated hybrid density functional theory. Phys. Rev. B 2015, 92, 035202.

(70) Brawand, N. P.; Vörös, M.; Govoni, M.; Galli, G. Generalization of dielectric-dependent hybrid functionals to finite systems. Phys. Rev. X 2016, 6, 041002.

(71) Baer, R.; Head-Gordon, M. Chebyshev expansion methods for electronic structure calculations on large molecular systems. J. Chem. Phys. 1997, 107, 10003–10013.

(72) Vlcek, V.; Baer, R.; Neuhauser, D. Stochastic Time-Dependent DFT with Optimally Tuned Range-Separated Hybrids: Application to Excitonic Effects in Large Phosphorene Sheets. arXiv preprint arXiv:1903.04604 2019.

(73) Bruneval, F.; Sottile, F.; Olevano, V.; Del Sole, R.; Reining, L. Many-Body Perturbation Theory Using the Density-Functional Concept: Beyond the GW Approximation. Phys. Rev. Lett. 2005, 94, 186402.

(74) Dorset, D. L.; McCourt, M. P. Disorder and the molecular packing of C60 buckminsterfullerene: a direct electron-crystallographic analysis. Acta Cryst. A 1994, 50, 344–351.

(75) Casalegno, M.; Zanardi, S.; Frigerio, F.; Po, R.; Carbonera, C.; Marra, G.; Nicolini, T.; Raos, G.; Meille, S. V. Solvent-free phenyl-C61-butyric acid methyl ester (PCBM) from clathrates: insights for organic photovoltaics from crystal structures and molecular dynamics. Chem. Commun. 2013, 49, 4525–4527.

(76) Troullier, N.; Martins, J. L. Efficient pseudopotentials for plane-wave calculations. Phys. Rev. B 1991, 43, 1993.

(77) Wang, X.-B.; Woo, H.-K.; Wang, L.-S. Vibrational cooling in a cold ion trap: Vibrationally resolved photoelectron spectroscopy of cold C60 anions. J. Chem. Phys. 2005, 123, 051106.

(78) Sun, H.; Ryno, S.; Zhong, C.; Ravva, M. K.; Sun, Z.; Körzdörfer, T.; Bredas, J.-L. Ionization energies, electron affinities, and polarization energies of organic molecular crystals: quantitative estimations from a polarizable continuum model (PCM)-tuned range-separated density functional approach. J. Chem. Theory Comput. 2016, 12, 2906–2916.

(79) Akaike, K.; Kanai, K.; Yoshida, H.; Tsutsurni, J.; Nishi, T.; Sato, N.; Ouchi, Y.; Seki, K. Ultraviolet photoelectron spectroscopy and inverse photoemission spectroscopy of [6, 6]-phenyl-C 61-butyric acid methyl ester in gas and solid phases. J. Appl. Phys. 2008, 104, 023710.

(80) Larson, B. W.; Whitaker, J. B.; Wang, X.-B.; Popov, A. A.; Rumbles, G.; Kopidakis, N.; Strauss, S. H.; Boltalina, O. V. Electron affinity of phenyl-C61–butyric acid methyl ester (PCBM). J. Phys. Chem. C 2013, 117, 14958–14964.

(81) Towns, J.; Cockerill, T.; Dahan, M.; Foster, I.; Gaither, K.; Grimshaw, A.; Hazlewood, V.; Lathrop, S.; Lifka, D.; Peterson, G. D. XSEDE: Accelerating scientific discovery,. Comput. Sci. Eng. 2014, 16, 62.