Quantum-quantum and quantum-quantum-classical schemes with projection-based-embedded GW-Bethe–Salpeter Equation

Vivek Sundaram[†] and Björn Baumeier[*,†]

†*Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600MB Eindhoven, The Netherlands*

‡*Institute for Complex Molecular Systems, Eindhoven University of Technology, P.O. Box 513, 5600MB Eindhoven, The Netherlands*

¶*Department of Applied Physics and Science Education, Eindhoven University of Technology, P.O. Box 513, 5600MB Eindhoven, The Netherlands*

E-mail: b.baumeier@tue.nl

Abstract

We present quantum-quantum and quantum-quantum-classical schemes based on many-body Green's functions theory in the GW approximation with the Bethe–Salpeter equation (GW-BSE) employing projection-based-embedding (PbE). Such approaches allow defining active and inactive subsystems of larger, complex molecular systems, with only the smaller active subsystem being explicitly treated by GW-BSE offering significant computational advantages. However, as PbE can modify the single-particle states in the ground state calculation and screening effects from inactive region are not automatically included in GW-BSE, results from such PbE-GW-BSE calculations can

1

deviate from a full-system reference. Here, we scrutinize in detail, e.g., the individual and combined effects of different choices of active regions, the influence of screening from the inactive region, and strategies for basis set truncation on frontier orbital and near-gap electron-hole excitation energies. As prototypical systems, we consider a diketopyrrolopyrrole bicyclic ring including side-chains, a polarity-sensitive dye (prodan) in aqueous environment, and a π-stacked dimer of benzene and tetracyanoethylene in water, respectively, covering a variety of excitation characters in molecular systems with complex chemical environments and photoinduced processes.

1 Introduction

Many-body Green's functions Theory employing the GW approximation and the Bethe–Salpeter equation (BSE)[1] has been a widely established method for the determination of electronic excitations in solid-state physics. Over the last decade, it has gradually found more and more application in traditionally molecular quantum chemistry settings[2–6]. It was shown that GW-BSE provides an effective single- and two-particle picture with accurate energies for charged and neutral excitations of different character, e.g., photoionization and localized vs charge-transfer type excitations, without the need for any adaptations[4,7,8]. Even though its scaling (dependent on details of the implementation) is favorable compared to wave-function based methods such as ADC(2)[9] or CC2[10], the direct application of GW-BSE to many complex molecular systems remains computationally challenging. Examples for such molecular systems are polymers with complex internal architecture, either solvated[11] or pure or mixed blends[12], more general solvent-solute systems with nonequilibrium relaxation dynamics[13], or molecular aggregates as in organic semiconductor films[14].

To make systems like these accessible, hybrid methods combining quantum and classical methods are often used[4,13,15–19], sometimes combined with machine-learning models[20]. While effective, such approaches rely on, e.g., some intuitive partitioning of the supramolecular system into fragments that only interact via classical electrostatics and a careful choice
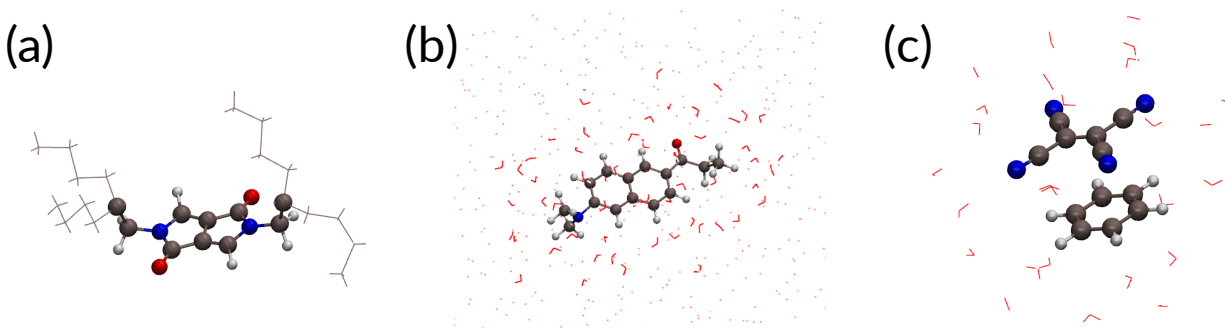
Figure 1: Molecular structures used as test systems for PbE-*GW*-BSE: (a) DPP bicyclic ring with branched alkyl side chains, (b) prodan in close (quantum) and distant (classical) water, (c) a water-solvated benzene-TCNE dimer.

and parametrization of the environment model, and may fail for covalently or hydrogen-bonded systems with partial charge transfer. In these cases, a quantum-quantum embedding approach might be advantageous, which allows defining active and inactive partitions, with only the smaller active partition being explicitly treated by *GW*-BSE, while interaction with the inactive one is on the level of density-functional theory (DFT)[21–23]. Recently, Tölle *et al.*[24,25] have reported *GW*-BSE calculations based on subsystem-DFT (sDFT)[26–30] for a series of weakly interacting molecular clusters. sDFT starts from Kohn–Sham-like calculations on fragments and determines the full-system density from them, in which the effective potential for a fragment contains contributions from non-additive terms in the kinetic energy and exchange-correlation potential. This intrinsic partitioning of the supramolecular system into small fragments allowed the authors to also partition the screening contributions to the correlation part of the self-energy in *GW* and to approximately include environment polarization effects into the calculations, and it could be shown to recover to a large extent the full-system frontier orbital energies and the transitions between them.

Projection-based-embedding[31–36] (PbE) is an alternative to sDFT, which partitions the full system based on a full-system reference Kohn–Sham calculation into active and inactive parts, and to subsequently restrict the *GW*-BSE calculation to the active part. PbE has some advantages over sDFT, e.g., it is a formally exact partitioning and also works for active/inactive partitioning through covalent bonds, which is important for studies of

3

macromolecular assemblies with electronically active vs inactive functional groups. However, for large inactive parts involving many molecules or molecular fragments adopting the environment screening correction proposed in[24] becomes cumbersome, as it would require either a screening calculation for a still intractably large single inactive region or further decomposition of the latter.

Arguably there are scenarios in which the exact agreement between a (hypothetical) full supramolecular and a PbE-GW-BSE calculation is not required. For dynamical processes such as charge or exciton transfer or conversions between localized and charge-transfer type excitations, for instance, only relative energy gaps are relevant. Against this background, we scrutinize in this work in detail different schemes of PbE-GW-BSE calculations: plain calculations, calculations in which a truncated atomic orbital basis is used, and calculations in which the PbE-GW-BSE calculation is further embedded in a classical, atomistic polarizable region. We aim to elucidate the individual and combined effects of different choices of active regions, the influence of screening (or lack thereof) from the inactive region, and impact of basis set truncation on energies of local and charge-transfer excitations, respectively. As prototypical systems covering different types of excitations in a variety of chemical environments, we consider the three test systems shown in Figure 1: (a) a single diketopyrrolopyrrole (DPP) bicyclic ring with branched alkyl side-chains, (b) prodan, a polarity-sensitive dye, solvated in water, and (c) a benenze-TCNE donor-acceptor pair in water. For all systems, we study the effects of the PbE with (trPbE) or without basis truncation on the frontier orbitals as well as the selected local or CT excitations. We pay special attention to the differences in contributions of the exchange and correlation parts of the self-energy to quasiparticle energies and of exchange and direct terms in the electron-hole interaction to the BSE energies between full and embedded calculations. In addition to this, each of the three test systems is here chosen with specific objectives: For the DPP molecule with branched side chains, we intend to demonstrate a PbE-GW-BSE calculation for a system in which the two regions are connected by a covalent bond. We investigate

4

the sensitivity of the embedding results on choice the choice of the active region. Prodan in water has been chosen to showcase the quantum-quantum-classical PbE-GW-BSE/MM approach. Also here, we consider the influence of the choice of the active regions on the predicted excitation energies. The donor-acceptor benzene-TCNE dimer solvated in water is used to evaluate the differences in embedding effects on localized and charge-transfer type excitations.

This paper is organized as follows: Section 2 summarizes the essentials of many-body Green's functions theory, projection-based-embedding and basisset truncation methods, as well as the coupling to classical polarizable environments. Computational details are given in Section 3 before the results for the three test systems are presented in Section 4 and the overall findings discussed in Section 5. A brief summary concludes the paper.

2 Methodology

2.1 Many-Body Green's Functions Methods for Electronically Excited States

Kohn–Sham (KS) DFT[37,38] provides the starting point for the effective single- or two-particle formulations for electronic excitations and their energies within the framework of perturbation theory with many-body Green's functions. One first obtains KS wavefunctions $\phi_i^{\mathrm{KS}}(\mathbf{r})$ and energies $\varepsilon_i^{\mathrm{KS}}$ from

$$\left\{ -\frac{1}{2}\Delta + v_{\mathrm{ext}}(\mathbf{r}) + v_{\mathrm{H}}(\mathbf{r}) + v_{\mathrm{xc}}[n](\mathbf{r}) \right\} \phi_i^{\mathrm{KS}}(\mathbf{r}) = \varepsilon_i^{\mathrm{KS}} \phi_i^{\mathrm{KS}}(\mathbf{r}), \tag{1}$$

where v_{ext} is the external potential, v_{H} the Hartree potential, and v_{xc} the exchange-correlation potential, and they define together with the kinetic energy operator the effective KS Hamiltonian $\hat{H}_{\mathrm{KS}}$.

Hedin[39,40] introduced the GW approximation of many-body Green's functions theory, in

which electron self-energy is written as $\Sigma = iGW$, and allows to derive a set of effective single-particle eigenvalue problems known as the *quasiparticle* (QP) equations

$$
\left[\hat{H}^{\text{KS}} - v_{\text{xc}}(\mathbf{r})\right] \phi_i^{\text{QP}}(\mathbf{r}) + \int \Sigma(\mathbf{r}, \mathbf{r}', \varepsilon_i^{\text{QP}}) \phi_i^{\text{QP}}(\mathbf{r}') \, d^3 r'
$$
$$
= \varepsilon_i^{\text{QP}} \phi_i^{\text{QP}}(\mathbf{r}'). \tag{2}
$$

Typically, the QP wave functions $\phi_i^{\text{QP}}(\mathbf{r})$ are approximated by the KS wave functions, which allows to write the QP energies as

$$
\varepsilon_i^{\text{QP}} = \varepsilon_i^{\text{KS}} + \left\langle \phi_i^{\text{KS}} \left| \Sigma(\varepsilon_i^{\text{QP}}) \right| \phi_i^{\text{KS}} \right\rangle. \tag{3}
$$

The self-energy is calculated in frequency space as

$$
\Sigma(\mathbf{r}, \mathbf{r}', \omega) = \frac{\text{i}}{2\pi} \int G(\mathbf{r}, \mathbf{r}', \omega + \omega') W(\mathbf{r}, \mathbf{r}', \omega') e^{\text{i}\omega'\eta} \, d\omega', \tag{4}
$$

from the Green's function based on the Kohn–Sham solution

$$
G(\mathbf{r}, \mathbf{r}', \omega) = \sum_m \frac{\phi_m^{\text{KS}}(\mathbf{r}) \phi_m^{\text{KS}*}(\mathbf{r}')}{\omega - \varepsilon_m^{\text{KS}} - \text{i}\eta \,\text{sgn}(E_{\text{F}} - \varepsilon_m^{\text{KS}})} \tag{5}
$$

and the screened Coulomb interaction W in the random-phase approximation

$$
W(\mathbf{r}, \mathbf{r}', \omega) = \int \epsilon^{-1}(\mathbf{r}, \mathbf{r}'', \omega) v_{\text{C}}(\mathbf{r}'', \mathbf{r}') \, d^3 r''. \tag{6}
$$

Evaluating Eq. (6) in turn requires the microscopic, frequency-dependent dielectric function given by

$$
\epsilon(\mathbf{r}, \mathbf{r}', \omega) = \delta(\mathbf{r}, \mathbf{r}') - \int v_{\text{C}}(\mathbf{r}, \mathbf{r}'') \chi_0(\mathbf{r}'', \mathbf{r}', \omega) \, d\omega' \tag{7}
$$

6

containing the irreducible polarizability χ_0:

$$
\begin{aligned}
\chi_0(\mathbf{r}, \mathbf{r}', \omega) = \sum_v^{\text{occ}} \sum_c^{\text{unocc}} \Bigg\{ & \frac{\phi_v^{\text{KS*}}(\mathbf{r})\phi_c^{\text{KS}}(\mathbf{r})\phi_c^{\text{KS*}}(\mathbf{r}')\phi_v^{\text{KS}}(\mathbf{r}')}{\omega - (\varepsilon_c^{\text{KS}} - \varepsilon_v^{\text{KS}}) + i\eta} \\
& - \frac{\phi_v^{\text{KS}}(\mathbf{r})\phi_c^{\text{KS*}}(\mathbf{r})\phi_c^{\text{KS}}(\mathbf{r}')\phi_v^{\text{KS*}}(\mathbf{r}')}{\omega + (\varepsilon_c^{\text{KS}} - \varepsilon_v^{\text{KS}}) - i\eta} \Bigg\} .
\end{aligned}
\tag{8}
$$

As the self-energy is energy-dependent, and thus depends on $\varepsilon_i^{\text{QP}}$, the solution of Eq. (3) must be found self-consistently. From Eq. (7) it is possible to split the self-energy $\Sigma = iGW$ into its bare exchange part

$$
\begin{aligned}
\Sigma^{\text{x}}(\mathbf{r}, \mathbf{r}') &= \frac{i}{2\pi} \int G(\mathbf{r}, \mathbf{r}'\omega + \omega')v_{\text{C}}(\mathbf{r}, \mathbf{r}')e^{i\omega'\eta}\, d\omega' \\
&= -\sum_v^{\text{occ}} \phi_v^{\text{KS}}(\mathbf{r})\phi_v^{\text{KS*}}(\mathbf{r}')v_{\text{C}}(\mathbf{r}, \mathbf{r}')
\end{aligned}
\tag{9}
$$

and the explicitly frequency-dependent correlation part

$$
\Sigma^{\text{c}}(\mathbf{r}, \mathbf{r}', \omega) = \frac{i}{2\pi} \int G(\mathbf{r}, \mathbf{r}', \omega + \omega')\left(W(\mathbf{r}, \mathbf{r}', \omega') - v_{\text{C}}(\mathbf{r}, \mathbf{r}')\right)\, d\omega'.
\tag{10}
$$

With $\omega_i = \varepsilon_i^{\text{KS}} + \langle \phi_i^{\text{KS}}| \Sigma^{\text{x}} - v^{\text{KS}} |\phi_i^{\text{MD}} \rangle$ and $\langle \phi_i^{\text{MF}}| \Sigma^{\text{c}}(\omega) |\phi_i^{\text{KS}} \rangle = \Sigma_i^{\text{c}}(\omega)$, we can rewrite Eq. (3) into the fixed-point problem

$$
\omega - \omega_i = \Sigma_i^{\text{c}}(\omega).
\tag{11}
$$

Due to the pole structure of the self-energy, there are in general several solutions to Eq. (11). In this situation, the spectral weight, defined as

$$
Z(\omega) = \left(1 - \frac{d\Sigma^{\text{c}}(\omega)}{d\omega}\right)^{-1},
\tag{12}
$$

is used to identify the "true" QP energy by $Z(\omega) \approx 1$, or $|d\Sigma^{\text{c}}(\omega)/d\omega| \approx 0$.

In the above, when evaluating the self-energy, the KS eigenvalues and eigenfunctions

7

are used to construct G and W, which is also known as a "one-shot" G_0W_0 calculation. Alternatively, it is possible to use updated QP energies until eigenvalue self-consistency is reached (evGW)[41–43].

Charge-neutral excitations that involve excitonic effects (electron-hole pair interaction) are not accounted for, and can instead to obtained so solutions to the *Bethe–Salpeter equation* (BSE)

$$\underline{\mathbf{H}}^{\text{BSE}} \ket{\zeta_S} = \Omega_S \ket{\zeta_S}, \tag{13}$$

in which the electron-hole wave functions $\ket{\zeta_S}$ are typically expressed in a basis of resonant and antiresonant products of single-particle functions

$$\zeta_S(\mathbf{r}, \mathbf{r}') = A_{vc}^S \phi_c(\mathbf{r}) \phi_v^*(\mathbf{r}') + B_{vc}^S \phi_v(\mathbf{r}) \phi_c^*(\mathbf{r}'). \tag{14}$$

With that the BSE explicitly reads in matrix form

$$\begin{pmatrix} \underline{\mathbf{H}}^{\text{res}} & \underline{\mathbf{K}} \\ -\underline{\mathbf{K}} & -\underline{\mathbf{H}}^{\text{res}} \end{pmatrix} \begin{pmatrix} \mathbf{A}^S \\ \mathbf{B}^S \end{pmatrix} = \Omega_S \begin{pmatrix} \mathbf{A}^S \\ \mathbf{B}^S \end{pmatrix}, \tag{15}$$

with elements

$$H_{vc,v'c'}^{\text{res}}(\omega) = D_{vc,v'c'} + K_{vc,v'c'}^{\text{x}} + K_{vc,v'c'}^{\text{d}} \tag{16}$$

$$K_{cv,v'c'}(\omega) = K_{cv,v'c'}^{\text{x}} + K_{cv,v'c'}^{\text{d}}. \tag{17}$$

8

and

$$D_{vc,v'c'} = (\varepsilon_c - \varepsilon_v)\delta_{vv'}\delta_{cc'} \tag{18}$$

$$K^{\mathrm{x}}_{vc,v'c'} = \int \phi^*_c(\mathbf{r})\phi_v(\mathbf{r})v_{\mathrm{C}}(\mathbf{r},\mathbf{r}')\phi_{c'}(\mathbf{r}')\phi^*_{v'}(\mathbf{r}')\mathrm{d}^3\mathbf{r}\,\mathrm{d}^3\mathbf{r}' \tag{19}$$

$$K^{\mathrm{d}}_{vc,v'c'} = \int \phi^*_c(\mathbf{r})\phi_{c'}(\mathbf{r})\phi_v(\mathbf{r}')\phi^*_{v'}(\mathbf{r}') \tag{20}$$

$$\times W(\mathbf{r},\mathbf{r}',\omega = 0)\mathrm{d}^3\mathbf{r}\,\mathrm{d}^3\mathbf{r}'\,. \tag{21}$$

If the ground state is a spin-singlet state and spin-orbit coupling is small compared to the electron-hole coupling, the BSE solutions can be classified as spin-singlet and spin-triplet excitations. This allows in turn, to solve the BSE separately for the spin type of interest, with

$$H^{\mathrm{BSE}}_{\mathrm{singlet}} = D + K^{\mathrm{d}} + 2K^{\mathrm{x}} \tag{22}$$

$$H^{\mathrm{BSE}}_{\mathrm{triplet}} = D + K^{\mathrm{d}}, \tag{23}$$

respectively.

2.2 Projection-based-embedding

Projection-based embedding[31] (PbE) is a formally exact DFT-in-DFT embedding scheme. It begins with a standard DFT calculation on the complete reference system in the full-molecule AO basis. The N occupied canonical molecular orbitals from this calculation, $\phi_i(\mathbf{r})$ for $i = 1,\dots N$, are then first transformed into localized orbitals, $\phi^{\mathrm{LO}}_i(\mathbf{r})$, with a unitary transformation that leaves the total density of the system unchanged. With the specification of a set of atoms in the active region A, one constructs an initial active density $n^{\mathrm{A}}(\mathbf{r})$ from those localized orbitals that have a significant Mulliken population $q^{\mathrm{A}}_i > q_t$ on these atoms

9

(q_t is typically 0.4):

$$n^{\mathrm{A}}(\mathbf{r}) = 2 \sum_{\substack{i=1 \\ q_i^{\mathrm{A}} > q_t}}^{N} \left| \phi_i^{\mathrm{LO}}(\mathbf{r}) \right|^2. \tag{24}$$

With $n^{\mathrm{A}}(\mathbf{r})$ given, one can determine the density of the inactive region as $n^{\mathrm{B}}(\mathbf{r}) = n(\mathbf{r}) - n^{\mathrm{A}}(\mathbf{r})$. From this initial partitioning of the total reference density, one considers $n^{\mathrm{A}}(\mathbf{r})$ variable, denoted here as $\tilde{n}^{\mathrm{A}}(\mathbf{r})$. The Fock matrix in the full-molecule AO basis for an embedded (A-in-B) calculation on the electrons in subsystem A is given by

$$\underline{\mathbf{F}}^{\mathrm{A\text{-}in\text{-}B}} = \underline{\mathbf{h}}_{\mathrm{core}}^{\mathrm{A\text{-}in\text{-}B}}[n^{\mathrm{A}} + n^{\mathrm{B}}] + \underline{\mathbf{J}}[\tilde{n}^{\mathrm{A}}] + \underline{\mathbf{K}}[\tilde{n}^{\mathrm{A}}] + \underline{\mathbf{V}}_{\mathrm{xc}}[\tilde{n}^{\mathrm{A}}], \tag{25}$$

where $\underline{\mathbf{h}}_{\mathrm{core}}^{\mathrm{A\text{-}in\text{-}B}}$ is the embedded core Hamiltonian based on the partitioned initial densities n^{A} and n^{B}. The density-dependent terms $\underline{\mathbf{J}}$, $\underline{\mathbf{K}}$, and $\underline{\mathbf{V}}_{\mathrm{xc}}$ are updated in each iterative step and hence depend on the updated active density $\tilde{n}^{\mathrm{A}}$. The embedded core Hamiltonian reads

$$\begin{aligned}
\underline{\mathbf{h}}_{\mathrm{core}}^{\mathrm{A\text{-}in\text{-}B}}[n^{\mathrm{A}} + n^{\mathrm{B}}] = {} & \underline{\mathbf{h}}_{\mathrm{core}} + \underline{\mathbf{J}}[n^{\mathrm{A}} + n^{\mathrm{B}}] - \underline{\mathbf{J}}[n^{\mathrm{A}}] \\
& + \underline{\mathbf{K}}[n^{\mathrm{A}} + n^{\mathrm{B}}] - \underline{\mathbf{K}}[n^{\mathrm{A}}] \\
& + \underline{\mathbf{V}}_{\mathrm{xc}}[n^{\mathrm{A}} + n^{\mathrm{B}}] - \underline{\mathbf{V}}_{\mathrm{xc}}[n^{\mathrm{A}}] + \mu \mathbf{P}_{\mathrm{B}}
\end{aligned} \tag{26}$$

and contains the core Hamiltonian of the full system, the difference between the Hartree, exchange, and exchange-correlation terms for the full system and the initially chosen active subsystem, respectively, as well as a projection term $\mu \underline{\mathbf{P}}_{\mathrm{B}}$ with the projection operator

$$\underline{\mathbf{P}}_{\mathrm{B}} = \underline{\mathbf{S}} \underline{\mathbf{D}}^{\mathrm{B}} \underline{\mathbf{S}} \tag{27}$$

based on the density matrix $\underline{\mathbf{D}}^{\mathrm{B}}$ of the environment and the atomic orbital overlap matrix $\underline{\mathbf{S}}$ to ensure orthogonality between the occupied states of the environment and the rest of the active subsystem. The level-shift parameter μ is typically of the order of $10^5 - 10^6$ Hartree. In the limit $\mu \to \infty$, the two subsystems are exactly orthogonal.

10

Performing a PbE calculation allows us to limit a *GW*-BSE calculation for the electronic excitations on the active region. As the embedding potential is already included in the preceding PbE-DFT calculation, there are no changes to the procedure of the *GW*-BSE steps as outlined in Section 2.1, except that all quantities involved use the embedded Kohn–Sham molecular orbitals and their energies as starting point. We will discuss the impact on the results of such calculations in Section 4.

2.3 Truncation of the full atomic orbital basis

All matrices composing the Fock matrix $\underline{\mathbf{F}}^{\text{A-in-B}}$ are so far expressed in the full molecular atomic orbital basis. The most noteworthy computational gain of the embedding lies in the fact that only a smaller number of occupied states need to be explicitly determined in the self-consistent procedure.

Reducing the atomic orbital basis for the actual embedding step should not only offer computational savings by decreasing the dimension of the eigensystem, but also affect the virtual orbital space and, in an extreme case, localize the orbitals intrinsically in the active region.

Miller *et al.* have shown that a reduced atomic orbital basis can be constructed by truncating the full basis of the reference calculation via manipulation of the projection term[35,36]. The procedure begins with an additional classification of atoms in the inactive region into *border atoms* and *distant atoms*. This distinction is based on whether any of the atomic basis functions centered at an atom of the inactive region contribute beyond a threshold to the density of the active region. In practice, this is determined based on the net Mulliken population of an atomic orbital α in the active density matrix $\underline{\mathbf{D}}^{\text{A}}$

$$q_\alpha = D^{\text{A}}_{\alpha\alpha} S_{\alpha\alpha}, \tag{28}$$

where $\underline{\mathbf{S}}$ is the overlap matrix. If any of the q_α exceeds a threshold value (typically 10^{-4}),

11

the atom associated with basis function α is added to the list of border atoms. All remaining atoms are distant atoms. Subsequently, the originally assigned inactive molecular orbitals are also split into border and distant MOs. Border MOs are inactive molecular orbitals that have a Mulliken population larger than a threshold (again, typically 0.4) on any of the border atoms. All remaining MOs are distant molecular orbitals.

This splitting into border and distant molecular orbitals also allows a similar splitting of the projection operator via the respective density matrices $\underline{\mathbf{D}}^{\text{border}}$ and $\underline{\mathbf{D}}^{\text{distant}}$

$$\mathbf{P}_{\text{B}} = \mathbf{P}_{\text{B}}^{\text{border}} + \mathbf{P}_{\text{B}}^{\text{distant}} = \underline{\mathbf{S}}\underline{\mathbf{D}}^{\text{border}}\underline{\mathbf{S}} + \underline{\mathbf{S}}\underline{\mathbf{D}}^{\text{distant}}\underline{\mathbf{S}}. \tag{29}$$

The split projector is now used in Eq. (26), however, with different values for the level-shift, $\mu^{\text{border}} \sim 100\,\text{Ha}$ and $\mu^{\text{border}} \sim 10^5\,\text{Ha}$. These steps allow evaluating the Fock matrix in Eq. (25) in a reduced, truncated, basis which only includes the basis functions centered at the active and border atoms. We refer to this as truncated projection-based-embedding (trPbE), in the remainder.

2.4 Classical Polarizable Embedding

To account for the effects of a complex molecular environment on electronic excitations, a quantum (QM) region with the excited state complex is embedded in a classical, polarizable atomistic (MM) model for the environment. The QM/MM scheme in VOTCA-XTP makes use of a distributed atomic multipole representation for molecules in the MM region, which allows treatment of both the effects of static electric fields and the polarization response as a self-consistent reaction field. Specifically, this classical MM energy for the system is evaluated as

$$E_{\text{MM}} = \frac{1}{2}\sum_{\substack{I,J \\ I \neq J}}\sum_{i \in I}\sum_{j \in J}\sum_{tu}(Q_t^i + \Delta Q_t^i)T_{tu}^{ij}Q_u^j, \tag{30}$$

12

where I and J indicate individual molecules in the system, i and j atoms in the respective molecules, Q_t^i are the static atomic multipole moments of rank t associated to atom i, and T_{tu}^{ij} is the tensor describing the interactions between the multipoles moments Q_t^i and $Q_u^{j\,44}$. The induced moments ΔQ_t^i are generated by the electric field created by moments t' of atom $i' \neq i$ in molecule I and the one generated by the moment u of atom j in molecule J:

$$\Delta Q_t^i = -\sum_{\substack{I,J\in\mathcal{S}\\I\neq J}}\sum_{j\in J}\sum_{\substack{i'\in I\\i'\neq i}}\sum_{tt'u}\alpha_{tt'}^{ii'}T_{t'u}^{i'bj}(Q_u^j + \Delta Q_u^j), \tag{31}$$

with $\alpha_{tt'}^{ii'}$ the atomic polarizability on each site. To avoid the effects of spurious overpolarization, a damped version of the interaction tensor (Thole damping[44]) is used. Then, the static and induced multipoles in the MM region also interact with the electron density in the QM region via an additional external potential to Eq. (1). At the same time, the explicit electrostatic field from the QM density is included in polarizing the MM region. The total density of excited state S is evaluated from the excited-state wavefunction ζ^S as

$$\rho^S(\mathbf{r}) = \rho_{\mathrm{DFT}}(\mathbf{r}) + \rho_e^S(\mathbf{r}) - \rho_h^S(\mathbf{r}), \tag{32}$$

with

$$\begin{aligned}
\rho_e^S(\mathbf{r}) &= \rho_e^S(\mathbf{r}_e) = \int d\mathbf{r}_h |\zeta_S(\mathbf{r}_e, \mathbf{r}_h)|^2 \\
\rho_h^S(\mathbf{r}) &= \rho_h^S(\mathbf{r}_h) = \int d\mathbf{r}_e |\zeta_S(\mathbf{r}_e, \mathbf{r}_h)|^2.
\end{aligned} \tag{33}$$

To obtain the polarization response of both the QM and MM regions, a self-consistent procedure is employed. At step p of this procedure, the total energy of the coupled QM/MM system for the state S of interest (ground state $S = 0$, or excited states $S > 0$) is determined as

$$E_{\mathrm{QM/MM}}^{S,p} = E_{\mathrm{QM}}^{S,p} + E_{\mathrm{MM}}^{S,p}, \tag{34}$$

13

with

$$E_{\text{QM}}^{S,p} = E_{\text{DFT}}^{S,p} + \Omega_S^p, \tag{35}$$

and $\Omega_S^p = 0$ for the ground state case. The whole procedure is repeated until the change of total energy is less than a preselected accuracy, typically 10^{-5} Ha. The excitation energy $\Omega_S^{\text{QM/MM}}$ of a complex in the polarizable environment is then obtained as the difference

$$\Omega_S^{\text{QM/MM}} = E_{\text{QM/MM}}^S - E_{\text{QM/MM}}^0. \tag{36}$$

As in this the interactions between the quantum and classical regions are purely represented by electrostatic potentials, it is straightforward to combine (tr)PbE and classical polarizable embedding approaches into one quantum-quantum-classical embedding scheme. The static moments of the MM regions continue acting as an additional background potential to the quantum-quantum region. Similarly, the electric field acting on the polarizable sites in the MM region is created by the total electron density (plus the nuclei) of the subsystem QM region. It is worth highlighting that when a polarizable model is used in the PbE-GW-BSE/MM, the outer SCF coupling the quantum and classical regions implies that during such a calculation also the density of the inactive region can respond to the polarization of the MM region.

3 Computational Details

All calculations have been done using the VOTCA-XTP package[18,19] which interface to the ORCA software[45] for the full system reference DFT calculations. The def2-TZVP basis-set[46] with an optimized auxiliary basis[47] along with the PBE0 hybrid functional[48] has been used in all DFT and subsequent GW-BSE calculations. For the construction of localized orbitals required in the projection-based-embedding calculations, we employ the Pipek–Mezey (PM) localization scheme[49], which maximizes the atomic Mulliken population subject to the con-

Table 1: Overview of the number of basis functions (N_{basis}), number of functions in the auxiliary basis (N_{aux}), number of occupied states included in the quasiparticle calculation and the BSE product basis (N_{occ}), idem for virtual states (N_{virt}), number of transitions in the RPA (N_{RPA}), and dimension of the BSE Hamiltonian (N_{BSE}), for the full-GW-BSE, PbE-GW-BSE, and trPbE-GW-BSE calculations on the three testsystems.

	N_{basis}	N_{aux}	N_{occ}	N_{virt}	N_{RPA}	N_{BSE}
DPP+alkyl						
full-GW-BSE	1194	2940	115	229	124085	52670
PbE-GW-BSE	1194	2940	51	101	55029	10302
trPbE-GW-BSE	1152	2835	51	101	52887	10302
aqueous prodan						
full-GW-BSE	2994	7377	336	672	893088	451584
PbE-GW-BSE(dye)	2994	7377	61	121	162138	14762
PbE-GW-BSE	2994	7377	96	191	255168	36672
trPbE-GW-BSE(dye)	2453	6035	61	191	129137	14762
aqueous benzene-TCNE						
full-GW-BSE	1650	4062	183	365	268461	133590
PbE-GW-BSE	1650	4062	53	105	77751	11130
trPbE-GW-BSE(10^{-4})	1384	3400	53	105	63653	11130
trPbE-GW-BSE(10^{-5})	1638	4032	53	105	77115	11130

straint of keeping the total density fixed. For the actual maximization step, we make use of the unitary optimization algorithm as described by Lehtola and Jonsson[50]. If not stated otherwise, the G_0W_0 variant has been chosen with the Plasmon-Pole model[51] (PPM) for self-energy calculation in the GW step. To maintain consistency, we use in each case the full spectrum of single-particle states in the RPA (N_{RPA}), and consider all occupied (N_{occ}) and the lowest $N_{\text{virt}} = 2N_{\text{occ}} - 1$ unoccupied states for the calculation of the QP corrections and the expansion of the BSE product basis, i.e., $N_{\text{BSE}} = 2N_{\text{occ}}(2N_{\text{occ}} - 1)$. The explicit numbers for all three test systems are summarized in Table 1.

4 Results

4.1 General considerations for Projector-based-embedded GW-BSE calculations

While the DFT-in-DFT calculation can be shown to reproduce the full reference total energy exactly, we have seen that there are changes in the molecular orbitals, and we therefore cannot in general expect a GW-BSE calculation after PbE-DFT (from now on referred to for short as PbE-GW-BSE) to yield the same excitation energies as a full GW-BSE calculation. One can get an indication of what the general changes are by considering, e.g., the expressions for the exchange part Σ^{x} (Eq. (9)) and correlation part Σ^{c} (Eq. (10)) of the self-energy.

The exchange part is affected by (i) summing over fewer occupied states in the subsystem-GW-BSE calculation and (ii) the changes in the molecular orbitals themselves. Note that while Σ^{x} itself only depends on the occupied orbitals, it enters the quasiparticle energies of both occupied and unoccupied states as evaluated by Eq. (3). Therefore, even though the virtual molecular orbitals are unchanged in the subsystem-DFT calculation (using the full basis), their GW quasiparticle energies may have different contributions arising from Σ^{x}.

For the frequency-dependent correlation part, a similar analysis is more complicated, as the expression in Eq. (10) involves the single-electron Green's function (Eq. (5)), and the screened Coulomb interaction W determined with the help of the irreducible polarizability χ_0 as in Eq. (8). Subsystem embedding changes both G_1 (which also leads to the discussed changes in Σ^{x}) and χ_0 via the different orbitals and their energies. For polarizability, embedding implies several noteworthy modifications. Even if the $\phi(\mathbf{r})$ and energies ε were unchanged, the sum over occupied orbitals is limited to the active occupied orbitals (the ones from the inactive one are found in the virtual space at high energy, and should be excluded from the sum over virtual orbitals). As a result, the screening only has contributions from transitions between occupied orbitals in the active subsystem and virtual orbitals of the combined system (in the full basis calculation), while contributions from transitions from oc-

16

cupied orbitals in the inactive region to all virtual orbitals are removed. The inactive region therefore can be considered static from the perspective of the screened Coulomb interaction, similar to a QM/MM embedding with only static moments in the MM region, as discussed in Section 2.4. If additionally a truncated basis restricted ideally to the active region is used, this will also affect the virtual orbitals and essentially limit the transitions to those within the active region, removing charge-transfer-like transitions between the subsystems from the response. The effect on the calculated contribution of these CT-like transitions to Σ^{c} is expected to be small, however. Both considerations regarding the modifications type of transitions excluded in the screening in subsystem-GW are in general then combined with additional effects of changed orbitals and their energies in the active region. From the lack of screening from the now inactive region, one can generally expect the contributions of Σ^{c} to the quasiparticle energies to be smaller (in absolute values) in the subsystem-GW calculation compared to the full GW case. In other words, even when the orbitals themselves are only minimally affected (for weakly interacting, non-bonded molecular structures, for instance), one can expect to find the occupied (virtual) quasiparticle energies from the embedded calculation to be below (above) the ones from the full calculations. In particular, the HOMO-LUMO gap in subsystem-GW is then larger than the respective gap in full-system GW.

One can make similar examinations on the level of the BSE. Naturally, as the number of occupied orbitals is reduced, the electron-hole transitions used to expand the two-particle wavefunctions are limited to the transitions starting from the active subsystem. Any changes to the quasiparticle energies as a result of the points discussed above will directly impact the free transition term $D_{vc,v'c'}$ from Eq. (18) in the BSE Hamiltonian. The effects on the exchange and direct (screened) terms of the electron-hole interaction kernel $K^{\mathrm{x}}_{vc,v'c'}$ and $K^{\mathrm{d}}_{vc,v'c'}$ are similar to those discussed for Σ^{x} and Σ^{c}, respectively. Especially, the reduced screening can be expected to result in stronger electron-hole attraction compared to the full-system calculation and might in turn compensate to some degree the larger quasiparticle
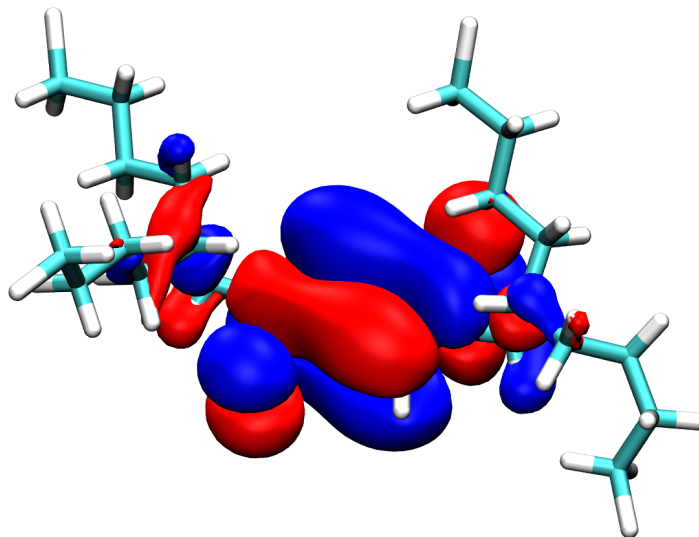
17

Figure 2: Isosurfaces of the KS HOMO (isovalues $\pm 0.01\,\mathrm{a_B^{-3}}$) of the single DPP bicyclic ring with branch alkyl side chain as obtained from regular KS-DFT.

gap in the free transition.

4.2 DPP bicyclic ring with branched alkyl side-chains

As a first test system, we consider a single DPP unit from the class of DPP2PymT polymers studied in this work. Alkyl side chains with a branched structure are attached to the respective nitrogen atoms. A short C_2H_2 group contained the branching point, and each branch is formed by C_4H_9, as can also be seen in Figure 1(a). The geometry of this structure is optimized in vacuum (DFT with the PBE0 functional and def2-TZVP basis). In (conjugated) polymer systems, it is often assumed that the frontier orbitals relevant for charge transport are localized on the actual functional backbone, and that the side-chains do not participate in the electronic processes. For the testing of the PbE-GW-BSE approach, such chemical intuition suggests to actually select only the DPP unit including the nitrogen atoms as the active region and the complete two branched alkyl side chains into the inactive one. However, as can be seen from the isosurfaces of the HOMO from a full KS calculation in Figure 2, the occupied frontier orbital extends further into the side chains, even slightly beyond the branching atom.
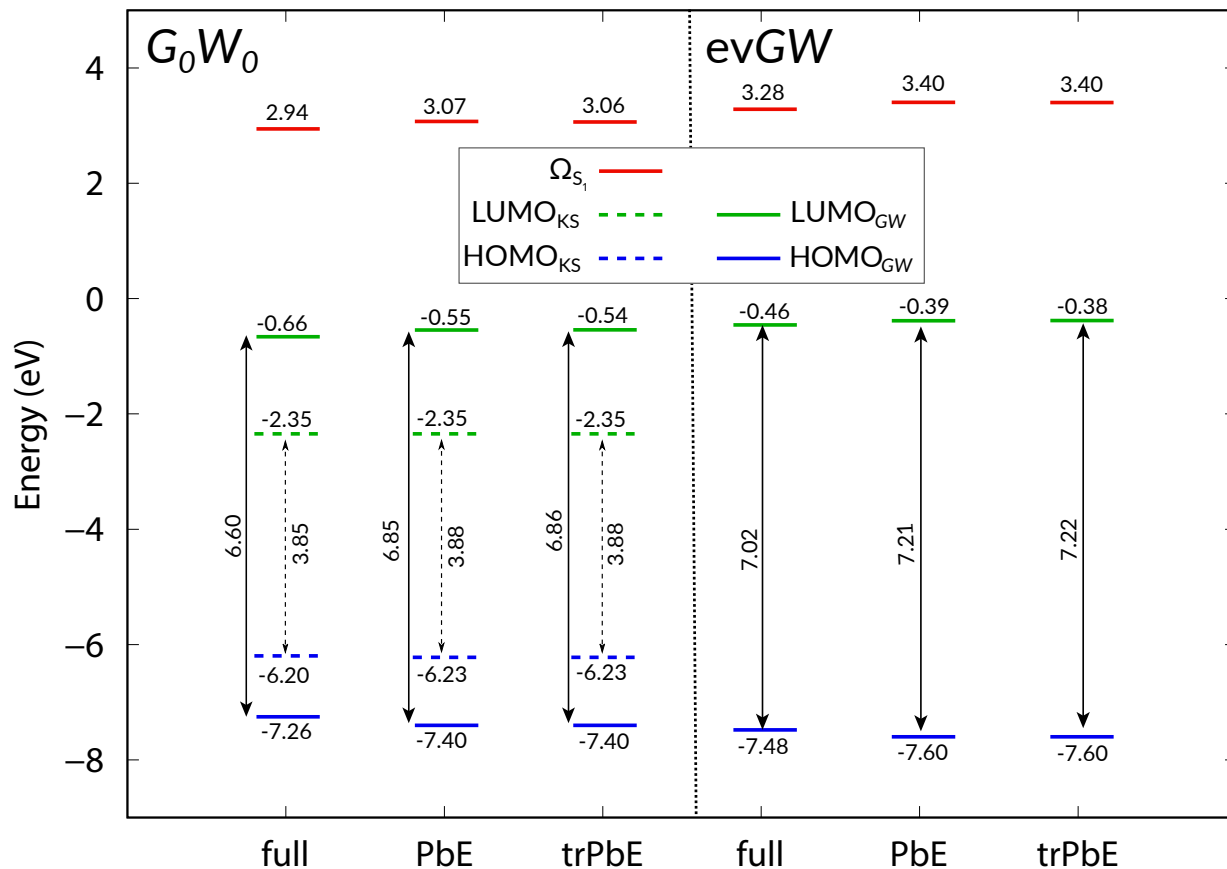
18

Figure 3: Energy level (in eV) diagram for a DPP bicyclic ring with branched alkly side-chains showing HOMO (blue), LUMO (green), and lowest electron-hole excitation (red) energies on KS level (dashed) or GW-BSE level (solid), as resulting from full, PbE, and trPBE calculations, respectively. Left panel shows results from G_0W_0 calculations, right panel from evGW calculations. Double-headed arrows additionally indicate the respective HOMO-LUMO gaps. See also Table 2 for details.

The choice of only the DPP core as the active region is therefore expected to yield considerable modifications to the occupied electronic states. With the objective to first inspect the different effects in the GW-BSE steps upon embedding without significant modification of the respective orbitals, we extended the active region to include the C_2H_2 groups until the respective branching carbon atom for all following PbE calculations. Figure 3 shows the results of G_0W_0 (left) and evGW (right) calculations for the HOMO, LUMO and Ω_{S_1} energies for full, PbE, and trPbE calculations, respectively. The results are also collected in Table 2, in which also the individual contributions to the quasiparticle energies according to

https://doi.org/10.26434/chemrxiv-2024-wp120 ORCID: https://orcid.org/0000-0002-6077-0467 Content not peer-reviewed by ChemRxiv. License: CC BY-NC-ND 4.0

Table 2: Results of full-G_0W_0-BSE, PbE-G_0W_0-BSE and trPbE-G_0W_0-BSE calculations for a DPP bicyclic ring with branched alkyl side-chains: KS and QP HOMO and LUMO energies together with the individual contributions from the exchange-correlation potential V^{xc}, and the self-energy split in Σ^x, and Σ^c, according to $\varepsilon^{QP} = \varepsilon^{KS} - V^{xc} + \Sigma^x + \Sigma^c$, as well as KS and QP HOMO-LUMO gaps. The lowest electron-hole excitation energy from the respective BSE calculation Ω_{S_1} is also split into the free transition energy (D) and contributions from the exchange (K^x) and direct (K^d) terms of the electron-hole interaction. All energies in eV.

		full	PbE	$\Delta_{\text{full}}^{\text{PbE}}$	trPbE	$\Delta_{\text{PbE}}^{\text{trPbE}}$
DFT	$\varepsilon_{\text{H}}^{\text{KS}}$	-6.20	-6.23	-0.03	-6.23	0.00
	$\varepsilon_{\text{L}}^{\text{KS}}$	-2.35	-2.35	0.00	-2.35	0.00
	$E_{\text{gap}}^{\text{KS}}$	3.85	3.88	0.03	3.88	0.00
G_0W_0	V_{H}^{xc}	-12.52	-12.47	0.05	-12.47	0.00
	$\Sigma_{\text{H}}^{\text{x}}$	-14.08	-14.02	0.06	-14.02	0.00
	$\Sigma_{\text{H}}^{\text{c}}$	0.50	0.38	-0.12	0.38	0.00
	V_{L}^{xc}	-11.25	-11.23	0.02	-11.23	0.00
	$\Sigma_{\text{L}}^{\text{x}}$	-7.52	-7.50	0.02	-7.50	0.00
	$\Sigma_{\text{L}}^{\text{c}}$	-2.04	-1.92	0.12	-1.92	0.00
	$\varepsilon_{\text{H}}^{\text{QP}}$	-7.26	-7.40	-0.14	-7.40	0.00
	$\varepsilon_{\text{L}}^{\text{QP}}$	-0.66	-0.55	0.11	-0.54	0.01
	$E_{\text{gap}}^{\text{QP}}$	6.60	6.85	0.25	6.86	0.01
BSE	D	7.63	8.00	0.37	8.00	0.00
	K^{x}	0.54	0.56	0.02	0.56	0.00
	K^{d}	-5.23	-5.49	-0.26	-5.50	-0.01
	Ω_{S_1}	2.94	3.07	0.13	3.06	-0.01

$\varepsilon^{QP} = \varepsilon^{KS} - V^{xc} + \Sigma^x + \Sigma^c$, the HOMO-LUMO gap, and contributions of the free transition energy (D), the exchange (K^x) and direct (K^d) terms of the electron-hole interaction to the electron-hole excitation energy Ω_{S_1} are listed. The quantity $\Delta_{\text{full}}^{\text{PbE}}$ is the difference of the PbE calculation to the full one, and $\Delta_{\text{PbE}}^{\text{trPbE}}$ measures the additional change of the basis set truncation with respect to the PbE calculation.

We first start the discussion of the KS HOMO and LUMO energies on DFT level only. They are shown in Figure 3 as dashed lines (blue: HOMO, green: LUMO) in the G_0W_0 panel (note that they are identical in the evGW case). With the choice of the active region as discussed above, the change in the energy of the HOMO level upon embedding is small (-0.03 eV), and the LUMO is unaffected as expected from the theoretical basis given in

20

Section 2.2. Using a truncated atomic orbital basis does not yield any changes of the KS electronic structure at the shown accuracy. As one can see from the values N_{basis} in Table 1, the basis truncation procedure as described in Section 2.3 removes only 42 of the 1194 functions, or about $3.5\,\%$.

Considering from now on the actual G_0W_0 results, one can first identify the typical effects of quasiparticle corrections on the HOMO and LUMO energies. In the full approach, the HOMO energy is lowered by $1.06\,\text{eV}$, and the LUMO energy raised by $1.69\,\text{eV}$, such that the HOMO-LUMO gap increases by $2.75\,\text{eV}$. Upon PbE, however, the respective shifts are more pronounced, by $0.14\,\text{eV}$ ($0.11\,\text{eV}$) for the HOMO (LUMO), cf. Table 2. The quasiparticle gap $E_{\text{gap}}^{\text{QP}}$ is hence increased by $0.25\,\text{eV}$. Within the GW formalism, a larger gap can often be associated with reduced screening. This notion is corroborated by the data provided for the contributions from V^{xc}, Σ^{x}, and Σ^{c} to ε^{QP} in Table 2. From the respective $\Delta_{\text{full}}^{\text{PbE}}$ for, e.g., the HOMO, one can see that the differences from the KS exchange-correlation potential and the exchange part of the self-energy almost compensate (note that V^{xc} is used with a negative sign in Eq. (3)), and that consequently the differences in quasiparticle energies between full- and PbE-G_0W_0 are practically determined by effects in the correlation part Σ^{c} alone. For the LUMO, very similar observations can be made. Again, due to the minimal reduction of the basis set, no significant changes are noted in trPbE-G_0W_0 compared to PbE-G_0W_0.

Turning now towards the respective BSE results, we first note in the full calculation that the excitation energy results as $2.94\,\text{eV}$, which consists of the effective free-transition energy $D = 7.63\,\text{eV}$, and the exchange ($K^{\text{x}} = 0.54\,\text{eV}$) and direct ($K^{\text{d}} = -5.23\,\text{eV}$) parts of the electron-hole interaction kernel. The fact that D exceeds $E_{\text{gap}}^{\text{QP}}$ by about one eV indicates that the electron-hole excitation is not exclusively given by a HOMO-LUMO transition. Qualitatively, the same holds also in the PbE (and trPbE) calculations. The S_1 energy in PbE-G_0W_0-BSE results with $3.07\,\text{eV}$ only $0.13\,\text{eV}$ higher than in the full reference. This is noteworthy because as discussed above $E_{\text{gap}}^{\text{QP}}$ is larger by almost twice this value. Upon inspection of the individual contributions to the BSE level given in Table 2, one first observes

21

that the free interlevel contribution D is larger by $0.37\,\text{eV}$ and exceeds the relative increase on $E_{\text{gap}}^{\text{QP}}$, indicating that the lack of screening in Σ^{c} is larger for states outside the fundamental gap. That Ω_{S_1} as obtained in the PbE approach is close compared to the full one despite these observations is due to the effect of the direct electron-hole interaction in the BSE Hamiltonian. This contribution, which is solely responsible for effective electron-hole binding, is with $K^{\text{d}} = -5.49\,\text{eV}$ in the PbE case stronger by $0.26\,\text{eV}$. As discussed in Section 2.1, the direct terms contains the screened Coulomb interaction W. Lack of screening from the inactive region then implies that the electron-hole pair is subject to a stronger, more bare-Coulomb like, electron-hole attraction. Such a stronger binding compensates at least to some extent (by $\sim 50\,\%$) the relatively larger energy differences of the free transitions.

In the right panel of Figure 3, the energy level diagram is given for results in which the evGW method is used in all three cases. Generally evGW leads here to a larger quasiparticle gap (by $0.42\,\text{eV}$) and a larger Ω_{S_1} (by $0.34\,\text{eV}$) than in G_0W_0 (full approach). Interestingly, the difference in $E_{\text{gap}}^{\text{QP}}$ upon PbE is here with $0.13\,\text{eV}$ somewhat smaller. Compared with a larger reference value in the full evGW calculation, this reduces its relative error from $3.8\,\%$ to $2.7\,\%$. For the first electron-hole excitation energy, the absolute deviation to the full calculation is with $0.12\,\text{eV}$ almost identical to the G_0W_0 case. Due to the larger reference value, we still find a slightly reduced relative deviation of $3.7\,\%$ compared to $4.4\,\%$. As before, basis truncation has no noticeable impact on the energy level diagram.

4.3   Prodan in water

The study on the DPP bicyclic ring with branched alkyl side chains in the previous section has indicated that the proper choice of the active region is important, that differences in the energy levels between full and PbE calculations are attributable to the lack of screening effects from the inactive region, and that basis truncation had a minimal effect. We will now turn to a different test system, to scrutinize if these findings are specific to the DPP system in which the active and inactive region were connected by a covalent bond. We will also present

22

and analyze the use of PbE in the GW-BSE/MM scenario for quantum-quantum-classical embedding.

The system we have chosen for this study consists of prodan, a polarity-sensitive dye, solvated in bulk water, as shown in Figure 1(b). Up excitation of the S_1 state absorption of a photon, the dipole moment of prodan in the excited state is significantly increased compared to the ground state. In polar solvents, such as water, additional screening effects originate from the structural relaxation of the solvent molecules, which in turn affect the excited state properties of the solute, and lead to a significant reduction of the emission energy. In Ref.[52], this process was simulated by an iterative GW-BSE+MD procedure. We take a single snapshot of one of those trajectories, and perform static and polarizable G_0W_0-BSE/MM calculations on this structure.

To begin with, the system is partitioned in quantum and classical regions. We assign the prodan molecule and any water molecule whose center-of-mass is within 1 nm of the solute center-of-mass to the QM region. As a result, the QM region contains 199 atoms (prodan and 55 water molecules), which can be considerably challenging for full GW-BSE calculations. This QM region is then embedded in another 1 nm wide shell, which contains 630 water molecules treated on MM level.

We begin with the results of full G_0W_0-BSE calculations in a static MM environment, as shown in the left panel of the energy level diagram in Figure 4 and data summarized in Table 3. From the computational details in Table 1 it can be inferred that the computational cost for such a calculation is significant. We will discuss this in more detail in Section 5.1 below. In the G_0W_0 step, we observe the typical lowering of the HOMO energy and increase of the LUMO energy with respect to the KS reference, resulting in a HOMO-LUMO gap of $E_{\text{gap}}^{\text{QP}} = 3.90\,\text{eV}$ as compared to $E_{\text{gap}}^{\text{KS}} = 1.70\,\text{eV}$. For the PbE calculations, we now first split the QM region into an active region containing only the dye molecule and an inactive region containing the 55 water molecules (setup indicated as "dye" in Table 1 and Table 3). One can see from the dashed lines in Figure 4 that the PbE lowers the HOMO energy by

23

Table 3: Results of full-G_0W_0-BSE/MM, PbE-G_0W_0-BSE/MM and trPbE-G_0W_0-BSE/MM calculations (all static) for aqueous prodan: KS and QP HOMO and LUMO energies together with the individual contributions from the exchange-correlation potential V^{xc}, and the self-energy split in Σ^{x}, and Σ^{c}, according to $\varepsilon^{\text{QP}} = \varepsilon^{\text{KS}} - V^{\text{xc}} + \Sigma^{\text{x}} + \Sigma^{\text{c}}$, as well as KS and QP HOMO-LUMO gaps. The lowest electron-hole excitation energy from the respective BSE calculation Ω_{S_1} is also split into the free transition energy (D) and contributions from the exchange (K^{x}) and direct (K^{d}) terms of the electron-hole interaction. All energies in eV.

|				dye			dye+water	
		full	**PbE**	$\boldsymbol{\Delta}_{\text{full}}^{\text{PbE}}$	**trPbE**	$\boldsymbol{\Delta}_{\text{PbE}}^{\text{trPbE}}$	**PbE**	$\boldsymbol{\Delta}_{\text{full}}^{\text{PbE}}$
DFT	$\varepsilon_{\text{H}}^{\text{KS}}$	-5.14	-5.22	-0.08	-5.45	-0.23	-5.15	-0.01
	$\varepsilon_{\text{L}}^{\text{KS}}$	-3.44	-3.44	0.00	-3.70	-0.26	-3.44	0.00
	$E_{\text{gap}}^{\text{KS}}$	1.70	1.78	0.08	1.75	-0.03	1.71	0.01
G_0W_0	V_{H}^{xc}	-13.02	-12.79	0.23	-12.80	-0.01	-12.97	0.05
	$\Sigma_{\text{H}}^{\text{x}}$	-16.08	-15.86	0.22	-15.85	0.01	-16.04	0.04
	$\Sigma_{\text{H}}^{\text{c}}$	2.13	1.98	-0.15	2.00	0.02	2.03	-0.10
	V_{L}^{xc}	-11.31	-10.93	0.38	-10.96	-0.03	-11.08	0.23
	$\Sigma_{\text{L}}^{\text{x}}$	-7.61	-7.21	0.40	-7.27	-0.06	-7.37	0.24
	$\Sigma_{\text{L}}^{\text{c}}$	-2.44	-2.15	0.29	-2.12	0.03	-2.23	0.21
	$\varepsilon_{\text{H}}^{\text{QP}}$	-6.07	-6.32	-0.25	-6.50	-0.18	-6.19	-0.12
	$\varepsilon_{\text{L}}^{\text{QP}}$	-2.17	-1.86	-0.31	-2.13	-0.27	-1.96	0.21
	$E_{\text{gap}}^{\text{QP}}$	3.90	4.46	0.56	4.37	-0.09	4.23	0.33
BSE	D	4.14	4.76	0.62	4.71	-0.05	4.52	0.38
	K^{x}	0.11	0.18	0.07	0.18	0.00	0.15	0.04
	K^{d}	-3.30	-3.77	-0.47	-3.86	-0.09	-3.64	-0.34
	Ω_{S_1}	0.96	1.17	0.21	1.03	-0.14	1.03	0.07
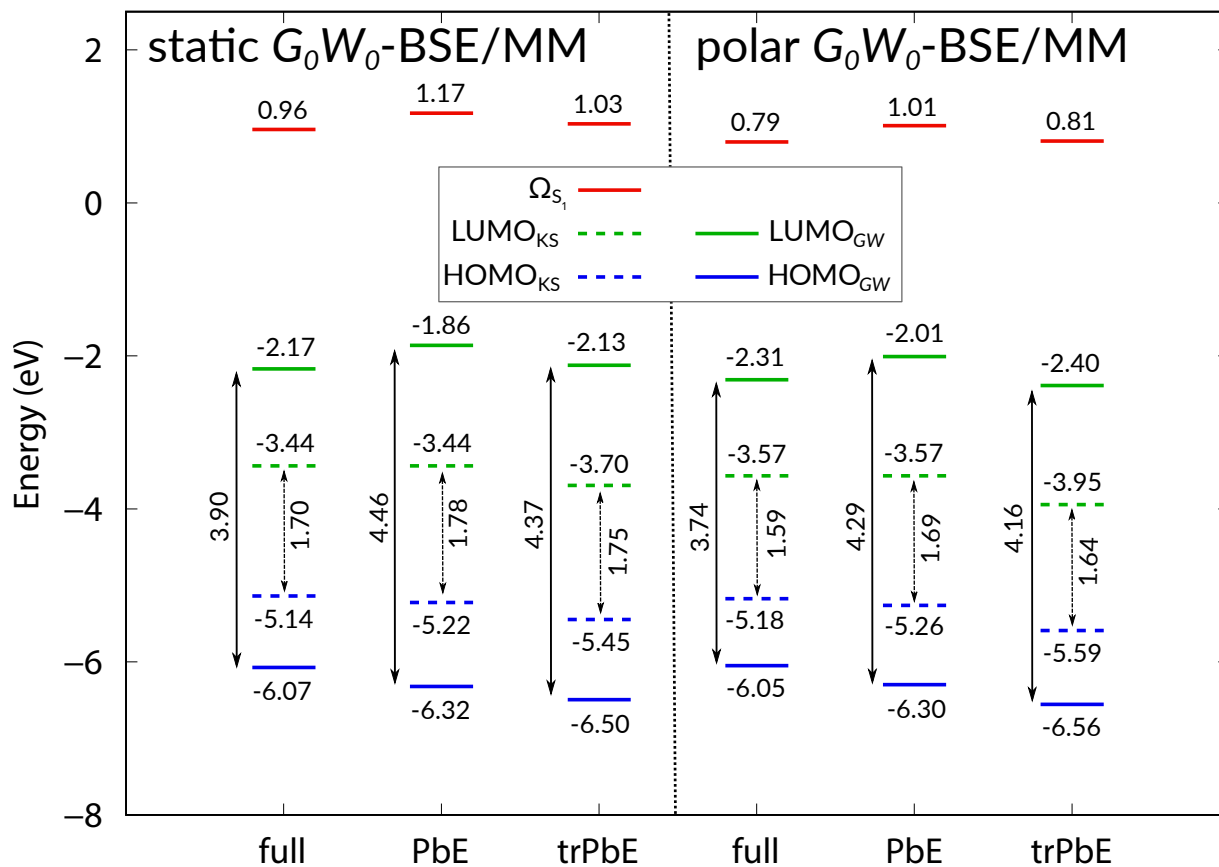
Figure 4: Energy level (in eV) diagram for aqueous prodan showing HOMO (blue), LUMO (green), and lowest electron-hole excitation (red) energies on KS level (dashed) or GW-BSE level (solid), as resulting from full, PbE, and trPBE calculations, respectively. Left panel shows results from static G_0W_0-BSE/MM calculations, right panel from polar G_0W_0-BSE/MM calculations. Double-headed arrows additionally indicate the respective HOMO-LUMO gaps. See also Table 3 for details.

0.08 eV and yields a correspondingly increased HOMO-LUMO gap as the LUMO energy remains constant. From the isosurface plot as shown in Figure 5, one can see that the full KS HOMO is not exclusively localized on the dye alone but exhibits contributions from some close water molecules. With the restriction of the active region to prodan only, the neglect of these contributions changes the electronic state noticeably. Even though there is no covalent bond formed, this is very similar to the case of the DPP structure in the previous section, in which the HOMO extended over some carbon atoms of the side chain.

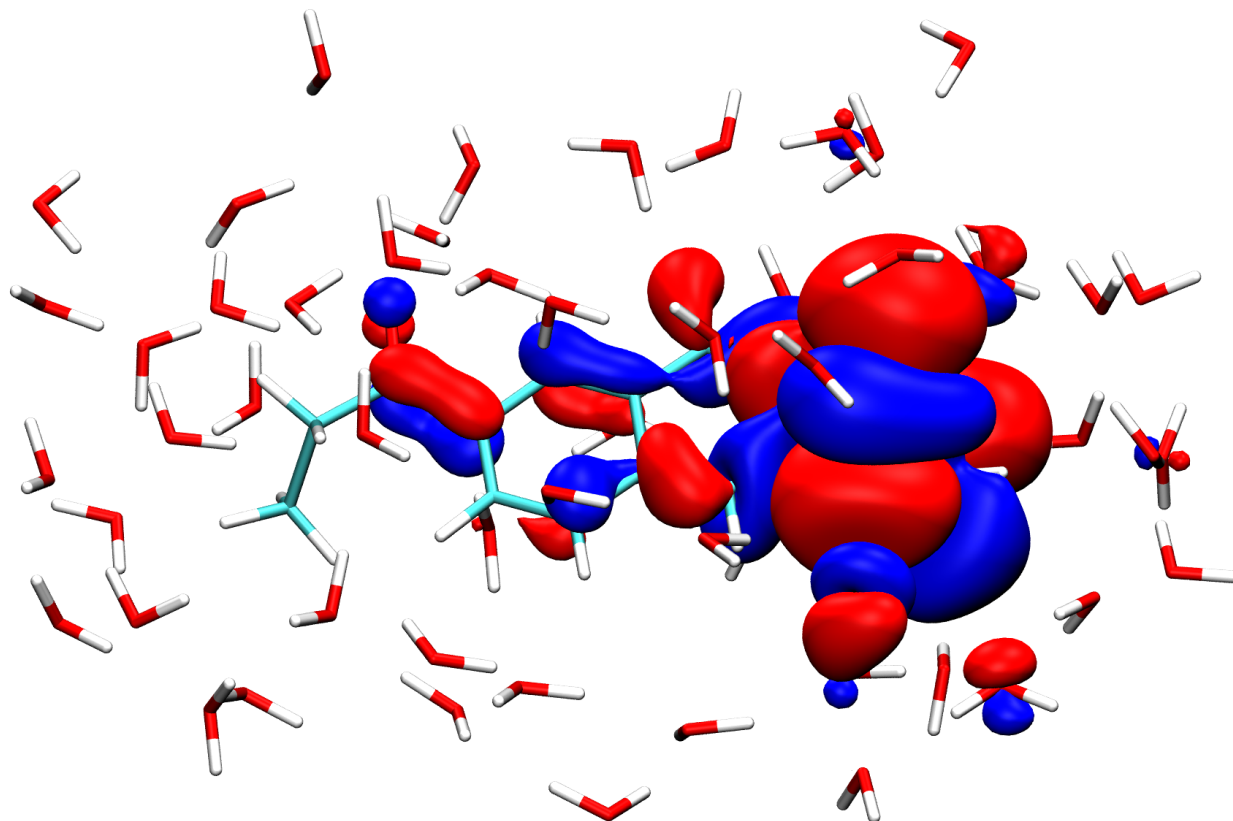The quasiparticle gap in static PbE-G_0W_0-BSE/MM results with 4.46 eV considerably

Figure 5: Isosurfaces of the KS HOMO (isovalues $\pm 0.01\,\mathrm{a_B^{-3}}$) of aqueous prodan chain as obtained from regular KS-DFT.

larger than in the full reference ($3.90\,\mathrm{eV}$), compatible with the earlier observations about reduced screening effects for the embedding. Indeed, the differences in the respective contributions from V^{xc} and Σ^{x} nearly compensate and the difference in quasiparticle corrections from full- to PbE-G_0W_0 mostly arises from effects in Σ^{c}. Basis set truncation removes 541 basis functions (see Table 1) in aqueous prodan and has a more pronounced impact on the electronic levels. As can be seen in Figure 4 and Table 3, both HOMO and LUMO levels on KS level result lower in energy in trPbE as compared to the reference system (by $0.31\,\mathrm{eV}$ and $0.26\,\mathrm{ev}$, respectively) and also to the respective PbE results. As a consequence of the almost constant shift, $E_{\mathrm{gap}}^{\mathrm{KS}}$ is only $0.05\,\mathrm{eV}$ larger. On G_0W_0 level, we note that the differences in the contributions in the quasiparticle corrections are relatively small and lead to only a small reduction of the quasiparticle gap by $0.09\,\mathrm{eV}$. On BSE level, the calculated $\Omega_{\mathrm{S_1}}$ energy varies from $0.96\,\mathrm{eV}$ (full) via $1.17\,\mathrm{eV}$ (PbE) to $1.03\,\mathrm{eV}$ (trPbE). Inspection of the respective

26

contributions to the BSE energy from the data in Table 3 reveals the same qualitative behaviour as discussed in Section 4.2: the lack of screening from the inactive region increases the contribution from free transitions D, which is compensated to some extent by the for the same reason also increased electron-hole attraction in K^{d} ($\sim 0.5\,\mathrm{eV}$).

In the last two columns of Table 3, we additionally show results of PbE-G_0W_0-BSE results in which the six water molecules with substantial contribution to the full system HOMO orbital are moved from the inactive to the active region. As can be seen, the change in the HOMO energy is very small compared to the full system, indicating that the more extended active region is a suitable choice. In this scenario, the differences in the quasiparticle energies for the HOMO and LUMO are smaller, but still amount to a $E_{\mathrm{gap}}^{\mathrm{QP}}$ increased by $0.33\,\mathrm{eV}$. Different choices for the active/inactive region splitting do not affect the qualitative observation that this difference is effectively only given by Σ^{c} contributions. In the BSE calculation, the lowest electron-hole excitation energy is obtained as $1.03\,\mathrm{eV}$, only $0.07\,\mathrm{eV}$ higher than in the full G_0W_0-BSE calculation. Here, the stronger by $-0.34\,\mathrm{eV}$ electron-hole interaction in the BSE kernel compensates the larger by $0.38\,\mathrm{eV}$ free transition contribution.

When polarizable GW-BSE/MM is employed, the respective calculations of the coupled system in the self-consistent reaction field of the MM environment become state dependent. What is shown as polar G_0W_0-BSE/MM in the right panel of Figure 4 are the energy levels in the final step of a self-consistent procedure to evaluate the total energy of the S_1 excited state according to Eqs. (35)-(36). Note also that the PbE and trPbE results shown there are for the case in which only the prodan molecule is in the active region, as in the left panel. We refrain from analyzing the shown data in detail because even though the exact numbers are different, there are no fundamental differences in what has been observed for static G_0W_0-BSE. In other words, while the external potential is different in both cases, the intrinsic effects going from full to PbE or trPbE calculations are the same.

27

4.4 Benzene-TCNE dimer in water

The final system under consideration is a dimer of benzene and TCNE, as it is known to exhibit intermolecular charge-transfer type excitations in which the hole is predominantly located on benzene and the electron on TCNE molecule[53]. In a polar solvent, these CT excitations are massively lowered in energy compared to the vacuum case, and a proper treatment of the polarizable environment is essential. Not only is the type of electron-hole excitation different to the ones studied in Section 4.2 and Section 4.3 but also the localization on the contributing frontier orbitals. We scrutinize in the following if the observations regarding PbE- or trPbE-GW-BSE calculations made for the previous two test systems also hold in case of intermolecular excitations. To this end, we prepared first a dimer of benzene and TCNE molecules stacked with a separation of 3.7 Å. This initial structure was then solvated with water using `packmol`[54,55]. From this solvated system, the benzene-TCNE dimer and the 26 closest water molecules have been selected for the following calculations. Note that we are only interested in the trends of effects from using PbE or trPbE here, so a more involved procedure to obtain relaxed atomic positions is not required for this purpose.

In Figure 6 we show the resulting energy level diagram for the different variants of G_0W_0-BSE calculations on this benzene-TCNE dimer, including a reference calculation for the dimer in vacuum, see also Table 4. In the vacuum reference, the HOMO-LUMO gap is increased in G_0W0 by 3.44 eV compared to the KS value, and the CT excitation energy is obtained as 3.01 eV. When embedded in water, the full G_0W_0-BSE calculation exhibits a reduced $E_{\mathrm{gap}}^{\mathrm{KS}}$ (by 0.54 eV), which is predominantly caused by a shift in the HOMO level. The G_0W_0 gap is even reduced compared to vacuum by 0.81 eV, but interestingly here we also observe a downward shift of the LUMO energy albeit only by 0.19 eV compared to the upward shift by 0.62 eV of the HOMO. From the isosurfaces of the HOMO and LUMO in the full-G_0W_0 calculation as shown in Figure 7 one can clearly see the donor-acceptor character of the dimer in the distribution of the frontier orbitals on the respective molecules, corroborating the notion that a transition from HOMO to LUMO is of charge-transfer character. Its CT
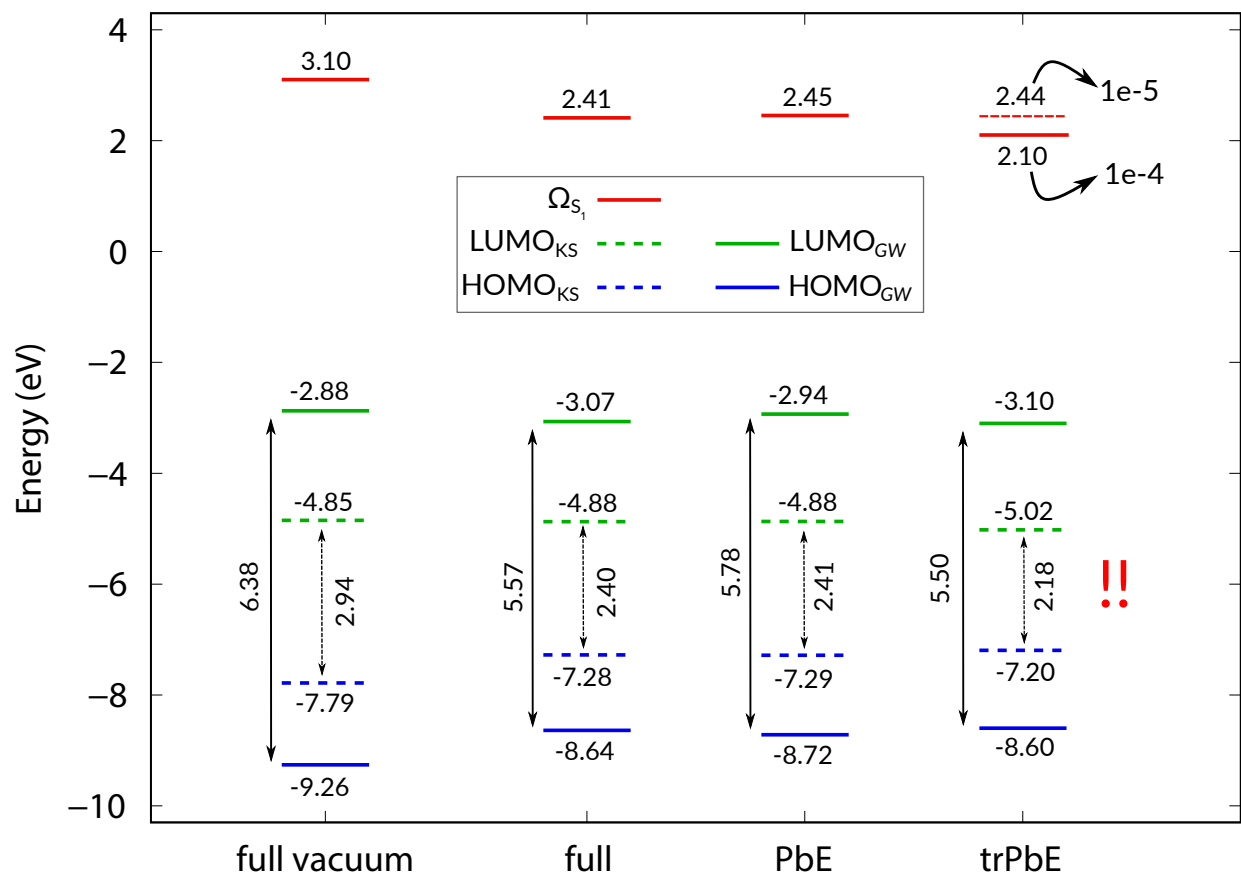
Figure 6: Energy level (in eV) diagram for benzene-TCNE dimer solvated in water showing HOMO (blue), LUMO (green), and lowest electron-hole excitation (red) energies on KS level (dashed) or G_0W_0-BSE level (solid), as resulting from full, PbE, and trPBE calculations, respectively. Double-headed arrows additionally indicate the respective HOMO-LUMO gaps. See also Table 4 for details.

excitation energy is obtained as $2.41\,\text{eV}$, lower by $0.69\,\text{eV}$ than in the vacuum case. It is noteworthy that this lowering of the CT energy is caused mainly by the reduction of the contribution from the free transition term D ($6.41\,\text{eV}$ in vacuum vs. $5.59\,\text{eV}$ in water) while the direct part of the electron-hole interaction is just reduced by $0.12\,\text{eV}$ due to the additional screening from the environment.

From Figure 7 it can also be seen that both frontier orbitals are not completely localized on the respective molecules, as the HOMO has minimal contributions at the TCNE and some close water molecules. The LUMO shows a similar pattern. For the PbE and trPbE calculations, we choose only the dimer as the active region. As one can see in Figure 6 and

29

Table 4: Results of full-G_0W_0-BSE, PbE-G_0W_0-BSE and trPbE-G_0W_0-BSE calculations for a benzene-TCNE dimer in water: KS and QP HOMO and LUMO energies together with the individual contributions from the exchange-correlation potential V^{xc}, and the self-energy split in Σ^{x}, and Σ^{c}, according to $\varepsilon^{\mathrm{QP}} = \varepsilon^{\mathrm{KS}} - V^{\mathrm{xc}} + \Sigma^{\mathrm{x}} + \Sigma^{\mathrm{c}}$, as well as KS and QP HOMO-LUMO gaps. The lowest electron-hole excitation energy from the respective BSE calculation Ω_{S_1} is also split into the free transition energy (D) and contributions from the exchange (K^{x}) and direct (K^{d}) terms of the electron-hole interaction. All energies in eV.

		vacuum	full	PbE	$\mathbf{\Delta}_{\mathrm{full}}^{\mathrm{PbE}}$	trPbE	$\mathbf{\Delta}_{\mathrm{PbE}}^{\mathrm{trPbE}}$
DFT	$\varepsilon_{\mathrm{H}}^{\mathrm{KS}}$	-7.79	-7.28	-7.29	-0.01	-7.20	0.09
	$\varepsilon_{\mathrm{L}}^{\mathrm{KS}}$	-4.85	-4.88	-4.88	0.00	-5.02	-0.14
	$E_{\mathrm{gap}}^{\mathrm{KS}}$	2.94	2.40	2.41	0.01	2.18	-0.23
G_0W_0	$V_{\mathrm{H}}^{\mathrm{xc}}$	-10.40	-10.46	-10.40	0.06	-10.43	-0.03
	$\Sigma_{\mathrm{H}}^{\mathrm{x}}$	-11.97	-12.02	-11.95	0.07	-11.95	0.00
	$\Sigma_{\mathrm{H}}^{\mathrm{c}}$	0.10	0.20	0.12	-0.08	0.11	-0.01
	$V_{\mathrm{L}}^{\mathrm{xc}}$	-11.54	-11.56	-11.52	0.04	-11.53	-0.01
	$\Sigma_{\mathrm{L}}^{\mathrm{x}}$	-7.88	-7.93	-7.89	0.04	-7.92	-0.03
	$\Sigma_{\mathrm{L}}^{\mathrm{c}}$	-1.68	-1.82	-1.69	0.13	-1.70	-0.01
	$\varepsilon_{\mathrm{H}}^{\mathrm{QP}}$	-9.26	-8.64	-8.72	-0.08	-8.61	0.11
	$\varepsilon_{\mathrm{L}}^{\mathrm{QP}}$	-2.88	-3.07	-2.94	0.13	-3.10	-0.16
	$E_{\mathrm{gap}}^{\mathrm{QP}}$	6.38	5.57	5.78	0.21	5.51	-0.27
BSE	D	6.40	5.59	5.80	0.21	5.52	-0.28
	K^{x}	0.01	0.01	0.01	0.00	0.01	0.00
	K^{d}	-3.31	-3.18	-3.36	-0.18	-3.43	-0.07
	Ω_{S_1}	3.10	2.42	2.45	0.03	2.10	-0.35

Table 4, the associated neglect of the contributions of the water molecules to the HOMO has very little effect on the KS HOMO energy in PbE, and concomitantly, the KS HOMO-LUMO gap. The quasiparticle energies in PbE-G_0W_0 are 0.08 eV lower for the HOMO and 0.13 eV higher for the LUMO, compared to the full-G_0W_0 result, and as a consequence the gap is larger by 0.21 eV. As observed for the DPP system and aqueous prodan, the differences in the respective contributions to the quasiparticle corrections from V^{xc} and Σ^{x} mostly cancel out. The CT excitation energy in PbE-G_0W_0-BSE is with 2.45 eV only 0.03 eV larger than in the full calculation, as the comparatively stronger electron-hole interaction almost completely compensates the larger free quasiparticle transition energies. Upon truncation of the basis set, we find unexpectedly larger deviations in trPbE-G_0W_0. Already on Kohn–Sham level, we find that the HOMO energy is increased to -7.20 eV and the LUMO lowered to 5.02 eV,
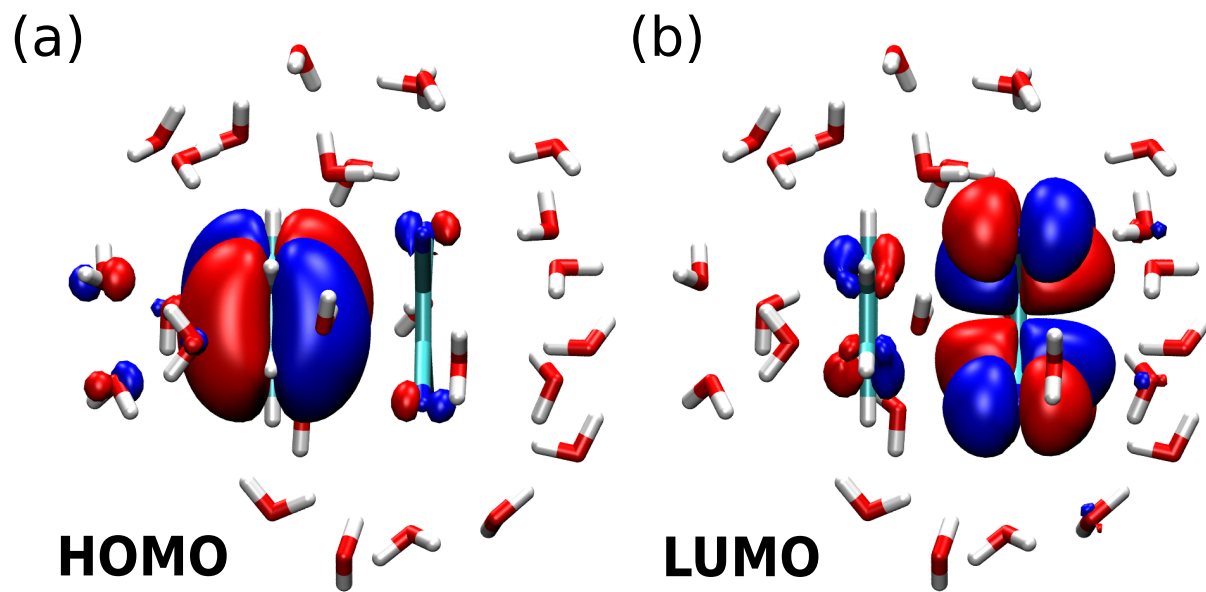
Figure 7: Isosurfaces of the (a) HOMO and (b) LUMO (isovalues $\pm 0.01\,\mathrm{a_B^{-3}}$) of the aqueous benzene-TCNE dimer as obtained from regular KS-DFT.

leading to E_{gap} being reduced by $0.23\,\mathrm{eV}$. At quasiparticle level, the gap energy is reduced by $0.27\,\mathrm{eV}$ as compared to the untruncated PbE case, so that the quasiparticle gap is very close to the full-G_0W_0 result. This should be considered coincidental. Indeed, the energy of the CT excitation is also lower by $0.35\,\mathrm{eV}$ and is as a consequence $0.32\,\mathrm{eV}$ smaller than in the full calculation reference, as the electron-hole interaction remains under-screened.

5   Discussion

We begin the discussion of the analysis of the PbE and trPbE techniques in application to the GW-BSE methodology by taking a broader view on the obtained quasiparticle energies than just the frontier orbital and fundamental gap energies. To this end we show in Figure 8 a comparison between the density of states (DOS) for the three systems studied in this work as obtained from full and PbE calculations on KS and G_0W_0 levels. On KS level (grey shaded area: full; black line: PbE) one can see in all three cases the frontier orbital peaks are well reproduced, and by construction also the full range derived from unoccupied orbitals.
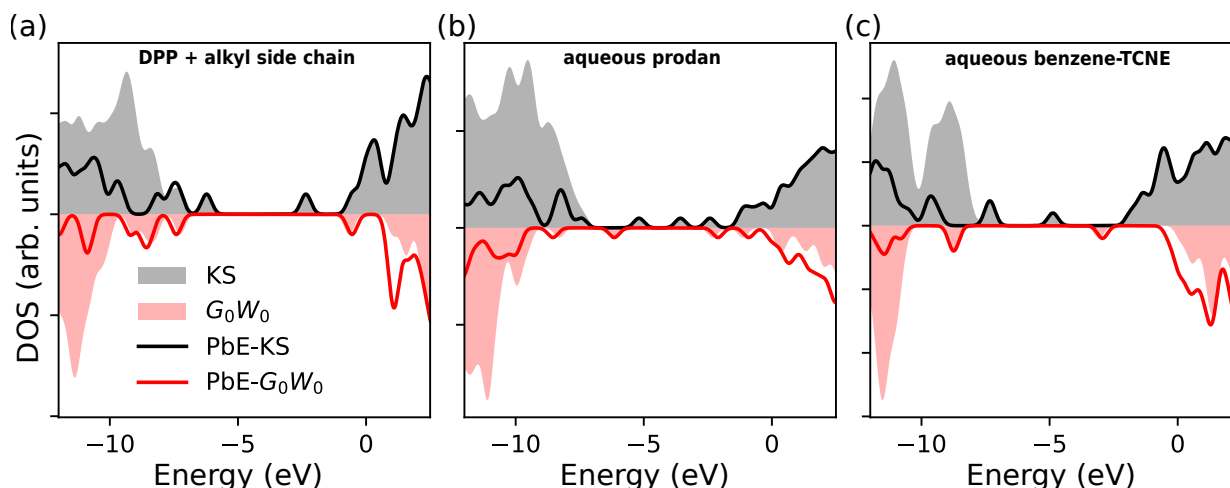
31

Figure 8: Density of states (DOS) for the three systems studied in this work: (a) DPP with branched alkyl side chains, (b) aqueous prodan, and (c) aqueous benzene-TCNE. Grey (red) shaded areas show the DOS as obtained by full KS (G_0W_0) calculations, while the black (red) lines indicate the respective PbE-KS (PbE-G_0W_0) DOS. A Gaussian broadening with standard deviation $0.2\,\text{eV}$ is used in all cases.

For the energy region lower than the respective HOMO energy, one can generally observe the presence of fewer states, as expected. Especially for the two water-solvated systems a significant part of the full DOS is removed by the embedding. When comparing the same data obtained on G_0W_0 level of theory, one can see the small deviations of the frontier orbital energies in PbE-G_0W_0 with respect to the full calculation. For lower energy occupied and higher energy unoccupied levels, the comparison is not so straightforward. When one compares PbE-G_0W_0 to PbE-KS, one can see similarities in the broad shape of the DOS, but also that not all orbital energies experience the same quasiparticle corrections. This seems to affect the virtual DOS above the LUMO more significantly. As one can see in PbE-G_0W_0 results in Figure 8(a) and (c) in particular, there are peaks in the DOS below those the full calculation DOS. This is an indication that the QP corrections for these levels, whose KS reference energy is the same in full- and PbE-KS, are less pronounced when an embedded calculation is performed. This is different to the observation that quasiparticle corrections are generally stronger for the frontier orbitals due to the lack of screening.

To elucidate we consider in more detail the differences between the PbE and full calcu-

32

Table 5: Results of full-G_0W_0-BSE/MM, PbE-G_0W_0-BSE/MM calculations (all polar) for aqueous prodan: KS and QP LUMO+1 and LUMO+2 energies together with the individual contributions from the exchange-correlation potential V^{xc}, and the self-energy split in Σ^{x}, and Σ^{c}, according to $\varepsilon^{\mathrm{QP}} = \varepsilon^{\mathrm{KS}} - V^{\mathrm{xc}} + \Sigma^{\mathrm{x}} + \Sigma^{\mathrm{c}}$. All energies in eV.

	LUMO+1			LUMO+2		
	full	**PbE**	$\boldsymbol{\Delta}^{\mathrm{PbE}}_{\mathrm{full}}$	**full**	**PbE**	$\boldsymbol{\Delta}^{\mathrm{PbE}}_{\mathrm{full}}$
$\varepsilon^{\mathrm{KS}}$	-2.32	-2.32	0.00	-1.06	-1.06	0.00
V^{xc}	-10.72	-10.47	0.25	-8.19	-5.91	2.29
Σ^{x}	-6.76	-6.52	0.24	-4.53	-3.11	1.43
Σ^{c}	-2.66	-1.58	0.24	-2.23	-1.58	0.65
$\varepsilon^{\mathrm{QP}}$	-1.02	-0.79	0.23	0.37	0.16	-0.21

lation split among the different contributions to the QP energies as done for HOMO and LUMO in Table 3, now for LUMO+1 and LUMO+2 from the final step in the polarizable G_0W_0-BSE/MM calculations. The results are summarized in Table 5. For the LUMO+1, we find qualitatively the same behavior as for the LUMO as discussed in Section 4.3: the contributions from V^{xc} and Σ^{x} nearly cancel out, and the too weak screening in PbE leads to a 0.23 eV higher quasiparticle energy as compared to the full calculation. For LUMO+2, the same does not hold: the difference in contributions from V^{xc} and Σ^{x} is significant, and contributes with -0.86 eV to the difference in quasiparticle energies. The contribution from the difference Σ^{x} is with 0.65 eV positive, consistent with the argument of too weak screening. Its magnitude is however much bigger than for LUMO and LUMO+1. In total, we find a quasiparticle energy that is *lower* by 0.21 eV in PbE as compared to the full calculation.

Inspecting the nature of the respective states might provide an indication of the origin of this different behavior. In Figure 9 we shows isosurfaces of LUMO+1 and LUMO+2, respectively. LUMO+1 (like LUMO) is predominantly localized on the dye molecule itself, with only small contributions from some close water molecules. In contrast, LUMO+2 is markedly different and extends to a large amount onto water molecules, that is into the inactive region from the point of view of projection-based-embedding. It stands to reason that in such a situation the removal of the occupied states of the inactive region from the expression for the exchange part of the self-energy introduces additional deviations.
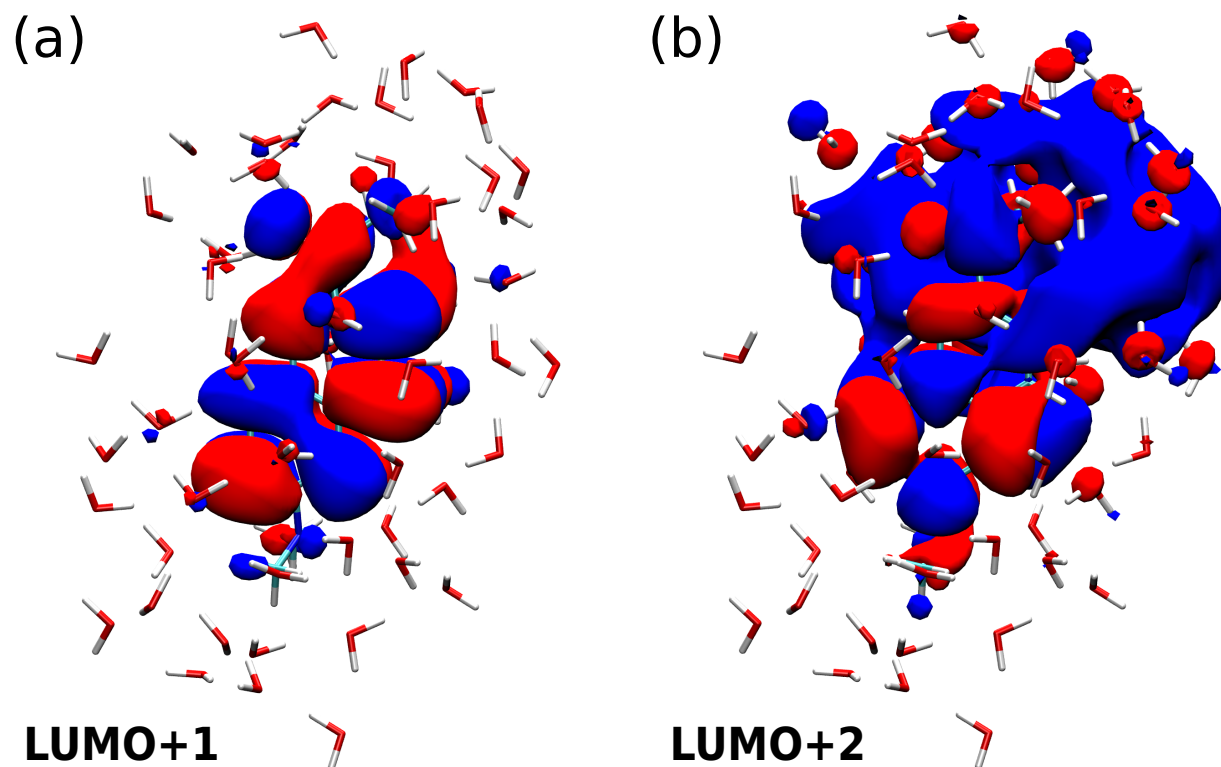
33

Figure 9: Isosurfaces of the (a) LUMO+1 and (b) LUMO+2 (isovalues $\pm 0.01\,\mathrm{a_B^{-3}}$) of the aqueous prodan as obtained from regular KS-DFT.

5.1 Embedding cost

In addition to the quality of the results of the different PbE-G_0W_0-BSE calculations compared to the full calculations, we consider the respective computational costs using the aqueous prodan system from Section 4.3 as an example. As can be seen from Table 1, large savings can be expected in the RPA steps (reduction of the number of transitions to less than 20 %), and in the BSE solution (reduction of the product basis size to about 3 %). Figure 10(a) shows the runtime of the respective calculation steps. For all the three variants, the underlying DFT calculation on the full system is performed with ORCA in about 1100 s. Localizing all 336 occupied orbitals with the unitary optimization requires in VOTCA-XTP around 1300 s, and performing the PbE-KS around 9600 s. Note that the internal DFT implementation in VOTCA-XTP is not optimized for performance and intended for development purposes instead. As expected, the most significant saving in computational time is the actual G_0W_0
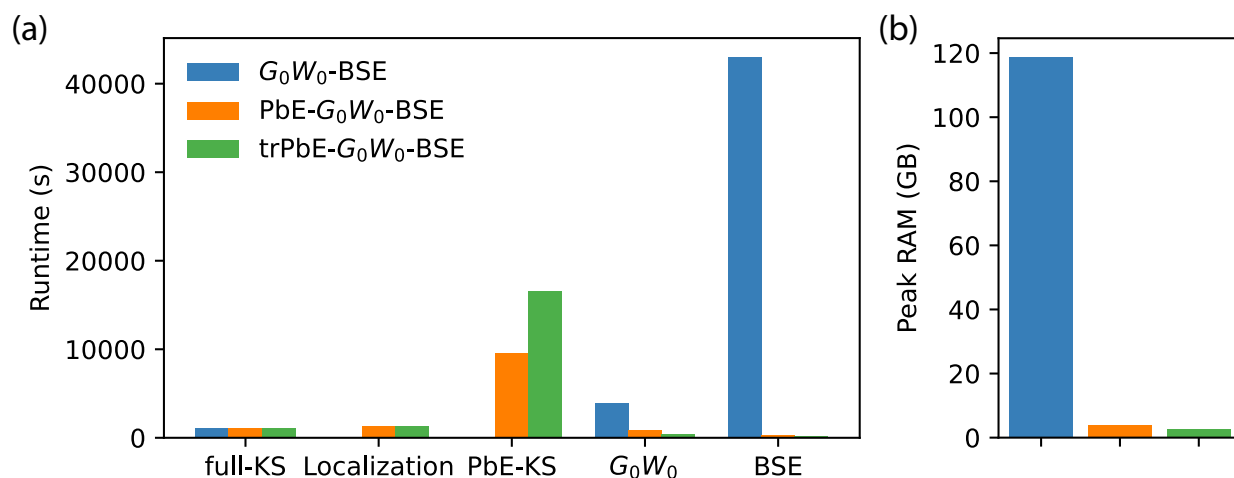
34

Figure 10: Computational costs of the different calculation steps in full-, PbE-, and trPbE-G_0W_0-BSE calculations for the aqueous prodan system from Section 4.3. (a) Runtime (in s) on 28 threads of an Intel(R) Xeon(R) Gold 5120 CPU @ 2.20GHz. (b) Peak RAM consumption in GB.

and BSE steps of the procedure. For the former, the reduced number of transitions in the RPA, combined with a lower number of states for which quasiparticle corrections have to be determined, reduced the time from 3900 s to around 800 s. Note that the absolute cost of the G_0W_0 step is, in fact, reasonably small (only a factor ~ 3.5 in the full calculation) compared to the DFT step, due to the use of the Plasmon-Pole Model for the frequency-dependence of the self-energy, which requires the explicit evaluation of the microscopic dielectric function only for two frequencies, see also Section 2.1. The most dramatic compute time reduction is observed for the BSE step. Here PbE leads to a reduction from 43000 s to just 270 s due to the massively reduced dimension of the product basis. Basis set truncation allows for another reduction of runtimes in G_0W_0 and BSE by a factor of 2, respectively. Apart from runtimes, the size of accessible systems for the GW-BSE approach is often also limited by the peak memory consumption, at least in the implementation in VOTCA-XTP. From the data shown in Figure 10(b), it is clear that the full and PbE approaches differ vastly in peak memory consumption. Embedding and the subsequent reduction in the in-memory storage of three-center Coulomb integrals after contraction with molecular orbitals requires only 3.8 GB of RAM compared to 118.7 GB. The truncated basis has a smaller effect on top

35

of this (2.5 GB). This clearly shows that PbE techniques can remove some computational bottlenecks of GW-BSE calculations, at the price of some deviations in the obtained results due to the lack of screening contributions from the inactive region.

6 Summary

In this paper, we have introduced and scrutinized projection-based-embedding techniques of GW-BSE calculations. Based on the analysis of the three test systems DPP ring with branched alkyl side chains, aqueous prodan, and an aqueous benzene-TCNE dimer, we could see that PbE can offer significant computational gains, making larger systems accessible to the many-body Green's functions based methodology. We have demonstrated that it can also be directly incorporated in quantum-classical embedding (GW-BSE/MM) schemes. We also found that the agreement with full calculations depends on the choice of the active region and is subject to effects from the neglect of screening contributions from the inactive electrons in the GW steps, which leads generally to an increased quasiparticle HOMO-LUMO gap. It was also noted that the lack of screening is in part compensated in the BSE as it manifests itself in an increased electron-hole attraction, so that deviations from full results for the electron-hole excitation energy are on the order of $0.1\,$eV for the different types of excitations studied here. It should be noted that all excitations are near-gap excitations, and it can not be guaranteed that the same quantitative agreement will hold for higher-energy excitations. Finally, we have seen that additional truncation of the basis set can reduce the computational costs by a factor of two with respect to full-basis PbE, but results appear to be sensitive to the chosen threshold values for removing basis functions.

7 acknowledgement

References

(1) Onida, G.; Reining, L.; Rubio, A. Electronic excitations: density-functional versus many-body Green's-function approaches. *Rev. Mod. Phys.* **2002**, *74*, 601–659.

(2) Ma, Y.; Rohlfing, M.; Molteni, C. Excited states of biological chromophores studied using many-body perturbation theory: Effects of resonant-antiresonant coupling and dynamical screening. *Phys. Rev. B* **2009**, *80*, 241405.

(3) Baumeier, B.; Andrienko, D.; Ma, Y.; Rohlfing, M. Excited States of Dicyanovinyl-Substituted Oligothiophenes from Many-Body Green's Functions Theory. *J. Chem. Theory Comput.* **2012**, *8*, 997–1002.

(4) Baumeier, B.; Rohlfing, M.; Andrienko, D. Electronic excitations in push-pull oligomers and their complexes with fullerene from many-body Green's functions theory with polarizable embedding. *J. Chem. Theory Comput.* **2014**, *10*, 3104–3110.

(5) Duchemin, I.; Guido, C. A.; Jacquemin, D.; Blase, X. The Bethe-Salpeter formalism with polarisable continuum embedding: Reconciling linear-response and state-specific features. *Chemical Science* **2018**, *9*, 4430–4443.

(6) Förster, A.; Visscher, L. Quasiparticle Self-Consistent GW-Bethe–Salpeter Equation Calculations for Large Chromophoric Systems. *J. Chem. Theory Comput.* **2022**, *18*, 6779–6793.

(7) Blase, X.; Attaccalite, C. Charge-Transfer Excitations in Molecular Donor-Acceptor

Complexes within the Many-Body Bethe-Salpeter Approach. *Appl. Phys. Lett.* **2011**, *99*, 171909.

(8) Sharifzadeh, S.; Darancet, P.; Kronik, L.; Neaton, J. B. Low-Energy Charge-Transfer Excitons in Organic Solids from First-Principles: The Case of Pentacene. *J. Phys. Chem. Lett.* **2013**, *4*, 2197–2201.

(9) Schirmer, J. Beyond the random-phase approximation: A new approximation scheme for the polarization propagator. *Phys. Rev. A* **1982**, *26*, 2395–2416.

(10) Christiansen, O.; Koch, H.; Jørgensen, P. The second-order approximate coupled cluster singles and doubles model CC2. *Chemical Physics Letters* **1995**, *243*, 409–418.

(11) Sundaram, V.; Lyulin, A. V.; Baumeier, B. Development and Testing of an All-Atom Force Field for Diketopyrrolopyrrole Polymers with Conjugated Substituents. *J. Phys. Chem. B* **2020**, *124*, 11030–11039.

(12) Sundaram, V.; Lyulin, A. V.; Baumeier, B. Effect of Solvent Removal Rate and Annealing on the Interface Properties in a Blend of a Diketopyrrolopyrrole-Based Polymer with Fullerene. *J. Phys. Chem. B* **2022**, *126*, 7445–7453.

(13) Baral, S.; Phillips, M.; Yan, H.; Avenso, J.; Gundlach, L.; Baumeier, B.; Lyman, E. Ultrafast Formation of the Charge Transfer State of Prodan Reveals Unique Aspects of the Chromophore Environment. *The Journal of Physical Chemistry B* **2020**, *124*, 2643–2651.

(14) Tirimbò, G.; de Vries, X.; Weijtens, C. H. L.; Bobbert, P. A.; Neumann, T.; Coehoorn, R.; Baumeier, B. Quantitative Predictions of Photoelectron Spectra in Amorphous Molecular Solids from Multiscale Quasiparticle Embedding. *Phys. Rev. B* **2020**, *101*, 035402.

(15) Duchemin, I.; Jacquemin, D.; Blase, X. Combining the GW formalism with the polarizable continuum model: A state-specific non-equilibrium approach. *J. Chem. Phys.* **2016**, *144*, 164106.

(16) Li, J.; D'Avino, G.; Duchemin, I.; Beljonne, D.; Blase, X. Combining the Many-Body GW Formalism with Classical Polarizable Models: Insights on the Electronic Structure of Molecular Solids. *J. Phys. Chem. Lett.* **2016**, *7*, 2814–2820.

(17) Li, J.; D'Avino, G.; Duchemin, I.; Beljonne, D.; Blase, X. Accurate description of charged excitations in molecular solids from embedded many-body perturbation theory. *Phys. Rev. B* **2018**, *97*, 35108.

(18) Wehner, J.; Brombacher, L.; Brown, J.; Junghans, C.; Çaylak, O.; Khalak, Y.; Madhikar, P.; Tirimbò, G.; Baumeier, B. Electronic Excitations in Complex Molecular Environments: Many-Body Green's Functions Theory in VOTCA-XTP. *J. Chem. Theory Comput.* **2018**, *14*, 6253–6268.

(19) Tirimbò, G.; Sundaram, V.; Ca¸ylak, O.; Scharpach, W.; Sijen, J.; Junghans, C.; Brown, J.; Ruiz, F. Z.; Renaud, N.; Wehner, J.; Baumeier, B. Excited-state electronic structure of molecules using many-body Green's functions: Quasiparticles and electron-hole excitations with VOTCA-XTP. *J. Chem. Phys.* **2020**, *152*, 114103.

(20) Caylak, O.; Baumeier, B. Machine Learning of Quasiparticle Energies in Molecules and Clusters. *Journal of Chemical Theory and Computation* **2021**, *17*, 4891–4900, PMID: 34314186.

(21) Fux, S.; Jacob, C. R.; Neugebauer, J.; Visscher, L.; Reiher, M. Accurate frozen-density embedding potentials as a first step towards a subsystem description of covalent bonds. *J. Chem. Phys* **2010**, *132*, 164101.

(22) Gomes, A. S. P.; Jacob, C. R.; Visscher, L. Calculation of local excitations in large

39

systems by embedding wave-function theory in density-functional theory. *Phys. Chem. Chem. Phys.* **2008**, *10*, 5353–5362.

(23) Jacob, C. R.; Neugebauer, J.; Visscher, L. A flexible implementation of frozen-density embedding for use in multilevel simulations. *J. Comput. Chem.* **2008**, *29*, 1011–1018.

(24) Tölle, J.; Deilmann, T.; Rohlfing, M.; Neugebauer, J. Subsystem-Based GW/Bethe–Salpeter Equation. *J.Chem. Theory Comput.* **2021**, *17*, 2186–2199.

(25) Tölle, J.; Deilmann, T.; Rohlfing, M.; Neugebauer, J. Correction to "Subsystem-Based GW/Bethe–Salpeter Equation". *J.Chem. Theory Comput.* **2023**, *19*, 2699–2702.

(26) Senatore, G.; Subbaswamy, K. Density dependence of the dielectric constant of rare-gas crystals. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1986**, *34*, 5754.

(27) Johnson, M.; Subbaswamy, K.; Senatore, G. Hyperpolarizabilities of alkali halide crystals using the local-density approximation. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1987**, *36*, 9202.

(28) Cortona, P. Self-consistently determined properties of solids without band-structure calculations. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1991**, *44*, 8454.

(29) Wesolowski, T. A.; Warshel, A. Frozen Density Functional Approach for ab Initio Calculations of Solvated Molecules. *J. Phys. Chem.* **1993**, *97*, 8050.

(30) Wesolowski, T. A.; Weber, J. Kohn-Sham equations with constrained electron density: an iterative evaluation of the ground-state electron density of interacting molecules. *Chem. Phys. Lett.* **1996**, *248*, 71.

(31) Manby, F. R.; Stella, M.; Goodpaster, J. D.; Miller, T. F. A Simple, Exact Density-Functional-Theory Embedding Scheme. *J. Chem. Theory Comput.* **2012**, *8*, 2564–2568.

(32) Lee, S. J. R.; Ding, F.; Manby, F. R.; Miller, T. F. Analytical gradients for projection-based wavefunction-in-DFT embedding. *J. Chem. Phys.* **2019**, *151*, 064112.

(33) Claudino, D.; Mayhall, N. J. Automatic Partition of Orbital Spaces Based on Singular Value Decomposition in the Context of Embedding Theories. *J. Chem. Theory Comput.* **2019**, *15*, 1053–1064.

(34) Hégely, B.; Nagy, P. R.; Ferenczy, G. G.; Kállay, M. Exact density functional and wave function embedding schemes based on orbital localization. *J. Chem. Phys.* **2016**, *145*, 064107.

(35) Barnes, T. A.; Goodpaster, J. D.; Manby, F. R.; Miller, T. F. Accurate basis set truncation for wavefunction embedding. *J. Chem. Phys.* **2013**, *139*, 024103.

(36) Bennie, S. J.; Stella, M.; Miller, T. F.; Manby, F. R. Accelerating wavefunction in density-functional-theory embedding by truncating the active basis set. *J. Chem. Phys.* **2015**, *143*, 024105.

(37) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133–A1138.

(38) Kohn, W. Nobel Lecture: Electronic structure of matter—wave functions and density functionals. *Rev. Mod. Phys.* **1999**, *71*, 1253–1266.

(39) 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.

(40) Hedin, L.; Lundqvist, S. In *Effects of Electron-Electron and Electron-Phonon Interactions on the One-Electron States of Solids*; Seitz, F., Turnbull, D., Ehrenreich, H., Eds.; Solid State Physics; Academic Press, 1970; Vol. 23; pp 1–181.

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

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

(43) Golze, D.; Dvorak, M.; Rinke, P. The GW Compendium: A Practical Guide to Theoretical Photoemission Spectroscopy. *Frontiers in Chem.* **2019**, *7*, 377.

(44) Stone, A. *The Theory of Intermolecular Forces*, 2nd ed.; Oxford University Press: Oxford, 2013; p 352.

(45) Neese, F. The ORCA program system. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012**, *2*, 73–78.

(46) Weigend, F.; Köhn, A.; Hättig, C. Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. *J. Chem. Phys.* **2002**, *116*, 3175–3183.

(47) Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. RI-MP2: Optimized auxiliary basis sets and demonstration of efficiency. *Chem. Phys. Lett.* **1998**, *294*, 143–152.

(48) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.

(49) Pipek, J.; Mezey, P. G. A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. *J. Chem. Phys.* **1989**, *90*, 4916–4926.

(50) Lehtola, S.; Jónsson, H. Unitary Optimization of Localized Molecular Orbitals. *J.Chem. Theory Comput.* **2013**, *9*, 5365–5372.

(51) Rohlfing, M.; Krüger, P.; Pollmann, J. Efficient Scheme for GW Quasiparticle Band-Structure Calculations with Applications to Bulk Si and to the Si(001)-(2x1) Surface. *Physical Review B* **1995**, *52*, 1905–1917.

(52) Baral, S.; Baumeier, B.; Lyman, E. Predicting Spectral Properties of Polarity Sensitive Dyes with QM/MM Simulation. *Biophys. J.* **2018**, *114*, 272a.

(53) Emery, L. C.; Sheldon, J. M.; Edwards, W.; McHale, J. L. Intermolecular ground and excited state potential energy surfaces and charge-transfer spectrum of benzene: TCNE

42

electron donor—acceptor complex. *Spectrochimica Acta Part A: Molecular Spectroscopy* **1992**, *48*, 715–724.

(54) Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. Packmol: A package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* **2009**, *30*, 2157–2164.

(55) Martínez, J. M.; Martínez, L. Packing optimization for automated generation of complex system's initial configurations for molecular dynamics and docking. *J. Comput. Chem.* **2003**, *24*, 819–825.