Long-Range Embedding of Molecular Ions and Excitations in a Polarizable Molecular Environment

Carl Poelking[*,†] and Denis Andrienko[*,†]

†Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany

‡Heidelberg Graduate School of Fundamental Physics, INF 226, 69120 Heidelberg, Germany

E-mail: carl.poelking@mpip-mainz.mpg.de; denis.andrienko@mpip-mainz.mpg.de

Abstract

We present a method for evaluating electrostatic and polarization energies of a localized charge, charge transfer state, or exciton embedded in a neutral molecular environment. The approach extends the Ewald summation technique to polarization effects, rigorously accounts for the long-range nature of the charge-quadrupole interactions, and addresses aperiodic embedding of the charged molecular cluster and its polarization cloud in a periodic environment. We illustrate the method by evaluating the density of states and ionization energies in thin films and heterostructures of organic semiconductors. By accounting for long-range mesoscale fields, we obtain the ionization energies in both crystalline and mesoscopically amorphous systems with high accuracy.

1

1 Introduction

Knowledge of the energy landscape for charge and energy transport is key to the optimization of organic optoelectronic devices,[1] understanding of photosynthesis in biological systems,[2] molecular catalysis,[3] and chemical sensors.[4] It is, however, still a challenge in computational materials science to predict these landscapes and to establish their connection to device characteristics.[5–10] One of the major issues here is the limited system sizes used in today's simulations. They are frequently insufficient to sample the tail of the densities of states[8,11–13] or long-wavelength spatial correlations of the energy landscape.[14,15] The need to go beyond a nanometer length scale is also given by the fact that in ordered molecular systems the interaction of a charge with the molecular surrounding is long-range. For charge-quadrupole interactions, for instance, the corresponding sum is only conditionally convergent in 3D[16–20] and converges extraordinarily slow in 2D-periodic systems, such as thin organic films.

In illustration of this, Fig. 1a shows the electrostatic contribution to the energies of a neutral molecule, its cation and anion in a molecular crystal of the solar-cell donor compound D5M.[21] The dependence on the size of the molecular cluster, d_c, reflects the electrostatic energy convergence in 3D- and 2D-periodic systems. Fig. 1a may give the false impression that energy levels are converged for a cluster size of 8 nm, when, in fact, the interaction sum is only *conditionally* converged. It corresponds to spherical shell-by-shell growth of the cluster and will differ for other (cylindrical, cuboidal, etc.) cluster shapes. Fig. 1b shows how these energies change in a crystalline *thin film*. For *charges* embedded in a 20 nm-thin film (Fig. 1b), convergence is absolute but is not achieved even for cluster sizes larger than 100 nm. Finally, in Fig. 1c the electrostatic contribution is shown for a charge transfer (CT) state at a donor-acceptor (D5M-C_{60}) interface comprised of two thin films of D5M and the fullerene C_{60}. The electrostatic contribution to the energy of CT states converges faster, whereas the individual contributions of the CT-hole and CT-electron still exhibit the same slow thin-film convergence behavior.

As long-range crystalline and liquid-crystalline ordering characterizes many organic ma-
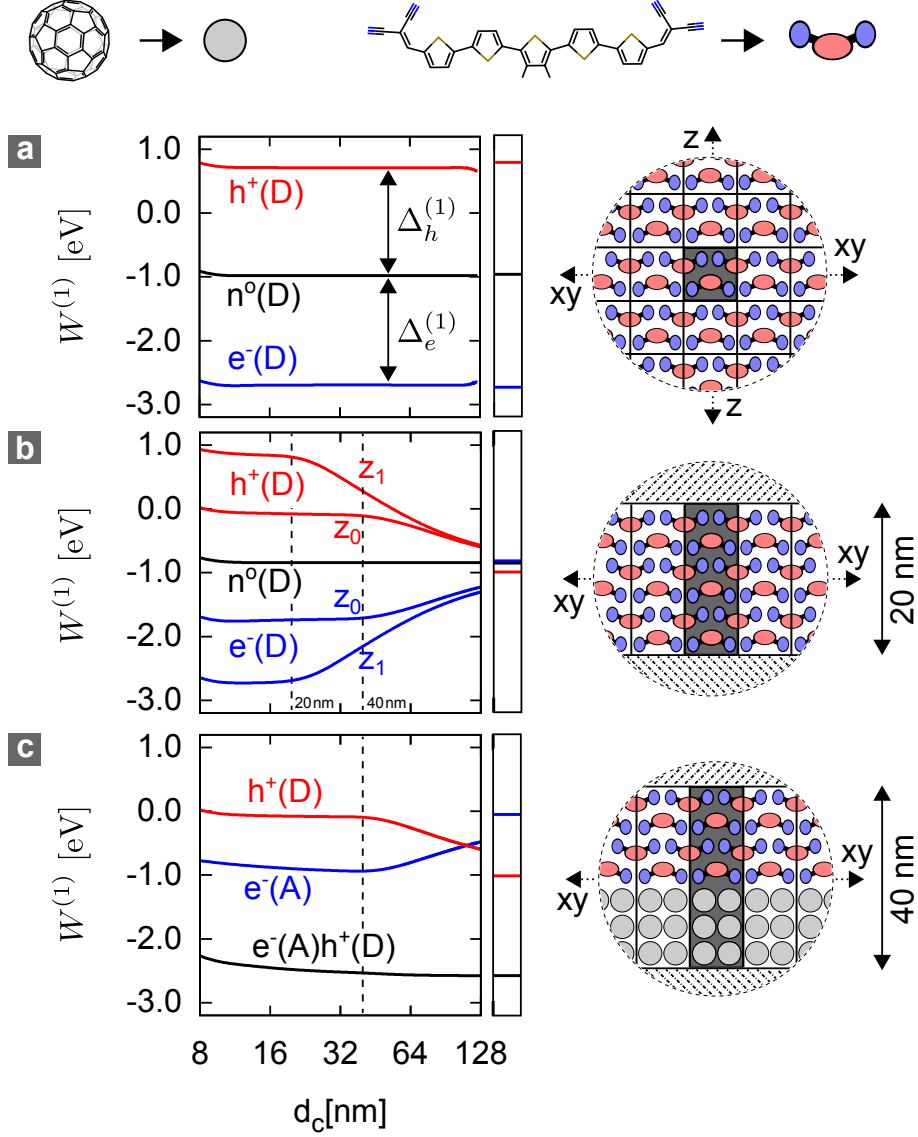
2

Figure 1: Variation of electrostatic energies $W^{(1)}$ with the cluster size d_c for a bulk (a) and thin-film (b,c) setup. Only static atomic multipoles up to quadrupoles are accounted for. In (a) and (b), black, red, and blue lines denote trends for neutral (n^o), hole (h^+) and electron (e^-) states, localized on a donor (D, here: D5M) or acceptor (A, here: C_{60}) molecule, respectively. In (b), z_0 and z_1 refer to a molecular ion positioned at the center and surface of the film, respectively. In (c), the black line denotes the convergence for an interfacial charge transfer state. Energies $W^{(1)}$ for the infinite systems are shown on the very right of each plot, where for the bulk setup a spherical shape factor has been used to remove the conditional convergence, which corresponds to the spherical shell-by-shell growth of the molecular cluster of size d_c. The schematics on the right-hand side indicate the system under study including cartoons of the two molecules, D5M and fullerene, with a repeat unit colored in black, and periodicity indicated by dashed arrows.

terials,[22,23] an account of long-range effects becomes indispensable in heterogeneous environments, in particular at interfaces. Such structures are at the heart of many functional devices, e.g., planar heterojunction solar cells, field effect transistors, and organic light-emitting diodes. Typical film thicknesses employed in these devices are on the order of tens of nanometers, whereas surface areas exceed square micrometers. Hence, the molecular arrangement in the out-of-plane dimension can be simulated to scale, whereas the in-plane dimensions have to be modeled effectively, through periodic boundary conditions.

Following up on applications to organic interfaces[16,17] and mixtures,[24] in this work we focus on the technical implementation, further verification, and complementary illustrations of the method designed to evaluate ionization energies (IEs) and electron affinities (EAs) of molecules embedded in such a *periodic* molecular environment. In what follows, we will hence focus on molecular ions embedded in a neutral environment; the method is, however, applicable to any type of embedded molecular excitation, including charge transfer and excited states.

Addressing both electrostatic and polarization contributions, the approach adapts and extends the Ewald summation technique[25,26] in three ways. First, it incorporates induction, which is not originally part of the formalism, but can be added.[27] Second, it addresses the long-range nature of the charge-quadrupole interaction. Third, it adapts the Ewald method to "broken" periodicities that result when embedding the charged molecular cluster and its polarization cloud in an otherwise periodic system. We illustrate that a long-range treatment is crucial for the description not only of crystalline systems, but also of energy correlations and level alignment in mesoscopically amorphous materials.

2   General formalism

In this section, we outline the implementation of our approach. To make simulations of large, atomistically-resolved systems computationally feasible, we use polarizable force fields based
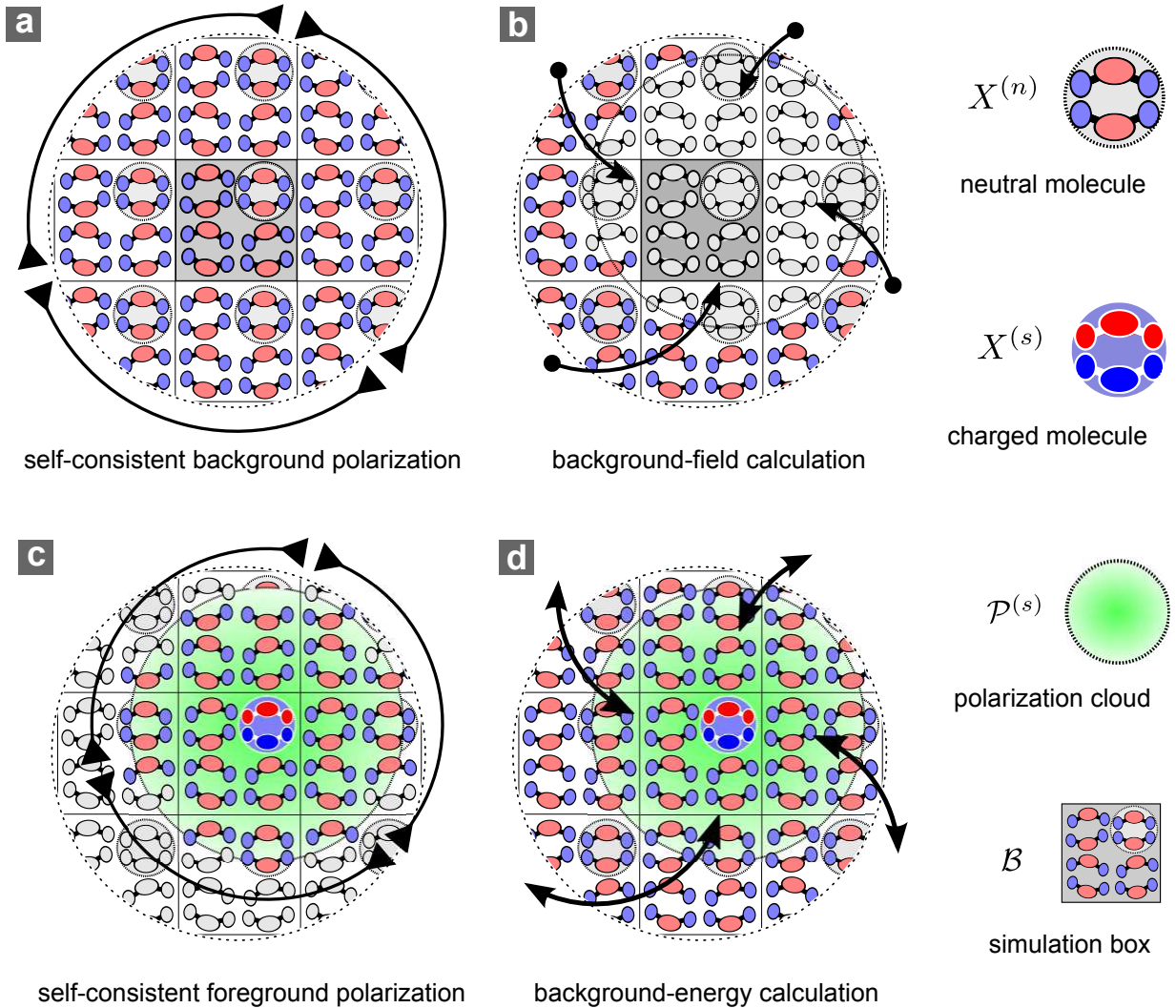
Figure 2: Computational workflow. (a) Computation of the polarization state of the periodic ground-state system. The simulation box $\mathcal{B}$, colored in gray, incorporates the molecular cluster in a neutral state, $X^{(n)}$. (b) Calculation of background fields acting on the polarization cloud (foreground) $P^{(s)}$ centered around the charged cluster $X^{(s)}$. (c) Self-consistent polarization within $\mathcal{P}^{(s)}$ and evaluation of the self-energy of $\mathcal{P}^{(s)}$. (d) Evaluation of the interaction energy between the foreground $\mathcal{P}^{(s)}$ and the polarized background $\tilde{\mathcal{B}}^*$.

on distributed atomic properties.[28–31] The basics of this concept are recapitulated in the next section. Readers familiar with the perturbative treatment of electrostatic interactions and polarization energies can move on directly to Section 2.2.

2.1 Polarization Energy and Work

In weakly interacting molecular assemblies, key corrections to energy levels of molecular ions result from the electrostatic and polarization interaction with the environment. Both interactions lead to energy contributions on the order of $1\,\text{eV}$ and can be treated perturbatively.[32] The electrostatic and polarization contributions correspond to the first- and second-order corrections $W_s^{(1)}$ and $W_s^{(2)}$, respectively, where the subscript s denotes the state (hole, electron, neutral) of the molecule. The energy correction Δ_s to IEs ($s = h$) and EAs ($s = e$) due to the environment then reads

$$\Delta_s = \Delta_s^{(1)} + \Delta_s^{(2)}, \tag{1}$$

where $\Delta_s^{(i)} = W_s^{(i)} - W_n^{(i)}$; n references the neutral ground state. The IE of a molecule, for example, follows as $\text{IE} = \text{IE}_0 + \Delta_h$: here, IE_0 denotes the gas-phase ionization energy, to be calculated on a quantum-mechanical level. For electron, excitonic or charge-transfer states, analogous expressions hold.

The perturbative correction to the molecular site energy, $W_s = W_s^{(1)} + W_s^{(2)}$, is calculated in a classical expansion of the molecular field and field response in terms of distributed multipoles[28] and polarizabilities,[33] respectively, positioned at (typically) atomic expansion sites. W_s then follows from a variational principle, which replaces Poisson's equation $\nabla(\varepsilon \nabla \phi) \sim \rho$ in this microscopic, particle-based picture:

$$\frac{\delta W_s}{\delta \Delta Q_{lm}^{a(s)}} = 0. \tag{2}$$

Its self-consistent solution consists of the set of multipolar (here: dipolar) moments $\Delta Q_{lm}^{a(s)}$ induced at the (atomic) expansion sites, a, with local polarizabilities α_{lm}^a, in response to the permanent multipolar moments $Q_{lm}^{a(s)}$ that approximate the molecular, unperturbed charge densities. The set $\Delta Q_{lm}^{a(s)}$ constitutes the polarization state of the system and, together with

$Q_{lm}^{a(s)}$, determines $W_s^{(1)}$ and $W_s^{(2)}$, and finally Δ_s. We note that $W_s^{(2)}$, different from $\Delta_s^{(2)}$, only takes negative values, as is characteristic of a second-order perturbative term.

The self-interaction energy W_s of the molecular system is normally decomposed into external and internal contributions, the latter reflecting the positive polarization work[32]

$$W_s = W_{\text{ext}} \left[\mathcal{D}^{(s)}; \mathcal{D}^{(s)} \right] + W_{\text{int}} \left[\mathcal{D}^{(s)} \right], \tag{3}$$

$$W_{\text{ext}} \left[\mathcal{D}^{(s)}; \mathcal{D}^{(s)} \right] = \frac{1}{2} \sum_{M^{(s)}} \sum_{M'^{(s)} \neq M^{(s)}} \left(Q_t^{a(s)} + \Delta Q_t^{a(s)} \right) T_{tu}^{aa'} \left(Q_u^{a'(s)} + \Delta Q_u^{a'(s)} \right), \tag{4}$$

$$W_{\text{int}} \left[\mathcal{D}^{(s)} \right] = \frac{1}{2} \sum_M \Delta Q_t^{a(s)} \left(\alpha^{-1} \right)_{tu}^{ab(s)} \Delta Q_u^{b(s)}. \tag{5}$$

Here, $\mathcal{D}^{(s)}$ is a discrete, multipolar charge density – it denotes the entire set of static and induced multipoles in the system; the notation $W_{\text{ext}} \left[\mathcal{D}^{(s)}; \mathcal{D}^{(s)} \right]$ emphasizes that the summation is performed over all pairs of molecules.

Following Stone's notation,[32] Q_t^a and $Q_u^{a'}$ are multipole moments in spherical-tensor representation, with angular and magnetic quantum numbers contracted into a single index. $T_{tu}^{aa'}$ are tensors that mediate the interaction between multipole moments Q_t^a and $Q_u^{a'}$ of atoms a and a'.[34] For both atomic indices ($a, b \in M$ and $a' \in M'$) as well as tensorial components (t and u), Einstein sum conventions are in place.

For computational efficiency, we truncate the expansion of the molecular charge density (distributed multipole analysis[28]) after rank $l = 2$, such that atomic quadrupoles are still accounted for. Also, we employ distributed polarizabilities in a local-dipole approximation ($\alpha_{tu}^{ab} \rightarrow \alpha_t^a \neq 0$ if $t \in \{1x, 1y, 1z\}$) as developed by Thole,[30,35] hence disregarding charge-flow effects.[29] In organic solids with their spatially rapidly fluctuating fields, charge flow is expected to play only a minor role. A local-dipole scheme is therefore sufficient to capture polarization effects. In order to avoid an unphysical overpolarization to which atomic-dipole schemes are susceptible,[30] the Thole model requires damping of induced-induced interactions at short separations. The interaction tensors $T_{tu}^{aa'}$ in Eq. 4 are hence modified such that terms with a distance scaling of $R^{-\nu}$ are multiplied by a damping function $\Lambda_{2\nu+1}$ listed in Section S3

of the Supporting Information. Furthermore, due to the larger polarizabilities of *conjugated* molecules in comparison to biological compounds, the set of Thole polarizabilities[35] is scaled iteratively in order to match the volume of the polarizability ellipsoid ($\sim 1/\Pi_{i=1}^{3}\sqrt{\tilde{\alpha}_i}$, where $\tilde{\alpha}_i$ is the i-th eigenvalue of the molecular polarizability tensor) calculated here using density functional theory.

2.2 Aperiodic-Periodic Decomposition

In the following, we detail how Eq. 2 is solved for an *aperiodic* (i.e., not periodically repeated) charged molecule embedded in a *periodic*, neutral, polarizable environment and how the associated perturbative contributions $\Delta_s^{(1)}$ and $\Delta_s^{(2)}$ are evaluated with an infinite interaction range applied to all particles.

Since a fully self-consistent solution of Eq. 2 for the entire system (incorporating all periodic images as well as a single charged molecule) is difficult to achieve even in principle, we first introduce a cutoff for the polarization of the environment by the excess charge. This cutoff controls up to where the introduction of this charge modifies the polarization state of the environment. The assumption of such a finite "polarization cloud" is readily justified: the charge – induced-dipole energy contribution has a $1/r^4$ decay and is hence absolutely convergent. The polarization cloud then has to be chosen only large enough to prevent discontinuities in the fields across its boundary.

To deal with the broken periodicity, we spatially partition the multipolar charge density $\mathcal{D}^{(s)}$ onto several subsets. Once again, the superscript s denotes the state of the embedded molecule; e.g., $s = n$ would correspond to the neutral state. Now let us denote by $\mathcal{B}$ the set of molecules in the simulation box, with all molecules in their *neutral* charge configuration. The entire ground-state system $\mathcal{B}^*$ then consists of $\mathcal{B}$ and its periodic replicas. Charging of the embedded molecule, $X^{(s)}$, modifies this charge density by polarizing the surrounding molecules. We correspondingly denote the set of molecules in the polarization cloud by $\mathcal{P}^{(s)}$. We will refer to the rest of the system, $\tilde{\mathcal{B}}^* = \mathcal{B}^* \setminus \mathcal{P}^{(n)}$, as the background. The system

partitioning is illustrated in Fig. 2.

Based on Eq. 3, the energy correction W_s then reads

$$W_s = W_{\text{ext}}[\mathcal{P}^{(s)}; \mathcal{P}^{(s)}] + W_{\text{ext}}[\mathcal{P}^{(s)}; \tilde{\mathcal{B}}^*] + W_{\text{ext}}[\tilde{\mathcal{B}}^*; \tilde{\mathcal{B}}^*] + W_{\text{int}}[\mathcal{P}^{(s)}] + W_{\text{int}}[\tilde{\mathcal{B}}^*]. \tag{6}$$

Here, W_{ext} denotes the intermolecular field interaction energy, and W_{int} the intramolecular induction work, as defined in Eqs. 4,5. The polarization cloud (Fig. 2c) is chosen large enough to screen the excitation from $\tilde{\mathcal{B}}^*$, whose polarization state is thus assumed unaffected. Therefore no state index s is used for the contributions $W_{\text{ext}}[\tilde{\mathcal{B}}^*; \tilde{\mathcal{B}}^*]$ and $W_{\text{int}}[\tilde{\mathcal{B}}^*]$, which will in fact cancel when taking the difference $\Delta_s^{(i)} \equiv W_s^{(i)} - W_n^{(i)}$ and hence need not be calculated. The surviving terms from Eq. 6 read as follows:

$$W_{\text{ext}}[\mathcal{P}^{(s)}; \mathcal{P}^{(s)}] = \frac{1}{2} \sum_P \sum_{P' \neq P} \left(Q_t^{p(s)} + \Delta Q_t^{p(s)} \right) T_{tu}^{pp'} \left(Q_u^{p'(s)} + \Delta Q_u^{p'(s)} \right), \tag{7}$$

$$W_{\text{ext}}[\mathcal{P}^{(s)}; \tilde{\mathcal{B}}^*] = \sum_P \sum_{\tilde{B}^*} \left(Q_t^{p(s)} + \Delta Q_t^{p(s)} \right) T_{tu}^{pb} \left(Q_u^{b(n)} + \Delta Q_u^{b(n)} \right), \tag{8}$$

$$W_{\text{int}}[\mathcal{P}^{(s)}] = \frac{1}{2} \sum_P \Delta Q_t^{p(s)} \left(\alpha^{-1} \right)_t^{a(s)} \Delta Q_t^{p(s)}. \tag{9}$$

Here, $p(s)$ denotes an atom of a molecule $P^{(s)}$ in a polarization cloud $\mathcal{P}^{(s)}$, s is the state of the molecule (neutral, anion, cation), and b enumerates atoms in the polarized background $\tilde{B}^*$. In these expressions, induced moments have to be calculated self-consistently on the basis of Eq. 2, yielding the linear system of equations[27,32]

$$\Delta Q_t^{p(s)} = - \sum_{\tilde{B}^*} \alpha_t^{p(s)} T_{tu}^{pb} \left(Q_u^{b(n)} + \Delta Q_u^{b(n)} \right) - \sum_P \alpha_t^{p(s)} T_{tu}^{pp'} \Delta Q_u^{p'(s)}. \tag{10}$$

Note that in Eqs. 7 and 9, $W_{\text{ext}}[\mathcal{P}^{(s)}; \mathcal{P}^{(s)}]$ and $W_{\text{int}}[\mathcal{P}^{(s)}]$ only count interactions within $\mathcal{P}^{(s)}$: direct evaluation is therefore possible at reasonable computational expense. $W_{\text{ext}}[\mathcal{P}^{(s)}; \tilde{\mathcal{B}}^*]$ (Eq. 8), by contrast, comprises the interaction of $\mathcal{P}^{(s)}$ with the infinite, but *semiperiodic* set $\tilde{\mathcal{B}}^*$. In order to use the Ewald summation technique, designed for periodic systems, we add

the foreground density $\mathcal{P}^{(n)}$ in its neutral charge and polarization state to $\tilde{\mathcal{B}}^*$, thus allowing for the transformation into reciprocal space, and subsequently correct for this addition in real space. taking spatial derivatives with regard to the coordinates of the source and target atomic positions. $W_{\text{ext}}[\mathcal{P}^{(s)}; \tilde{\mathcal{B}}^*]$ is then obtained as the sum over five contributions (the derivation is sketched in Sections S8 and S9 of the Supporting Information),

$$W_{\text{ext}}[\mathcal{P}^{(s)}; \tilde{\mathcal{B}}^*] = W_k[\mathcal{P}^{(s)}; \mathcal{B}] + W_r[\mathcal{P}^{(s)}; \tilde{\mathcal{B}}^*] - W_{\text{si}}[\mathcal{P}^{(s)}; \mathcal{P}^{(n)}] - W_c[\mathcal{P}^{(s)}; \mathcal{P}^{(n)}] + W_*[\mathcal{P}^{(s)}; \mathcal{B}],$$

(11)

The first four terms on the right-hand side are:

$$W_k[\mathcal{P}^{(s)}; \mathcal{B}] = \frac{1}{4\pi\varepsilon_0} \frac{4\pi}{V} \sum_{\vec{k} \neq 0}^{\infty} S\left(\vec{k}; [\mathcal{P}^{(s)}]\right) S^*\left(\vec{k}; [\mathcal{B}]\right) A(k),$$

(12)

$$W_r[\mathcal{P}^{(s)}; \tilde{\mathcal{B}}^*] = \frac{1}{4\pi\varepsilon_0} \sum_{\tilde{B}^*} \sum_{\mathcal{P}^{(s)}} \sum_{\nu=0}^{4} T_\nu^{pb} \Lambda_{2\nu+1}^{pb} B_\nu(R_L^{pb}),$$

(13)

$$W_{\text{si}}[\mathcal{P}^{(s)}; \mathcal{P}^{(n)}] = \frac{1}{4\pi\varepsilon_0} \sum_{\mathcal{P}} \left(\frac{2\beta}{\sqrt{\pi}} q^{p(s)} q^{p(n)} + \frac{4\beta^3}{3\sqrt{\pi}} \vec{\mu}^{p(s)} \cdot \vec{\mu}^{p(n)} + \frac{16\beta^5}{5\sqrt{\pi}} \tilde{\theta}^{p(s)} : \tilde{\theta}^{p(n)} \right),$$

(14)

$$W_c[\mathcal{P}^{(s)}; \mathcal{P}^{(n)}] = \frac{1}{4\pi\varepsilon_0} \sum_{P(s)} \sum_{P(n) \neq P(s)} \sum_{\nu=0}^{4} T_\nu^{pp'} C_\nu(R_0^{pp'}).$$

(15)

The first three contributions in Eq. 11 (Eqs. 12-14) are the standard Ewald terms for the reciprocal- and real-space interaction, and self-interaction correction, respectively. Their interaction kernels take into account higher-order permanent and induced moments[36,37] and, in the case of $W_r[\mathcal{P}^{(s)}; \tilde{\mathcal{B}}^*]$, incorporate short-range damping functions $\Lambda_{2\nu+1}$ as discussed above.

For the real- and reciprocal-space terms, the anisotropic interaction kernels S and T_ν (the former is the structure factor of the multipolar density) capture the orientation dependence of the interaction between two sets of atomic multipoles in the respective space; they are listed in Section S1 of the Supporting Information. The associated isotropic interaction kernels $A(k)$ and $B_\nu(R)$ (see Section S2 of the Supproting Information) yield the

10

distance dependence. Finally, the real-space interaction is damped by the damping function $\Lambda_{2\nu+1}$ prescribed by $B_\nu(R)$, which scales as $R^{-(2\nu+1)}$ for sufficiently small distances (see Section S3 of the Supporting Information for details). In the above equations, the set of atom-centered multipoles incorporates the atomic charge q, dipole $\vec{\mu}$ and quadrupole $\tilde{\theta}$, now in their Cartesian representation. Note that the quadrupole is here defined as $\theta_{\alpha\beta} = \sum_a q^a \left[\frac{1}{2} r_\alpha^a r_\beta^a - \frac{1}{6} \delta_{\alpha\beta} (r^a)^2 \right]$, which differs from the conventional definition by a factor $\frac{1}{3}$. Also note that the sum $\sum_{\tilde{\mathcal{B}}^*}$ is in practice implemented as a double sum over image-box vectors $\vec{L}$ and the periodic density $\mathcal{B}^*$, with molecules participating in the polarization cloud $\mathcal{P}^{(s)}$ and $\mathcal{P}^{(n)}$ being excluded.

The fourth term in Eq. 11, $W_c[\mathcal{P}^{(s)}; \mathcal{P}^{(n)}]$, denotes an aperiodic subtraction which corrects for the overlap between $\mathcal{P}^{(s)}$ and $\mathcal{P}^{(n)}$ (see Section S9 of the Supporting Information for details). Here, the anisotropic kernel is identical to T_ν from Eq. 13. The isotropic kernel C_ν, however, differs from B_ν (see Section S2 of the Supporting Information) in that it involves derivatives of the long-ranged $\mathrm{erf}(\beta r)/r$ rather than the short-ranged $\mathrm{erfc}(\beta r)/r$ part of the interaction. The $P^{(n)} = P^{(s)}$ terms have already been accounted for in the self-interaction term, Eq. 14.

The fifth term in Eq. 11, $W_*[\mathcal{P}^{(s)}; \mathcal{B}]$, is a shape ($k = 0$) contribution that tackles the conditionality of the interaction sum. An analogous conditionality arises in overall neutral systems with a net dipole moment of the periodically repeated charge density $\mathcal{B}$,[20] as is often encountered in molecular systems. As a result, the convergence of the interaction sum depends on the (macroscopic) summation shape. Here, we treat molecular solids, where a second conditionality arises from the net charge that resides in $\mathcal{P}^{(s)}$ and interacts with a net-quadrupolar background. Shape terms for the case of a macroscopic cube and slab summation shape are derived as (see Section S8 of the Supporting Information for details)

$$W_*^{\mathrm{cube}}[\mathcal{P}^{(s)}; \mathcal{B}] = -\frac{1}{4\pi\varepsilon_0} \frac{4\pi}{3V} \left(Q_0^{\mathcal{P}^{(s)}} \mathrm{Tr}[\tilde{\Theta}^{\mathcal{B}(n)}] + Q_0^{\mathcal{B}(n)} \mathrm{Tr}[\tilde{\Theta}^{\mathcal{P}(s)}] - \vec{M}^{\mathcal{P}(s)} \cdot \vec{M}^{\mathcal{B}(n)} \right), \quad (16)$$

$$W_*^{\mathrm{slab}}[\mathcal{P}^{(s)}; \mathcal{B}] = -\frac{1}{4\pi\varepsilon_0} \frac{4\pi}{V} \left(Q_0^{\mathcal{P}(s)} \Theta_{zz}^{\mathcal{B}(n)} + Q_0^{\mathcal{B}(n)} \Theta_{zz}^{\mathcal{P}(s)} - M_z^{\mathcal{P}(s)} M_z^{\mathcal{B}(n)} \right), \quad (17)$$

11

where Eq. 16 implies summation over cubic/spherical, and Eq. 17 over infinitely thin slab-shaped shells. The net multipolar moments of the densities $\mathcal{P}^{(s)}$ and $\mathcal{B}$ that feature in these expressions are defined as

$$Q_0^{\mathcal{P}(s)} = \sum_P q^{p(s)}, \tag{18}$$

$$\vec{M}^{\mathcal{P}(s)} = \sum_P \left(q^{p(s)} \vec{r}_p + \vec{\mu}^{p(s)} \right), \tag{19}$$

$$\tilde{\Theta}^{\mathcal{P}(s)} = \sum_P \left(\frac{1}{2} q^{p(s)} \vec{r}_p \otimes \vec{r}_p + \vec{\mu}^{p(s)} \otimes \vec{r}_p + \tilde{\theta}^{p(s)} \right).$$

Analogous expressions are used for $Q_0^{\mathcal{B}}$, $\vec{M}^{\mathcal{B}}$ and $\tilde{\Theta}^{\mathcal{B}}$.

Equation 11 is the key result of the method: with the cube/sphere shape correction, Eq. 16, it gives energies of charged clusters in a 3D-periodic lattice with all three periodic directions treated identically in the summation. In order to mimic a 2D-infinite slab (e.g. for calculating IE and EA of thin films), the slab shape correction, Eq. 17, shall be used.

2.3 Computational procedure

Combining our results for energy, field and polarization calculations, Fig. 2 summarizes the procedure to solve Eq. 2 and compute W_s for a molecular system in a state s. First (Fig. 2a), the polarization state of the neutral system is computed according to Eq. 10 and Eq. S17 in the Supporting Information. This step usually needs to be carried out only once for each coordinate configuration. Second (Fig. 2b), fields generated by the semiperiodic background across the polarization cloud are calculated, taking into account both permanent and induced moments. Third (Fig. 2c), the polarization cloud is polarized self-consistently in the fields of the semiperiodic background. Fourth (Fig. 2d), the interaction-energy between the semiperiodic background and the polarization cloud, as well as within the polarization cloud, are evaluated using Eqs. 7, 9 and 11.

The computational cost of the above procedure is mostly due to the self-consistent eval-

uation of induced dipoles of the polarization cloud (step 3). The cost of this step scales as r_{pc}^6 with the radius r_{pc} of the cloud. Typically, $r_{pc} = 4\,\text{nm}$ proves to be sufficient to converge level profiles, except for a homogeneous dielectric stabilization of the material beyond r_{pc}.

Finally, the long-range stabilization is recovered in an *ad-hoc* fashion:[38] To this end, we consider the polarization energy of a charge localized in a non-polarizable spherical cavity with radius r_{pc}, embedded in a dielectric of dielectric constant ε_1 in the half-space $z > 0$, in the vicinity of an interface with a second dielectric layer of dielectric constant ε_2 in $z < 0$ (as frequently encountered in organic electronic devices):

$$\Delta_s^{(\varepsilon)}(z > r_{pc}) = -\frac{1}{8\pi\varepsilon_0}\frac{q^2}{r_{pc}}\frac{\varepsilon_1 - 1}{\varepsilon_1}\left[1 + \frac{r_{pc}}{2z}\frac{\varepsilon_2 - \varepsilon_1}{(\varepsilon_1 - 1)(\varepsilon_1 + \varepsilon_2)}\right]. \tag{20}$$

This expression holds for excitations with a net-charge q as leading moment. Higher order moments, as they would apply to net-neutral excitations, in particular charge transfer states, could be readily treated on the same grounds.[38] In practice, the contribution from higher-order moments is, however, negligible – different from charges, where, e.g., $\Delta^{(\varepsilon)}$ evaluates to -0.11 eV for a hole or electron embedded in an environment of $\varepsilon_1 = \varepsilon_2 = 3$ and $r_{pc} = 4\,\text{nm}$.

3   Validation and applications

In this section, we first validate electrostatic energies obtained with our method against the existing 2D-Ewald approach by initially excluding all polarization contributions. We then briefly discuss published applications of the method, with benchmarks based on experimentally measured quantities: IEs of thin crystalline films,[16] energies of charge transfer states at donor-acceptor interfaces,[17] and density of states of mixtures of molecular crystals.[24] Finally, we apply our method to amorphous mesophases, as encountered in organic solar cells or light-emitting diodes, and show that residual ordering in a small simulation box combined with the cut-off based approaches can lead to false predictions of molecular energies in amorphous thin films.

3.1 2D Ewald vs. 3D-periodic plus the slab shape-term

For the case of a molecular film sandwiched between two vacuum layers, periodically repeated in all three dimensions, the thin film shape correction is inversely proportional to the volume of the simulation cell. Hence, by increasing its out-of-plane dimension, while keeping the slab thickness fixed, one can eventually recover the exact result for a 2D-periodic system even without applying a shape term[39,40] – at the cost, however, of a significantly denser k-vector spacing in reciprocal space. Another option is to use a 2D-Ewald sum[41,42] with the underlying formulae listed in Section S6 of the Supporting Information. The comparison of the 2D-Ewald versus 3D-periodic with the slab shape-term descriptions is presented in Section S7 of the Supporting Information. The comparison serves as a validation of the approach, at least for neutral non-polarizable systems with partial charges (the 2D-Ewald formalism has so far been formulated only for electrostatic sums with partial charges, thus excluding atomic polarizabilities and higher-rank multipoles).

The proposed 3D-periodic description has, however, several benefits as compared to the 2D-Ewald method. First, it is more efficient, since the reciprocal-space sum conveniently factorizes with respect to the two multipolar densities $\mathcal{P}^{(s)}$ and $\mathcal{B}$. Second, it is more flexible: it enables the simulation of both thin-film and bulk conditions within the same framework and, on the implementation side, is capable of treating higher-rank multipoles and polarization effects.

3.2 Ionization energies of crystalline films

The proposed technique is ideally suited for the evaluation of ionization energies (IEs) of thin organic films, routinely performed experimentally using ultraviolet photoelectron spectroscopy (UPS). To benchmark the accuracy of the method, the ionization energies of five different materials (see Fig. 3a) in face-on, long-edge-on and short-edge-on orientations in thin films were calculated as described in Ref.,[16] by adding the thin-film shape contribution to the energy of the periodic 3D-system. The resulting IEs agree remarkably well with the
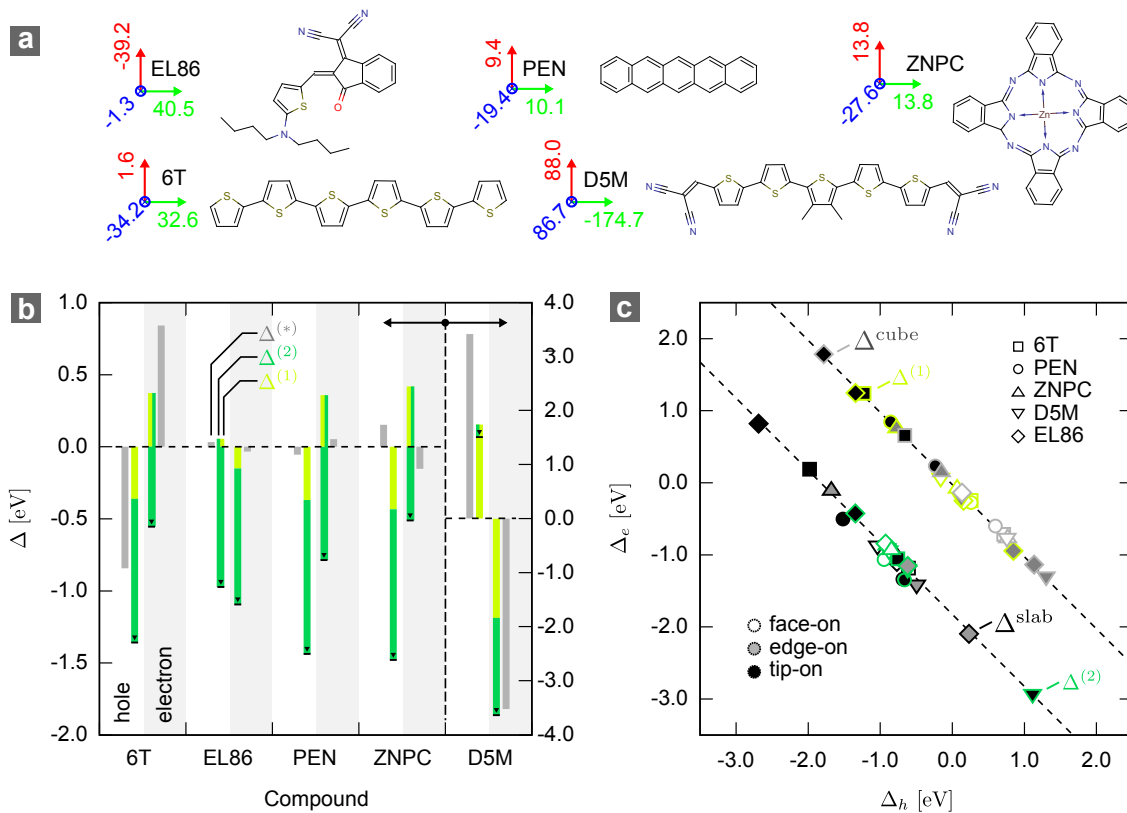
Figure 3: (a) Chemical structures of pentacene (PEN), sexithiophene (6T), zinc phthalocyanine (ZNPC), the merocyanine dye EL86,[43] and the acceptor-substituted oligothiophene D5M,[21] shown together with the eigenvectors of the quadrupole tensor (in atomic units). (b) Electrostatic ($\Delta^{(1)}$, light green), polarization ($\Delta^{(2)}$, dark green) and cube shape (Δ^{cube}, gray) contributions to electron and hole levels computed from experimental unit cells. The small black bars indicate the total solid-state contribution, Δ^{cube}. Note the different energy scale for D5M. (c) Correlation of electrostatic ($\Delta^{(1)}$) and polarization ($\Delta^{(2)}$) contributions to electron (Δ_e) and hole (Δ_h) site energies for thin-film levels calculated for differently oriented unit cells under application of the respective shape contribution. The symbol shape indicates the material (6T, ..., EL86), fill style the configuration (face-on, edge-on, tip-on).

experimentally measured energy levels:[16] in fact, the accuracy of the proposed method allows to deduce molecular orientation from a single IE measurement.

To illustrate the importance of different contributions to the ionization energy, Fig. 3b shows the electrostatic and polarization contributions $\Delta^{(1)}$ and $\Delta^{(2)}$ for both electrons and holes, as well as the shape contribution Δ^{cube} computed from Eq. 16. Note that the latter includes contributions from both electrostatics and polarization. First, it can be seen that the first-order correction $\Delta^{(1)}$ varies significantly among compounds, as does Δ^{cube}. $\Delta^{(2)}$ is reasonably constant across different materials, ranging between $-0.9\,\text{eV}$ and $-1.0\,\text{eV}$.

Fig. 3b seems to indicate that there are specific packing modes and molecular layouts, which together energetically favor either holes or electrons. In fact, apart from Δ^{cube}, these exact same results could have also been extracted from a cutoff-based description, which – as seen in Fig. 1a – tends to correspond to the cubic/spherical limit implied by Eq. 16. Such an approach has been used in the past and indeed led to the conclusion that some packing modes (e.g., herringbone vs. brickwork) will lead to lower bulk hole energies than others.[6]

A bulk description is, however, not appropriate when simulating devices. Instead, a thin-film description is needed, where IEs depend also on molecular orientation. For illustration, Fig. 3c correlates electron and hole contributions $\Delta^{(1)}$, $\Delta^{(2)}$, Δ^{cube}, as well as the solid state contribution Δ^{slab} obtained from the same unit cells as simulated in Fig. 3b, but with a slab shape term (Eq. 17) applied along the three unit-cell vectors. This procedure mimics different orientations (face-on, edge-on, tip-on) in a thin film, indicated by the fill style of the symbols, next to the symbol shape, which distinguishes between the five compounds. As becomes clear from a comparison of $\Delta^{(1)}$ (light-green symbols), general conclusions that link packing modes to electrostatic and polarization contributions are now impossible to formulate without at the same accounting for molecular orientation. Specifically, orientations with $Q_{20} > 0$ tend to produce a larger electrostatic stabilization of holes, as predicted by Eq. 17: For D5M, $Q_{20} > 0$ is associated with a face-on, for PEN, 6T and ZNPC with a tip-on orientation. Generally, the effect of orientation is sufficiently strong to produce both

16

negative and positive $\Delta^{(1)}$'s for the same carrier type.

3.3 Charge transfer states at donor-acceptor interfaces

Another practical application of our embeddig approach is the evaluation of energies of charge transfer (CT) states at the donor-acceptor (DA) interfaces of organic solar cells.[17] Here, the long-range charge quadrupole interactions result in an additional contribution, which either stabilizes or destabilizes the charge transfer state, depending on the orientation and ordering of molecular quadrupoles at the DA interface. In fact, this additional electrostatic potential can lead to a practically barrierless splitting of charge transfer states, which otherwise are strongly bound by the Coulomb attraction. Understanding the link between the photovoltaic gap and charge splitting and detrapping is of course crucial for the design of efficient photovoltaic cells.[17]

3.4 Density of states of mixtures

The proposed method is also suitable for evaluating the density of states of organic semiconductors. In particular, one can show that the long-range electrostatic effects can be exploited to tune the density of states of an organic semiconductor *continuously*, despite the fact that charges are spatially localized.[24] Such "bandstructure engineering", as confirmed by UPS spectroscopy, opens up new opportunities for the design of organic solar cells: the open-circuit voltage of organic solar cells can be continuously tuned by blending different absorber materials.[24]

3.5 Amorphous Systems

In this section, we will consider amorphous semiconductors used in organic light-emitting diodes, where disordered materials are preferred due to their processibility and fine miscibility in host-guest systems. Studying amorphous semiconductors in the context of long-range
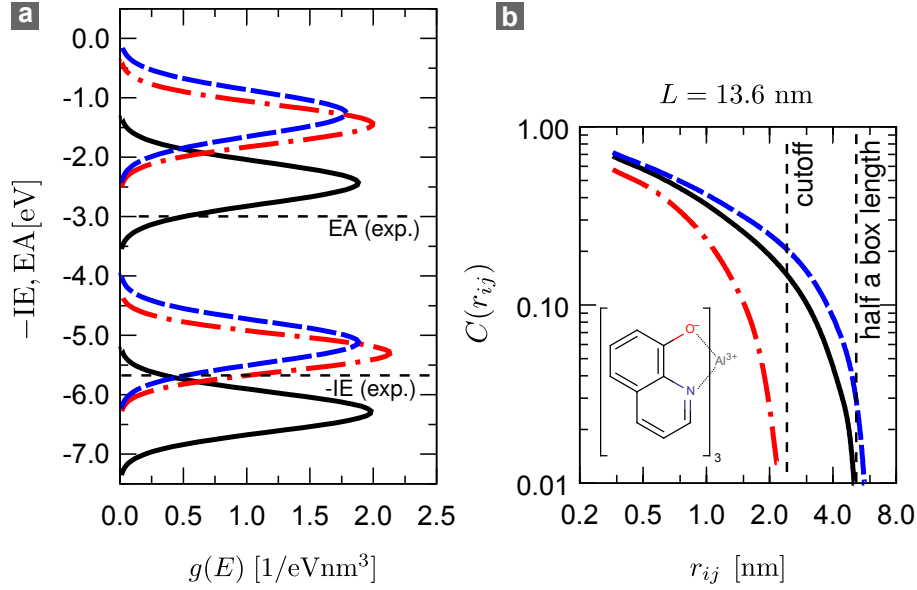
Figure 4: (a) Densities of ionization energies and electron affinities in Alq$_3$ computed with a cutoff (red dashed-dotted line), and with long-range embedding in the cubic (blue dashed line) or isotropic limit (black solid line). The latter yields much better agreement with ionization energies and electron affinities extracted from photoelectron spectroscopy[44,45] (horizontal dashed lines). Note that photoelectron spectroscopy probes the tail of the densities of states. (b) Spatial correlation function evaluated with the computational procedure indicated by the line style as described in (a). In the cutoff description, correlations are truncated at the cutoff and in the case of long-range embedding at half the box length, indicative of a finite-size effect.

interactions may appear paradoxical, since we have already identified (see Eqs. 16 and 17) that uncompensated multipolar moments are responsible for linking the energetics of microscopic states to mesoscopic order. In amorphous semiconductors, however, these moments should by definition tend to zero on large scales. Still, the question remains on what scale multipolar moments start to decay to give way to an electrostatically isotropic continuum - and how this reflects in the energy landscape of these materials.

The atomistic configurations of 4096 molecules of Alq_3[10] were prepared via molecular dynamics simulations with tailored force fields adapted from OPLS-AA.[46] The starting configurations were first randomized at high temperatures and then quenched to 300 K, with subsequent equilibration over a time period of several nanoseconds. The computed DOS, that is the hystogram of solid state IEs and EAs of all molecules, is shown in Fig. 4a. For both IEs and EAs, we employed three different computational procedures: a cutoff-based description (dashed blue lines) and the long-range embedding protocol from Sec. 2 with either a cubic shape term (dotted-dashed red lines) or no shape term at all (solid black lines). The latter should be thought of as an isotropic limit, which assumes all multipolar moments to average out on a mesoscopic scale.

First, we note that the cutoff-based description and long-range description in the cubic limit match closely. This is not surprising in that both frameworks are virtually identical with respect to their long-range behavior. In particular, a spherical cutoff always implicitly includes long-range contributions tackled by the shape term from Eq. 16, as also illustrated by the convergence scan from Fig. 1b. The agreement between these two approaches therefore serves as a mere consistency check. More striking is, however, the 1 eV difference between the DOS obtained in the cubic vs. isotropic limit – proving that this system, though structurally amorphous, features a conditionality in the interaction sum which is picked up by both the cutoff and cubic limit. This observation implies a challenge for simulations, as it is at this stage impossible to pinpoint the origin of the mesoscopic moments, as they can be due to subtle preferential ordering present also in the real system or to finite-size-induced

fluctuations, or both. If, for example, preferential ordering really plays a role, then how large do system sizes have to be in order to not truncate structural correlation functions before convergence (at least in a thin-film sense) is achieved? On the other hand, if finite-size effects are exclusively responsible for the observed differences, then the isotropic limit should be a more appropriate description. Indeed, this limit appears to agree better with energy levels extracted experimentally from the DOS onset,[44,45] as indicated by the dashed horizontal lines.

In addition to level positioning and alignment, we furthermore consider the spatial correlation function $C(R)$ of the energy landscape, see Fig. 4b. Correlations of this landscape have been shown to result in the characteristic Poole-Frenkel behavior of the mobility in many organic semiconductors. Fig. 4b, however, illustrates that atomistic simulations tend to truncate this correlation function in a cutoff-based description for pair separations larger than the cutoff r_c (here: 3 nm). Even in a long-range description, the finite system size forces correlations to zero at approximately half the box length, with the cubic limit resulting in stronger correlations. In amorphous semiconductors, calculating the DOS and spatial correlations thereof may hence be more involved than initially anticipated: In particular, large system sizes (to be tackled with advanced simulation protocols) will be required in order to disentangle finite-size artifacts from ordering effects. This challenge has a direct link to compound screening for organic light emitting diodes, where tuning of level alignment is crucial in order to guarantee device functionality and bypass degradation mechanisms.[47]

4   Conclusions

The long-range polarized embedding approach presented in this work targets the quantitative evaluation of the energy landscape of charge carriers, charge transfer, and excited states in molecular systems. Implemented in the VOTCA package,[10] it can be readily parametrized from first principles, while accounting for both polarization and electrostatic effects. In

particular, it successfully copes with the slowly-convergent charge-quadrupole interaction encountered in many molecular materials.

Using this approach, we illustrated the effect of molecular orientation and order on energy level profiles in thin crystalline films of organic semiconductors. We showed how orientational effects supersede packing effects, investigating how out-of-plane quadrupolar moments can cause an electrostatic stabilization or destabilization of charge carriers, irrespective of the packing motif. Our study of amorphous materials revealed that long-range effects persist even in structurally disordered systems, as either a finite-size induced simulation artifact or a result of residual structural correlations, or both.

Acknowledgement

Supporting Information Available

Derivation of the thin-film and bulk shape factors, electrostatic and induction terms, lists of damping kernels and 2D-Ewals potenials. This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(1) May, V.; Kühn, O. *Charge and Energy Transfer Dynamics in Molecular Systems*, 3rd ed.; Wiley-VCH: Weinheim, 2011.

(2) Miller Jr., J. H.; Villagrán, M. Y. S.; Maric, S.; Briggs, J. M. *Physica B: Condensed Matter* **2015**, *460*, 119–125.

(3) Kaneko, M.; Okada, T. In *Molecular Catalysts for Energy Conversion*; Okada, D. T., Kaneko, P. D. M., Eds.; Springer Series in Materials Science 111; Springer Berlin Heidelberg, 2009; pp 37–65.

(4) Lin, P.; Yan, F. *Adv. Mater. Weinheim* **2012**, *24*, 34–51.

(5) Yost, S. R.; Van Voorhis, T. *J. Phys. Chem. C* **2013**, *117*, 5617–5625.

(6) Ryno, S. M.; Risko, C.; Brédas, J.-L. *J. Am. Chem. Soc.* **2014**, *136*, 6421–6427.

(7) Idé, J.; Méreau, R.; Ducasse, L.; Castet, F.; Bock, H.; Olivier, Y.; Cornil, J.; Beljonne, D.; D'Avino, G.; Roscioni, O. M.; Muccioli, L.; Zannoni, C. *J. Am. Chem. Soc.* **2014**, *136*, 2911–2920.

(8) Kordt, P.; van der Holst, J. J. M.; Al Helwi, M.; Kowalsky, W.; May, F.; Badinski, A.; Lennartz, C.; Andrienko, D. *Adv. Funct. Mater.* **2015**, *25*, 1955–1971.

(9) Yavuz, I.; Martin, B. N.; Park, J.; Houk, K. N. *J. Am. Chem. Soc.* **2015**, *137*, 2856–2866.

(10) Rühle, V.; Lukyanov, A.; May, F.; Schrader, M.; Vehoff, T.; Kirkpatrick, J.; Baumeier, B.; Andrienko, D. *J. Chem. Theory. Comput.* **2011**, *7*, 3335–3345.

(11) Baumeier, B.; Stenzel, O.; Poelking, C.; Andrienko, D.; Schmidt, V. *Phys. Rev. B* **2012**, *86*, 184202.

(12) Stenzel, O.; Hirsch, C.; Brereton, T.; Baumeier, B.; Andrienko, D.; Kroese, D.; Schmidt, V. *Multiscale Model. Simul.* **2014**, *12*, 1108–1134.

(13) Kordt, P.; Stenzel, O.; Baumeier, B.; Schmidt, V.; Andrienko, D. *J. Chem. Theory Comput.* **2014**, *10*, 2508–2513.

(14) Gemünden, P.; Poelking, C.; Kremer, K.; Andrienko, D.; Daoulas, K. C. *Macro-molecules* **2013**, *46*, 5762–5774.

(15) Gemünden, P.; Poelking, C.; Kremer, K.; Daoulas, K.; Andrienko, D. *Macromol. Rapid Commun.* **2015**, *36*, 1047–1053.

(16) Poelking, C.; Tietze, M.; Elschner, C.; Olthof, S.; Hertel, D.; Baumeier, B.; Würthner, F.; Meerholz, K.; Leo, K.; Andrienko, D. *Nat. Mater.* **2014**, *14*, 434–439.

(17) Poelking, C.; Andrienko, D. *J. Am. Chem. Soc.* **2015**, 6320–6326.

(18) Leeuw, S. W. d.; Perram, J. W.; Smith, E. R. *Proc. R. Soc. Lond. A* **1980**, *373*, 27–56.

(19) Leeuw, S. W. d.; Perram, J. W.; Smith, E. R. *Proc. R. Soc. Lond. A* **1980**, *373*, 57–66.

(20) Smith, E. R. *Proc. R. Soc. Lond. A* **1981**, *375*, 475–505.

(21) Fitzner, R.; Mena-Osteritz, E.; Mishra, A.; Schulz, G.; Reinold, E.; Weil, M.; Körner, C.; Ziehlke, H.; Elschner, C.; Leo, K.; Riede, M.; Pfeiffer, M.; Uhrich, C.; Bäuerle, P. *J. Am. Chem. Soc.* **2012**, *134*, 11064–11067.

(22) Banerjee, R.; Novák, J.; Frank, C.; Lorch, C.; Hinderhofer, A.; Gerlach, A.; Schreiber, F. *Phys. Rev. Lett.* **2013**, *110*, 185506.

(23) Wynands, D.; Erber, M.; Rentenberger, R.; Levichkova, M.; Walzer, K.; Eichhorn, K.-J.; Stamm, M. *Org. Electron.* **2012**, *13*, 885–893.

(24) Schwarze, M.; Tress, W.; Beyer, B.; Gao, F.; Scholz, R.; Poelking, C.; Ortstein, K.; Günther, A. A.; Kasemann, D.; Andrienko, D.; Leo, K. *Science* **2016**, *352*, 1446–1449.

(25) Ewald, P. P. *Ann. Phys. (Berlin)* **1921**, *369*, 253–287.

(26) Giese, T. J.; Panteva, M. T.; Chen, H.; York, D. M. *J. Chem. Theory Comput.* **2015**, *11*, 436–450.

(27) Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933–5947.

(28) Stone, A. J. *J. Chem. Theory Comput.* **2005**, *1*, 1128–1132.

(29) D'Avino, G.; Muccioli, L.; Zannoni, C.; Beljonne, D.; Soos, Z. G. *J. Chem. Theory. Comput.* **2014**, *10*, 4959–4971.

(30) Thole, B. *Chem. Phys.* **1981**, *59*, 341–350.

(31) Bereau, T.; Andrienko, D.; von Lilienfeld, O. A. *J. Chem. Theory Comput.* **2015**, *11*, 3225–3233.

(32) Stone, A. J. *The Theory of intermolecular forces*; Clarendon Press: Oxford, 1997.

(33) Misquitta, A. J.; Stone, A. J. *J. Chem. Phys.* **2006**, *124*, 024111.

(34) Hättig, C.; Heß, B. A. *Mol. Phys.* **1994**, *81*, 813–824.

(35) van Duijnen, P. T.; Swart, M. *J. Phys. Chem. A* **1998**, *102*, 2399–2407.

(36) Leslie, M. *Mol. Phys.* **2008**, *106*, 1567–1578.

(37) Smith, W. *CCP5 Newsletter* **1998**, *46*, 18–30.

(38) Kirkwood, J. G. *J. Chem. Phys.* **1934**, *2*, 351.

(39) Hu, Z. *J. Chem. Theory Comput.* **2014**, *10*, 5254–5264.

(40) Pan, C.; Hu, Z. *J. Chem. Theory Comput.* **2014**, *10*, 534–542.

(41) Crozier, P. S.; Rowley, R. L.; Spohr, E.; Henderson, D. *J. Chem. Phys.* **2000**, *112*, 9253.

(42) De Leeuw, S. W.; Perram, J. W. *Mol. Phys.* **1979**, *37*, 1313–1322.

(43) Bürckstümmer, H.; Tulyakova, E. V.; Deppisch, M.; Lenze, M. R.; Kronenberg, N. M.; Gsänger, M.; Stolte, M.; Meerholz, K.; Würthner, F. *Angew. Chem. Int. Edit.* **2011**, *50*, 11628–11632.

(44) Sun, S.-S., Dalton, L. R., Eds. *Introduction to organic electronic and optoelectronic materials and devices*; Optical science and engineering 133; CRC Press: Boca Raton, 2008.

(45) Mori, T.; Kim, H.-G.; Mizutani, T.; Lee, D.-C. *Jpn. J. Appl. Phys.* **2001**, *40*, 5346–5349.

(46) Lukyanov, A.; Lennartz, C.; Andrienko, D. *Phys. Stat. Sol. (a)* **2009**, *206*, 2737–2742.

(47) May, F.; Baumeier, B.; Lennartz, C.; Andrienko, D. *Phys. Rev. Lett.* **2012**, *109*, 136401.

Graphical TOC Entry
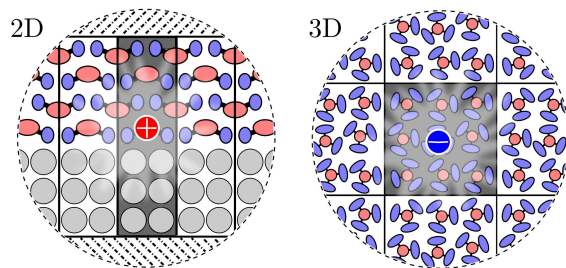
An ion embedded in a periodically repeated molecular slab or a bulk crystal.

Supporting Information

Long-Range Embedding of Molecular Ions and Excitations in a Polarizable Molecular Environment

Carl Poelking[1,2] and Denis Andrienko[2]

[1]Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany

[2]Heidelberg Graduate School of Fundamental Physics, INF 226, 69120 Heidelberg, Germany

S1   Anisotropic Kernels

Throughout the Supporting Information, we use the following definitions: $\vec{R}_L^{ab} = \vec{r}_a - \vec{r}_b + \vec{L}$ denotes the $\vec{L}$- (image-box-vector-) shifted particle-particle interaction vector. The outer products of real- and reciprocal-space vectors are abbreviated via $\mathbb{R} = \vec{R} \otimes \vec{R}$ and $\mathbb{K} = \vec{k} \otimes \vec{k}$, respectively. The dyadic product between matrices M and N is written as $M : N$. Finally, the atomic quadrupole is defined as $\tilde{\theta}_{\alpha\beta} = \sum_a q^a (\frac{1}{2} r_\alpha^a r_\beta^a - \frac{1}{6} \delta_{\alpha\beta} (r^a)^2)$, which differs from the conventional definition by a factor $\frac{1}{3}$, but avoids prefactors in the interaction kernels.

In real space, the T_ν^{pb} symbols from Eq. 13 describe the orientation dependence of the interaction between polar sites p and b with charge q, dipole $\vec{\mu}$ and quadrupole $\tilde{\theta}$ (in Cartesian representation, see the definition above). They read

$$T_0^{pb} = q^p q^b, \tag{S1}$$

$$T_1^{pb} = \vec{\mu}^p \cdot \vec{\mu}^b + (q^p \vec{\mu}^b - q^b \vec{\mu}^p) \cdot \vec{R}^{pb}, \tag{S2}$$

$$T_2^{pb} = 2\tilde{\Theta}^p : \tilde{\Theta}^b + (q^p \tilde{\Theta}^b + q^b \tilde{\Theta}^p) : \mathbb{R}^{pb} - (\vec{\mu}^p \cdot \vec{R}^{pb})(\vec{\mu}^b \cdot \vec{R}^{pb}) - \tag{S3}$$

$$- 2(\vec{\mu}^b \otimes \vec{R}^{pb}) : \tilde{\Theta}^p + 2(\vec{\mu}^p \otimes \vec{R}^{pb}) : \tilde{\Theta}^b,$$

$$T_3^{pb} = -4(\tilde{\Theta}^p \cdot \vec{R}^{pb})(\tilde{\Theta}^b \cdot \vec{R}^{pb}) - (\vec{\mu}^p \cdot \vec{R}^{pb})(\tilde{Q}^b : \mathbb{R}^{pb}) + (\vec{\mu}^b \cdot \vec{R}^{pb})(\tilde{Q}^p : \mathbb{R}^{pb}), \tag{S4}$$

$$T_4^{pb} = (\tilde{Q}^p : \mathbb{R}^{pb})(\tilde{Q}^b : \mathbb{R}^{pb}). \tag{S5}$$

The k-space anisotropy is absorbed in the structure factors S of the multipole density, featuring in Eq. 12:

$$S(\vec{k}; [\mathcal{P}^{(s)}]) = \sum_P \left(q^p + i\vec{\mu}^p \cdot \vec{k} - \tilde{\Theta}^p : \mathbb{K} \right) \exp \left(i\vec{k} \cdot \vec{r}_p \right). \tag{S6}$$

Note that the dipole moment $\vec{\mu}$ in the above expressions for T_ν^{pb} and S also incorporates the induced moment $\Delta\vec{\mu}$. If an energy splitting in terms of electrostatic and polarization contributions is desired, the T_ν^{pb} symbols, as well as S have be decomposed accordingly.

S2   Isotropic Kernels

The distance dependence of the interaction in Eqs. 12, 13, 15 is described by the isotropic kernels $A(k)$, $B_\nu(R)$ and $C_\nu(R)$, respectively, next to powers of k and R that already appear in the scalar contractions of S and T_ν^{pb} above:

$$A(k) = k^{-2} \exp \left(-\frac{k^2}{4\beta^2} \right), \tag{S7}$$

$$B_0(R) = R^{-1} \text{erfc}(\beta R), \tag{S8}$$

$$B_\nu(R) = R^{-2} \left[(2\nu - 1)B_{\nu-1}(R) + \frac{2^\nu \beta^{2\nu-1}}{\sqrt{\pi}} \exp(-\beta^2 R^2) \right], \tag{S9}$$

$$C_0(R) = R^{-1} \text{erf}(\beta R), \tag{S10}$$

$$C_\nu(R) = R^{-2} \left[(2\nu - 1)C_{\nu-1}(R) - \frac{2^\nu \beta^{2\nu-1}}{\sqrt{\pi}} \exp(-\beta^2 R^2) \right]. \tag{S11}$$

S3   Thole Damping Kernels

The Thole model damps induced-induced interactions at short distances. For faster convergence, damping of induced-permanent interactions – though not mandatory – may also be advised. With the polarizability $\tilde{\alpha}$ (here written more generally as a tensor quantity), an

effective interaction distance is defined as

$$u^{pb}(R) = \left(\frac{\tilde{\alpha}^p : \tilde{\alpha}^b}{3}\right)^{-1/6} R. \tag{S12}$$

With an exponential smearing function $\sim \exp(-\gamma u^3)$ as used in the AMOEBA force field,[27] where $\gamma = 0.39$, the damping functions $\Lambda_{2\nu+1}$ are

$$\Lambda_3(R) = 1 - \exp\left[-\gamma u^{pb}(R)^3\right], \tag{S13}$$

$$\Lambda_5(R) = 1 - \left[1 + \gamma u^{pb}(R)^3\right] \exp\left[-\gamma u^{pb}(R)^3\right], \tag{S14}$$

$$\Lambda_7(R) = 1 - \left[1 + \gamma u^{pb}(R)^3 + \frac{3}{5}\gamma^2 u^{pb}(R)^6\right] \exp\left[-\gamma u^{pb}(R)^3\right], \tag{S15}$$

$$\Lambda_9(R) = 1 - \left[1 + \gamma u^{pb}(R)^3 + \frac{18}{35}\gamma^2 u^{pb}(R)^6 + \frac{9}{35}\gamma^3 u^{pb}(R)^9\right] \exp\left[-\gamma u^{pb}(R)^3\right]. \tag{S16}$$

As higher orders of the damped interaction tensors are obtained from derivatives of tensors of lower order, the damping functions can be applied such that interaction terms that scale as $\frac{R_\alpha \cdots R_\beta}{R^n}$ (with $\alpha, \beta \in \{x, y, z\}$) are damped by Λ_n; or, put simpler, terms of power $-n$ in the isotropic distance R are damped by Λ_n. For the real-space interaction from Eq. 13, the damping is hence prescribed by the isotropic kernel B_ν, which scales as $R^{-(2\nu+1)}$ for $R \ll 1/\beta$. For the C_ν, no damping is required as long as the convergence parameter β is sufficiently small ($1/\beta \gtrsim$ nm), since we are then dealing with a purely long-ranged compensation term.

Special care has to be taken for the interaction tensors T from Eq. 7, where different summands appearing in T have to be multiplied by a different $\Lambda_{2\nu+1}$ according to the rule above.

S4 Field Calculation

For the evaluation of electric fields, a decomposition analogous to Eq. 11 holds:

$$\vec{E}(\vec{r}_p, [\tilde{\mathcal{B}}^*]) = \vec{E}_r(\vec{r}_p, [\tilde{\mathcal{B}}^*]) + \vec{E}_k(\vec{r}_p, [\mathcal{B}^*]) - \vec{E}_{\mathrm{si}}(\vec{r}_p, [\mathcal{P}^{(n)}]) - \vec{E}_c(\vec{r}_p, [\mathcal{P}^{(n)}]) + \vec{E}_s^*(\vec{r}_p, [\mathcal{B}^*])$$

(S17)

The individual terms correspond to the real-space $(\vec{E}_r)$ and reciprocal-space contribution $(\vec{E}_k)$, self-interaction correction $(\vec{E}_c')$, aperiodic subtraction $(\vec{E}_c)$ and shape contribution $(\vec{E}_s)$, respectively:

$$\vec{E}_r(\vec{r}_p, [\tilde{\mathcal{B}}^*]) = -\frac{1}{4\pi\varepsilon_0} \sum_{\tilde{\mathcal{B}}^*} \sum_{\nu=1}^{3} \vec{T}_\nu^{pb} \Lambda_{2\nu+1}^{pb} B_\nu(R_L^{pb}),$$

(S18)

$$\vec{E}_k(\vec{r}_p, [\mathcal{B}^*]) = -\frac{1}{4\pi\varepsilon_0} \frac{4\pi}{V} \sum_{\vec{k} \neq 0}^{\infty} i\vec{k} \exp(i\vec{k} \cdot \vec{r}_p) A(k) S^*(\vec{k}; [\mathcal{B}^*]),$$

(S19)

$$\vec{E}_c(\vec{r}_p, [\mathcal{P}^{(n)}]) = -\frac{1}{4\pi\varepsilon_0} \sum_{P^{(n)}}' \sum_{\nu=1}^{3} \vec{T}_\nu^{pb} \Lambda_{2\nu+1}^{pb} B_\nu(R_L^{pb}), \quad \vec{E}_{\mathrm{si}}(\vec{r}_p, [\mathcal{P}^{(n)}]) = -\frac{1}{4\pi\varepsilon_0} \frac{4\alpha^3}{3\sqrt{\pi}} \vec{\mu}^p.$$

(S20)

The vector symbols $\vec{T}_\nu^{pb}$ again absorb the orientation dependence of the interaction,

$$\vec{T}_1^{pb} = -q_b \vec{R}^{pb} + \vec{\mu}^b,$$

(S21)

$$\vec{T}_2^{pb} = 2\tilde{\Theta}^b \cdot \vec{R}^{pb} - (\vec{\mu}^b \cdot \vec{R}^{pb})\vec{R}^{pb},$$

(S22)

$$\vec{T}_3^{pb} = -(\tilde{\Theta}^b : \mathbb{R})\vec{R}^{pb}.$$

(S23)

The conditionality in the field calculation is removed through application of the appropriate shape term,

$$\vec{E}_s(\vec{r}_p, [\mathcal{B}^*]) = -\frac{1}{4\pi\varepsilon_0} \frac{4\pi}{3V} \vec{M}^{\mathcal{B}(n)},$$

(S24)

$$\vec{E}_s(\vec{r}_p, [\mathcal{B}^*]) = -\frac{1}{4\pi\varepsilon_0} \frac{4\pi}{V} M_z^{\mathcal{B}(n)} \hat{e}_z,$$

(S25)

4

where the former expression corresponds to a cubic limit, the latter to a thin-film limit with vanishing thickness.

S5   Reciprocal-Space Convergence

In order to detect potential reciprocal space resonances, k-vectors are sorted according to the rating functions

$$g_3(\vec{k}) = \frac{S(k_x \hat{e}_x) S(k_y \hat{e}_y) S(k_z \hat{e}_z)}{\left(\langle S(k_x' \hat{e}_x)\rangle_x \langle S(k_y' \hat{e}_y)\rangle_y \langle S(k_z' \hat{e}_z)\rangle_z\right)^{2/3}} A(k),$$

$$g_2(\vec{k}) = \frac{S(k_i \hat{e}_i) S(k_j \hat{e}_j)}{\left(\langle S(k_i' \hat{e}_i)\rangle_i \langle S(k_j' \hat{e}_j)\rangle_j\right)^{2/3}} A(k).$$

Here, g_3 is the rating function used for off-axis k-vectors with three non-zero components, g_2 is the rating function for in-plane k-vectors with two non-zero components. The ratings are hence constructed from structure factors $S(k_i \hat{e}_i)$ evaluated along the system axes $\hat{e}_x$, $\hat{e}_y$ and $\hat{e}_z$. The convergence criterion for off-axis and in-plane k-vectors is established as the root-mean square contribution of k-shells to fields and energies associated with the polarization cloud $\mathcal{P}^{(s)}$.

S6   Potentials for 2D-Periodic Embedding

For the 2D-periodic description employed in Fig. S1, the potential is calculated according to the standard Ewald-type decomposition[41,42]

$$\phi(\vec{r}_p, [\tilde{\mathcal{B}}^*]) = \phi_r(\vec{r}_p, [\tilde{\mathcal{B}}^*]) + \phi_k(\vec{r}_p, [\mathcal{B}^*]) - \phi_c(\vec{r}_p, [\mathcal{P}^{(n)}]) \tag{S26}$$

Shape contributions are explicitly included in the $(k = 0)$-term of $\phi_k(\vec{r}_p, [\mathcal{B}^*])$, where charge-quadrupole contributions are exempt, as these prove absolutely convergent in 2D-periodic

systems. The individual contributions to Eq. S26 read

$$\phi_r(\vec{r}_p, [\tilde{\mathcal{B}}^*]) = \frac{1}{4\pi\varepsilon_0} \sum_{\tilde{B}^*} \frac{q^{b(n)}\mathrm{erfc}\left(\beta R_L^{pb}\right)}{R_L^{pb}},$$

$$\phi_k(\vec{r}_p, [\mathcal{B}^*]) = \frac{1}{4\pi\varepsilon_0} \sideset{}{'}\sum_{\vec{k}\neq\vec{0}} \sum_B \frac{2\pi q^{b(n)}}{|\vec{L}_a \times \vec{L}_b|} \cos\left(\vec{k}\cdot\vec{R}^{pb}\right) \Bigg[$$

$$e^{-k(z_p-z_b)}\mathrm{erfc}\left(\frac{k}{2\beta} - \beta(z_p - z_b)\right) +$$

$$+e^{+k(z_p-z_b)}\mathrm{erfc}\left(\frac{k}{2\beta} + \beta(z_p - z_b)\right)\Bigg] -$$

$$-\frac{1}{4\pi\varepsilon_0} \sum_B \frac{2\pi q^{b(n)}}{|\vec{L}_a \times \vec{L}_b|} \left[\frac{e^{-\beta^2(z_a-z_b)^2}}{\sqrt{\pi}\beta} + (z_a - z_b)\mathrm{erf}\left(\beta(z_a - z_b)\right)\right], \qquad (S27)$$

$$\phi_c(\vec{r}_p, [\mathcal{P}^{(n)}]) = \frac{1}{4\pi\varepsilon_0} \sum_P \frac{q^{p(n)}\mathrm{erf}\left(\beta R_L^{pb}\right)}{R_l^{pb}}. \qquad (S28)$$

$\vec{L}_a$, $\vec{L}_b$ are the simulation-cell vectors in the periodic plane. Note that self-interaction terms with $R_L^{pb} = 0$ in $\phi_c(\vec{r}_p, [\mathcal{P}^{(n)}])$ are calculated using the limit $\mathrm{erf}(\beta R_L^{pb})/R_L^{pb} \to 2\beta/\sqrt{\pi}$ as $R_L^{pb} \to 0$.

Note that the standard Ewald decomposition is typically performed for the self-interaction energy of a discrete point-charge cloud. The expressions in real and reciprocal space then build on a double sum over point charges, which count all pairs twice. Hence, a factor $1/2$ is normally added in the formulae in order to avoid this double-counting. Here, however, we consider the potential generated by the periodic background density. This should not be multiplied by a factor $1/2$, and accordingly differs from the expressions given for the energy by, e.g., Crozier et al.[41]

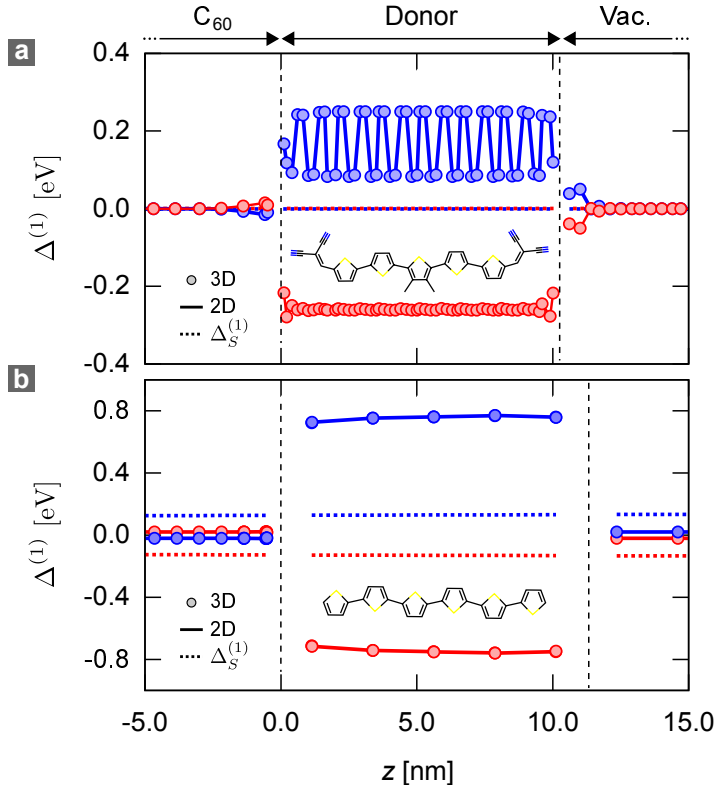# S7   2D- versus 3D-periodic description of thin films

Figure S1: **2D- versus 3D-periodic description of thin films.** Electrostatic energy level profile $\Delta^{(1)}(z)$ across a D5M (a) and 6T (b) 10 nm thin film for holes (red) and electrons (blue), once computed in a 2D-periodic (lines), and once in a 3D-periodic framework under application of the appropriate thin-film shape term from Eq. 17 (symbols). Dashed blue and red lines mark the contribution from this shape term, addition of which leads to perfect agreement between the 2D- and 3D-periodic description, already for a finite vacuum buffer of twice the thin-film thickness $d = 20$ nm, achieved by scaling the unit-cell dimension along the z-direction in the 3D-periodic framework.

Fig. S1 shows that the slab limit from Eq. 17 is indeed suited to mimic a 2D-periodic setup, with periodicity along the thin-film normal (aligned with the z-axis) effectively removed. To this end, we compare *electrostatic* level profiles $\Delta^{(1)}(z)$ across a C_{60}:D5M and C_{60}:6T interface, obtained once with the 3D-periodic framework detailed above and once with the 2D-periodic description. The results for both approaches (indicated via dots and lines, respectively) coincide already for small z-axis box vectors (as required in the 3D-periodic framework) of three times the film thickness, *if* the shape contribution is taken into account.

Note that these level profiles still exclude polarization effects, to be included by solving Eq. 2 as outlined in the section below.

S8 Shape factors

We will start from Eq. 3.12 of Smith,[20] which implies that the electrostatic potential for a system with zero-rank multipoles (charges) can be written as

$$\phi_{k=0}^{(0)}\left(\vec{r},[\mathcal{B}^*]\right) = -\frac{1}{4\pi\varepsilon_0}\sum_{\mathcal{B}} q^{b(n)}\mathcal{J}(\vec{r}-\vec{r}_b, P), \tag{S29}$$

where $\mathcal{J}$ is a shape factor defined via

$$\mathcal{J}(\vec{r}, P) = \frac{2}{\pi^2 V}\int_{-\infty}^{\infty} d^3v \frac{(\vec{r}\cdot\vec{v})^2}{|\vec{v}|^2}\int_P d^3\rho\, e^{2i\vec{\rho}\cdot\vec{r}}. \tag{S30}$$

P defines the "primitive" summation shape (cube, sphere, cylinder, . . .), whose surface is given by $P(\vec{r}) = 0$.

The shape factor can be readily evaluated for cube/sphere and slab structures:

$$\mathcal{J}_{\text{sphere}}(\vec{r}, P) = \frac{2\pi}{3V}|\vec{r}|^2, \tag{S31}$$

$$\mathcal{J}_{\text{slab}}(\vec{r}, P) = \frac{2\pi}{V}z^2. \tag{S32}$$

The generalization of Eq. S29 to higher-rank multipoles is achieved by substituting the charge scalar $q^{b(n)}$ by the operator $\mathcal{M}_{\vec{r}}^{b(n)} = q^{b(n)} - \vec{\mu}^{b(n)}\cdot\vec{\nabla} + \theta^{b(n)} : \left(\vec{\nabla}\otimes\vec{\nabla}\right)$, which performs the expansion up to the atomic quadrupole. The corresponding potential and shape contribution to the system energy are

$$\phi_{k=0}\left(\vec{r}\right) = -\frac{1}{4\pi\varepsilon_0}\sum_{\mathcal{B}} \mathcal{M}_{\vec{r}}^{b(n)}\mathcal{J}(\vec{r}-\vec{r}_b, P), \tag{S33}$$

$$W_* \left[P^{(s)}; \mathcal{B} \right] = -\frac{1}{4\pi\varepsilon_0} \sum_{\mathcal{B}} \sum_{P} \mathcal{M}_{\vec{r}}^{a(s)} \phi_{k=0} \left(\vec{r} \right) |_{\vec{r}_a}, \tag{S34}$$

which after some algebra results in Eqs. 16,17 of the main text.

S9 Aperiodic correction

Following the idea of Ewald, we split the electrostatic potential onto short- and long-range contributions,

$$\phi^{(0)} \left(\vec{r}_p; [\tilde{\mathcal{B}}^*] \right) = \frac{1}{4\pi\varepsilon_0} \sum_{\tilde{\mathcal{B}}^*} q^{b(n)} \left\{ \frac{\mathrm{erfc}(\beta|\vec{r}_p - \vec{r}_b|)}{\vec{r}_p - \vec{r}_b} + \frac{\mathrm{erf}(\beta|\vec{r}_p - \vec{r}_b|)}{\vec{r}_p - \vec{r}_b} \right\}. \tag{S35}$$

The first interaction term in the curly brackets is short-ranged, and hence efficiently summed up in real space, which eventually results in $W_r[\mathcal{P}^{(s)}; \tilde{\mathcal{B}}^*]$ in Eq. 11. We now add and subtract the neutral foreground $\mathcal{P}^{(n)}$ to the sum over the second, long-ranged, interaction term: $\sum_{\tilde{\mathcal{B}}^*} = \sum_{\tilde{\mathcal{B}}^* \cup \mathcal{P}^{(n)}} - \sum_{\mathcal{P}^{(n)}}$. The sum over $\tilde{\mathcal{B}}^* \cup \mathcal{P}^{(n)} \equiv \mathcal{B}^*$ results in the contributions $W_k[\mathcal{P}^{(s)}; \mathcal{B}]$ and $W_*[\mathcal{P}^{(s)}; \mathcal{B}]$, whereas the sum over $\mathcal{P}^{(n)}$ gives rise to $W_{\mathrm{si}}[\mathcal{P}^{(s)}; \mathcal{P}^{(n)}]$ and $W_c[\mathcal{P}^{(s)}; \mathcal{P}^{(n)}]$.

Specifically, for zero-rank multipoles (charges)

$$W_{\mathrm{si,c}}^{(0)} = \sum_{\mathcal{P}^{(s)}} q^{p(s)} \phi_{\mathrm{si,c}}^{(0)} \left(\vec{r}_p; [\tilde{\mathcal{B}}^*] \right) = -\frac{1}{4\pi\varepsilon_0} \sum_{\mathcal{P}} \sum_{\mathcal{P}'} q^{p(s)} q^{p'(n)} \frac{\mathrm{erf}(\beta|\vec{r}_p - \vec{r}_{p'}|)}{|\vec{r}_p - \vec{r}_{p'}|} = W_{\mathrm{si}}^{(0)} + W_c^{(0)},$$

where

$$W_{\mathrm{si}}^{(0)} = \frac{1}{4\pi\varepsilon_0} \sum_{\mathcal{P}} q^{p(s)} q^{p(n)} \lim_{r \to 0} \frac{\mathrm{erf}(\beta r)}{r},$$

$$W_c^{(0)} = \frac{1}{4\pi\varepsilon_0} \sum_{\mathcal{P}} \sum_{\mathcal{P}'} (1 - \delta_{pp'}) q^{p(s)} q^{p'(n)} \frac{\mathrm{erf}(\beta|\vec{r}_p - \vec{r}_{p'}|)}{|\vec{r}_p - \vec{r}_{p'}|}.$$

Generalization of these expressions to higher multipoles (up to quadrupoles) is possible

via the substitution

$$q^{b(n)} \rightarrow q^{b(n)} - \vec{\mu}^{b(n)} \cdot \vec{\nabla} + \theta^{b(n)} : \left(\vec{\nabla} \otimes \vec{\nabla} \right) , \tag{S36}$$

$$q^{p(s)} \rightarrow q^{p(s)} + \vec{\mu}^{p(s)} \cdot \vec{\nabla} + \theta^{p(s)} : \left(\vec{\nabla} \otimes \vec{\nabla} \right) , \tag{S37}$$

which finally results in Eqs. 14, 15 of the main text.

S10 Parameters of the polarizable force-field

Below, we list the set of atomic multipoles and polarizabilities used in the examples from the Validation section of the main text. The listing is structured according to compounds (ZnPc, D5M, 6T, EL86, PEN) and charge states (neutral, cation, anion). Each section defines the polarizable force field for the corresponding compound and charge state via a verbatim representation of the input file used by the VOTCA::CTP code.[10]

The file format largely resembles the format used by A. Stone in his GDMA code[28] and is structured as follows: The first two lines are comment lines, followed by a line which indicates the units of the atomic coordinates (Å in this example):

```
! Comment line 1
! Comment line 2
Units angstrom
```

After this header section, the atomic parameters are listed atom by atom. For a single atom, the associated entry reads, for example,

```
N +0.0002240 +9.8202378 +3.7242187 Rank 2
   -0.2150281
   -0.4548875 -0.0000536 +0.3131389
   +0.4246758 +0.0000774 -0.4983059 -0.1450614 -0.0000414
    P +2.1668289 +0.0000000 +0.0000000 +2.1668289 +0.0000000 +2.1668289
```

The first line of this section starts with the atomic element (here: N = nitrogen), followed by the atomic x-, y-, and z-coordinate in the appropriate units declared previously. The "Rank" keyword defines the rank up to which the multipolar expansion for this atom is performed (here: $l = 2$, i.e., up to quadrupoles).

The subsequent lines are one-by-one devoted to the $2l + 1$ multipole components of rank l in spherical-tensor notation and in ascending order of l, starting from the atomic charge Q_{00} in the first line (here: -0.2150281), then the atomic dipole Q_{1z}, Q_{1x}, Q_{1y} in the second line (here: -0.4548875 -0.0000536 +0.3131389), finally the atomic quadrupole Q_{20}, Q_{21c}, Q_{21s}, Q_{22c}, Q_{22s} (here: +0.4246758 +0.0000774 -0.4983059 -0.1450614 -0.0000414). Note that in the dipole line, the z-component is listed *first*. All multipole components are in atomic units. Multipole moments of rank l hence have to be multiplied by $e \cdot a_0^l$, where e is the positive elementary charge, and a_0 the Bohr radius.

The final line in the atom section, prefaced by "P", defines the atomic polarizability tensor $\tilde{\alpha}$ in upper-diagonal form, i.e., in the order α_{xx}, α_{xy}, α_{xz}, α_{yy}, α_{yz}, α_{zz}. The units used for the polarizabilities are Å^3. Note that in this work we used only isotropic atomic polarizabilities, obtained by iterative isotropic scaling of the Thole atomic polarizabilities,[35] such that the self-consistently calculated classical polarizability tensor matches the molecular tensor in terms of *polarizable volume*, as measured by the volume of the polarizability ellipsoid $\sim 1/\Pi_{i=1}^3 \sqrt{\tilde{\alpha}_i}$. Here, $\tilde{\alpha}_i$ is the i-th eigenvalue of the polarizability tensor. In this procedure, the polarizability of atoms in non-conjugated moieties of the molecules – such as side chains – are not scaled, in accordance with the original parametrization of the Thole model. The reference polarizability tensor was calculated from density functional theory via B3LYP/6-311+g(d,p).