JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Ultrafast Estimation of Electronic Couplings for Electron Transfer between π-Conjugated Organic Molecules

Fruzsina Gajdos,[†] Siim Valner,[†] Felix Hoffmann,[†,‡] Jacob Spencer,[†] Marian Breuer,[†] Adam Kubas,[†] Michel Dupuis,[§] and Jochen Blumberger*[,†]

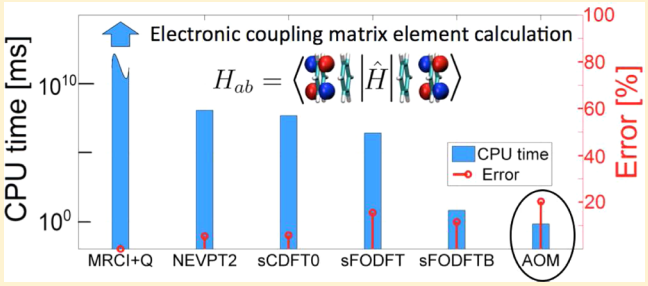
[†]Department of Physics and Astronomy, University College London, London WC1E 6BT, U.K.

[‡]Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, Universitätsstrasse 150, 44801 Bochum, Germany

[§]Pacific Northwest National Laboratory, Richland, Washington 99354, United States

Supporting Information

ABSTRACT: Simulation of charge transport in organic semiconducting materials requires the development of strategies for very fast yet accurate estimation of electronic coupling matrix elements for electron transfer between organic molecules (transfer integrals, H_{ab}). A well-known relation that is often exploited for this purpose is the approximately linear dependence of electronic coupling with respect to the overlap of the corresponding diabatic state wave functions for a given donor−acceptor pair. Here we show that a *single* such relation can be established for a large number of *different* π-conjugated organic molecules. In our computational scheme the overlap of the diabatic state wave function is simply estimated by the overlap of the highest singly occupied molecular orbital of donor and acceptor, projected on a minimum valence shell Slater-type orbital (STO) basis with optimized Slater decay coefficients. After calibration of the linear relation, the average error in H_{ab} as obtained from the STO orbital overlap is a factor of 1.9 with respect to wave function-theory validated DFT calculations for a diverse set of π-conjugated organic dimers including small arenes, arenes with S, N, and O heteroatoms, acenes, porphins, and buckyballs. The crucial advantage of the scheme is that the STO orbital overlap calculation is analytic. This leads to speedups of 6 orders of magnitude with respect to reference DFT calculations, with little loss of accuracy in the regime relevant to charge transport in organics.



1. INTRODUCTION

The simulation of charge transport (CT) in organic semiconductors (OSs) has attracted much interest in recent years.[1−7] This interest is fuelled by the many promising applications of OS materials in next generation microelectronics and energy production devices, e.g. organic light emitting diodes and flexible displays[8−10] and photovoltaic cells.[11−13] However, a serious limitation of present day OS materials is their rather limited charge mobility, in particular in their amorphous phase.[14] Here, modeling and simulation can offer valuable molecular insight into the factors determining charge transport such as microscopic structure, thermal fluctuations, and energy level alignment, the knowledge of which are all necessary for the design of improved materials.

The vast majority of simulations reported in the literature so far relied on the assumption that the charge carrier, electron or electron hole, is localized on a single molecule or molecular fragment and that transport occurs via charge hopping with rates given by Marcus or related theories.[1,2,15−18] Several groups including our own have noted that for most OS this view bears little theoretical justification.[3−6,19−21] The reorganization free energy or local electron−phonon coupling in these materials is often not larger than 0.2 eV, which is about an order of magnitude smaller than typical values for CT in aqueous media[22−25] and a factor of 3−6 smaller than for CT in proteins[26,27] On the other hand, electronic coupling between organic molecules, while on average an order of magnitude smaller, can become just as large as reorganization energy, in minimum energy or thermally accessible configurations (see e.g. refs 6 and 21, for the case of fullerenes). In this case the electronic interaction between the molecules can no longer be treated as a perturbation as is done in Marcus theory and the relevant states become delocalized. Unfortunately, neither is band theory applicable at relevant ambient temperatures, because the thermal fluctuations in OS materials are relatively large such that the mean free path is comparable to intermolecular lattice spacings.[28]

These considerations suggest that direct propagation schemes of the coupled electron−nuclear motion should give a more realistic picture of CT in OS materials,[4,19,29−33] in particular nonadiabatic molecular dynamics (NAMD) methods such as Ehrenfest and fewest switches surface hopping.[34−37] The latter two methods were developed primarily for

Received: June 17, 2014

   dx.doi.org/10.1021/ct500527v | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

simulation of photoexcited processes[37] (see e.g. refs 38−40 for application to relevant organic molecules) and have recently been adopted for simulation of thermal charge transport in DNA and proteins[38,41−44] and OS materials.[31−33] Due to their very high computational demand, NAMD simulation of CT in OS materials were typically carried out on relatively simple model systems, e.g. 1D or 2D chains of displaced harmonic oscillators described by model Hamiltonians.[19,30−32] Though a NAMD simulation of CT in a more detailed molecular model of a layer of organic molecules has also recently been reported.[33]

Studies with model Hamiltonians of reduced dimensions are undoubtedly useful for a theoretical understanding of CT in OSs, but it is not immediately clear if the findings carry over to real materials. A more realistic description would require the explicit NAMD simulation of at least several hundred organic molecules or fragments to allow for the case of strong excess charge delocalization in 3D and over time scales of at least ps-ns to account for a wide spectrum of nuclear relaxation times. Representing the electronic Hamiltonian in a localized site (or diabatic) basis, as is usually done, this would demand the calculation of many millions of diagonal elements (site energies) and off-diagonal elements (electronic coupling, H_{ab}) owing to the small integration time step for the nuclear motion (1 fs). Unfortunately, explicit electronic structure methods at the DFT or even at the semiempirical level (e.g., ZINDO), the latter typically taking seconds to minutes per matrix element, are not nearly fast enough for this purpose. Hence, to enable realistic NAMD simulation of OS materials, it is necessary to develop very fast yet reasonably accurate methods for estimation of matrix elements.

Here focus is placed on an efficient estimation of the off-diagonal matrix elements by investigating the approximation

$$H_{ab} = CS_{ab} \tag{1}$$

where S_{ab} is the overlap integral between two localized electronic states, and C is a constant (see section 2 for a more precise definition of these quantities). Of course, this is an old idea that dates back to the early days of extended Hueckel theory,[45] where a similar relation is used to approximate the resonance integral. The difference is that here H_{ab} is the coupling matrix element between two diabatic state electronic wave functions localized on two different molecules, whereas the resonance integral refers to the electronic interaction of atomic orbitals centered at nearest neighbor atoms on the same molecule. The relation eq 1 has been tested a number of times, and it was found to give an excellent approximation for ET in donor−acceptor pairs of DNA bases,[46] acenes,[15] and fullerenes,[20] with slopes determined to be in the order of C = 14 eV. It remains unclear, however, if C needs to be reparametrized for every single donor−acceptor system or if eq 1 is indeed of more general validity, in which case a large class of OS materials can be described by a single value for C.

Clearly, eq 1 is only useful if the overlap S_{ab} can be calculated in a highly efficient way. For this purpose we replace the diabatic state wave function by the highest singly occupied molecular orbital (SOMO) and represent the latter by a minimum Slater type orbital (STO) basis. This allows for an analytic calculation of S_{ab} in terms of atomic orbital overlaps between a relatively small number of orbitals. The slope C is then obtained by fitting the analytically calculated approximate overlaps against reference H_{ab} values obtained from DFT calculations for a training set of π-conjugated donor−acceptor complexes. The parametrized relation eq 1 is then tested on various sets of π-conjugated molecules outside the training set. We find that eq 1 with a single parameter C can describe rather well the electronic coupling of all molecules investigated, but only if the exponents of the minimum STO basis set are properly chosen. This leads to a speedup of about 6 orders of magnitude compared to DFT calculations, with little loss of accuracy.

The remainder of this paper is organized as follows. In section 2 we define all quantities of interest and describe in detail the projection of the SOMO on the minimum STO basis and the analytic calculation of S_{ab}. In section 3 we define the training set used to establish the correlation eq 1 and the test sets for investigation of the generality of this relation as well as the reference DFT calculations. In section 4 we first discuss the accuracy of the DFT H_{ab} values by comparing them to the results of correlated ab initio calculations, specifically n-electron valence state perturbation theory at second order (NEVPT2) level of theory.[47−49] We then present the results for the correlation eq 1 and discuss the gain of efficiency of this scheme. This work is concluded in section 5.

2. THEORY

The goal is to establish a linear correlation eq 1 between electronic coupling H_{ab} of the two diabatic electronic wave functions for initial and final electron transfer (ET) states, Ψ_a and Ψ_b, respectively

$$H_{ab} = \langle \Psi_a | \mathcal{H} | \Psi_b \rangle \tag{2}$$

and the corresponding overlap

$$S_{ab} = \langle \Psi_a | \Psi_b \rangle. \tag{3}$$

The correlation should be valid for any given geometry of a donor−acceptor pair and for a wide range of π-conjugated donor and acceptor molecules. The calculation of the exact overlap is computationally demanding as it requires the knowledge of the diabatic wave functions (see e.g. refs 50−53 where wave functions are approximated by Kohn−Sham determinants obtained from charge constrained DFT calculations). In the following we make a number of approximations to speed up the calculation of S_{ab} in eq 3. First, we assume that ET is mediated by only two orbitals, the SOMO of the isolated reduced donor fragment, ϕ'^{N}_{D}, and the SOMO of the isolated reduced acceptor fragment, ϕ'^{N}_{A}, respectively (notation as in refs 21 and 54).

$$S_{ab} = \langle \phi'^{N}_{D} | \phi'^{N}_{A} \rangle \tag{4}$$

The assumption that the SOMO orbital of the isolated fragments can be used to describe the ET and that lower lying orbitals do not significantly participate in the ET is often made in the literature, specifically in the fragment orbital DFT (FODFT) method or similar approaches for calculation of H_{ab}.[15,21,22,46,54,55] Recent benchmark calculations showed that this is a remarkably good approximation at least for π-conjugated organic molecules.[54] To enable a very fast estimation of the overlap eq 4 for many different donor−acceptor configurations, we carry out an explicit electronic structure calculation for donor and acceptor only once and project the corresponding SOMO orbitals ϕ'^{N}_{D} and ϕ'^{N}_{A} on a minimum Slater type orbital (STO) basis for the valence states

$$|\phi'^{N}_{D}\rangle \approx |\tilde{\phi}'^{N}_{D}\rangle = \sum_{k} c_k |\chi_k\rangle \tag{5}$$

dx.doi.org/10.1021/ct500527v | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

where $|\chi_k\rangle$ is the kth STO orbital, and c_k is the corresponding expansion coefficient. The index k runs over all STO orbitals on a given donor atom and over all donor atoms. For H, only the 1s orbital is included, and for second row atoms only the 2s, 2p_x, 2p_y, 2p_z orbitals are included. The expansion coefficients are given by $c_k = \langle \sum_l S^{-1}_{lk} \chi_l | \phi'^N_D \rangle$, where S^{-1}_{lk} are the elements of the inverse of the atomic overlap matrix $S_{lk} = \langle \chi_l | \chi_k \rangle$. The Slater coefficient μ_n of the radial part of the STOs

$$R_{n,l}(r) = N_{n,l} r^{n-1} e^{-\mu_n r/a_H} \tag{6}$$

is chosen so as to maximize the completeness of the projection, $p = \max_{\mu_n} \langle \tilde{\phi}'^N_D | \tilde{\phi}'^N_D \rangle \leq 1$. (In eq 6, n is the principal quantum number, l is the angular quantum number, a_H is the Bohr radius, and $N_{n,l}$ is a normalization factor.) The reconstruction of the SOMO orbitals in terms of a minimum Slater basis allows for an ultrafast analytic estimation of the overlap for a given donor−acceptor configuration, as further detailed below. For the atoms participating in π-conjugation we project the p orbitals in the π and σ contribution

$$|\tilde{\phi}'^N_D\rangle = \sum_{i\in D}^{\text{atoms}} c_{s,i}|s_i\rangle + c_{p\pi,i}|p_{\pi,i}\rangle + c_{p\sigma 1,i}|p_{\sigma 1,i}\rangle + c_{p\sigma 2,i}|p_{\sigma 2,i}\rangle \tag{7}$$

where p_π and $p_{\sigma 1}$, $p_{\sigma 2}$ are p orbitals perpendicular and parallel, respectively, to the plane of π-conjugation. The projection coefficients are given by

$$c_{p\pi,i} = \sum_{j=x,y,z} c_{pj,i} \langle p_{\pi,i} | p_{j,i} \rangle \tag{8}$$

and similarly for $c_{p\sigma 1,i}$ and $c_{p\sigma 2,i}$. As shown in section 4, the s and p_σ contributions in eq 7 can be neglected for the molecules studied here. After renormalization, one obtains the approximate SOMO orbital, $\bar{\phi}'^N_D$, in terms of the p_π contribution only

$$|\tilde{\phi}'^N_D\rangle \approx |\bar{\phi}'^N_D\rangle = \sum_{i\in D}^{\text{atoms}} c_{p\pi,i}|p_{\pi,i}\rangle \tag{9}$$

where $\langle \bar{\phi}'^N_D | \bar{\phi}'^N_D \rangle = 1$. A similar projection on the p_π orbitals is carried out for the SOMO of the acceptor. The overlap eq 4 can now be approximated by atomic overlaps between STO orbitals on the donor and acceptor fragments

$$S_{ab} \approx \bar{S}_{ab} = \langle \bar{\phi}'^N_D | \bar{\phi}'^N_A \rangle = \sum_{i\in D}^{\text{atoms}} \sum_{j\in A}^{\text{atoms}} c^*_{p\pi,i} c_{p\pi,j} \langle p_{\pi,i} | p_{\pi,j} \rangle \tag{10}$$

To calculate the atomic overlap between some arbitrarily oriented p_π orbital on atom i and an arbitrarily oriented p_π orbital on atoms j, a local coordinate system is introduced on atom i with the z axis pointing along the direction of the atoms i and j and the x and y axes pointing in orthogonal directions. The axes of the local coordinate system on atom j are chosen parallel to ones on atom i. The p_π orbital on atom i is then projected on p orbitals pointing along the axes of the local coordinate system

$$|p_{\pi,i}\rangle = c_{x,i}|p_{x,i}\rangle + c_{y,i}|p_{y,i}\rangle + c_{z,i}|p_{z,i}\rangle \tag{11}$$

and similarly for p_π on atom j. The overlaps between the STO orbitals on the RHS of eq 10 are then given by

$$\langle p_{\pi,i}|p_{\pi,j}\rangle = (c^*_{x,i}c_{x,j} + c^*_{y,i}c_{y,j}) S_{p\pi}(r, \mu_{p,i}, \mu_{p,j}) + c^*_{z,i}c_{z,j} S_{p\sigma}(r, \mu_{p,i}, \mu_{p,j}) \tag{12}$$

where $S_{p\sigma}$ is the overlap of the two coaxial p_z orbitals on atoms i and j, depending only on the distance between the atoms (r) and the Slater coefficients (μ), $S_{p\sigma}(r,\mu_{p,i},\mu_{p,j}) = \langle p_{z,i}|p_{z,j}\rangle$, and $S_{p\pi}$ is the overlap of the two parallel p_x (p_y) orbitals on atoms i and j, $S_{p\pi}(r,\mu_{p,i},\mu_{p,j}) = \langle p_{x,i}|p_{x,j}\rangle = \langle p_{y,i}|p_{y,j}\rangle$. The overlaps $S_{p\sigma}$ and $S_{p\pi}$ are computed analytically using the formulas provided by Mulliken.[56]

3. COMPUTATIONAL DETAILS

To establish the correlation eq 1, we have chosen a training set of π-conjugated molecular dimers. The performance of the linear correlation is then tested using six test sets, each of which includes dimer structures outside the training set. The molecules in the test and training sets are shown in Figure 1

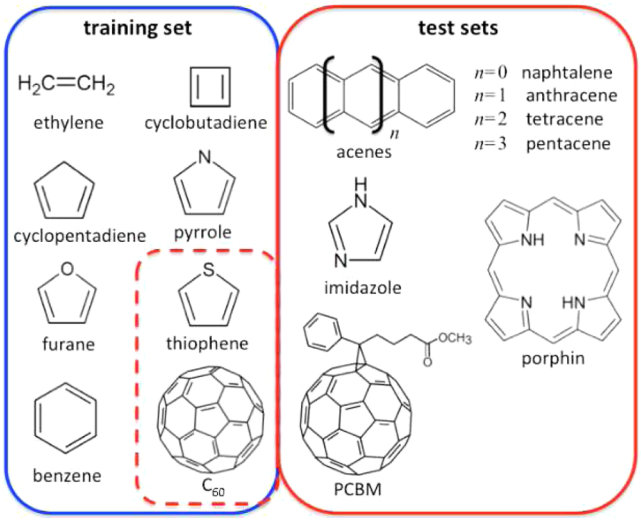


Figure 1. Training set of π-conjugated molecules used to establish the linear correlation eq 1 and test set of molecules used to test the linear correlation. The thiophene dimers used in the training and test sets are stacked and randomly orientated, respectively. The C_{60} dimers in training and test set differ in their respective distance and orientation. See section 3 for details.

and described in detail further below. As we would like to establish the correlation eq 1 for *any* donor−acceptor geometry, we are free to choose the geometry of each monomer for the calculation of the electronic coupling matrix element. Here the geometries of the neutral monomers are used. This differs for instance from ET rate calculations, where the coupling matrix element needs to be calculated for the transition state geometry of the donor−acceptor complex.[57] In cases where the SOMO orbitals are degenerate (benzene and C_{60}) we choose to compute the coupling matrix element for only one specific pair of SOMO orbitals on donor and acceptor. The calculation of the overlap in terms of the minimum STO basis is then carried out on the same pair of SOMO orbitals.

Training Set. The training set is comprised of a selection of electron hole transfer reactions between homodimers, taken from the HAB11 database,[54] $M1^+$-M2 → M1-$M2^+$, where M1 = M2 are ethylene, cyclobutadiene, cyclopentadiene, pyrrole, furane, thiophene, and benzene, see Figure 1. The geometries are in a stacked configuration and separated by 3.5, 4.0, 4.5, and

dx.doi.org/10.1021/ct500527v | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

5.0 Å. These dimer structures are referred to as the "HAB7" set. We have chosen these dimers because H_{ab} values for electron hole transfer in these systems have been recently computed using high level ab initio and DFT calculations. The H_{ab} values for these dimers span only a range of 50−500 meV, which is why we have included an additional set of charge shift reactions in the training set, ET in the C_{60} dimer anion for three different orientations of the C_{60} monomers with their centers of mass separated by 9, 10, 11, 12, 13, and 14 Å. These 15 structures are referred to as the "C_{60}" set, and the combined HAB7+C_{60} sets are referred to as the training set. The total number of charge shift reactions included in the training set is 43, and they span H_{ab} values from 0.1 to 500 meV.

Test Sets. The "thiophene" test set consists of 15 randomly oriented thiophene dimer cations (note, the thiophene dimers included in the training set are all stacked), the "acenes" test set consists of stacked naphthalene, anthracene, tetracene, and pentacene dimer cations, the "C_{60}^{r}" test set consists of 20 C_{60} dimer anions, randomly chosen from the 27,000 rotamers with equidistant C_{60} separation that were investigated in ref 21. (Note, these rotamers differ from the structures chosen for the test set.) The "PCBM" test set consists of 38 PCBM dimer anion structures obtained from the monoclinic and triclinic crystal structures[58] and optimized in ref 6. The "imidazole" test set consists of four stacked imidazole dimer cation structures included in the HAB11 database with separation distance 3.5, 4.0, 4.5, and 5.0 Å. Finally, the "porphin" test set consists of 5 stacked porphin dimer anion structures separated by 3, 4, 5, 6, and 8 Å.

Electronic Coupling Matrix Elements (H_{ab}). H_{ab} values for C_{60} and porphin dimer anions were calculated according to the FODFT method as implemented in the CPMD program package.[59] For all other dimers in the training and test sets we took the results from previous calculations.[54,60] The ionic cores are described by Goedecker-Hutter pseudo potentials, and the valence electron states are expanded in plane waves with a reciprocal space plane wave cutoff of 90 Ry. For C_{60} we used the PBE functional, and for the porphin system we used the PBE0 functional. The structures of the monomers were obtained from geometry optimization in the neutral state, and they were used to construct the dimer structures described above. The dimers were centered in the simulation box, and a vacuum of 4 Å was applied in each dimension. Increase of the vacuum layer did not significantly change H_{ab} values.

The H_{ab} values reported here are for the nonorthogonal orbital pair comprised of the SOMO of the reduced donor, ϕ'^{N}_{D}, and the SOMO of the reduced acceptor, ϕ'^{N}_{D}. This differs from our previous definition, where FODFT couplings were reported for the SOMO orbital pair obtained after Lowdin orthogonalization.[54] We found, however, that the difference in H_{ab} between nonorthogonal and orthogonal SOMO orbitals is very small, and the mean unsigned relative deviation is less than 1% for the dimers in the training and test sets. The FODFT couplings between the nonorthogonal orbitals were then scaled by a factor of 1.348 as prescribed in ref 54. The final coupling values are referred to as sFODFT (s for scaled), and they are summarized in Table 2.

4. RESULTS

Validation of H_{ab} Reference Values. At first we investigate the quality of the H_{ab} values obtained from FODFT, that will be used for the correlation eq 1. Electronic couplings are small energies (typically sub-meV to several 100 meV), which raises the question of the numerical accuracy that can be achieved. We have carried out careful checks in this regard and found that FODFT H_{ab} values are converged to 0.1 meV at relatively moderate reciprocal space plane wave cutoffs of 90 Ry compared to calculations at 210 Ry, see Table 1. The reason for this excellent convergence behavior is that electronic coupling matrix elements are determined by the tails of the frontier orbitals, which are already well represented at moderate cutoffs (as opposed to total energy). This is in contrast to localized basis set calculation where very large basis sets need to be employed to obtain similarly good convergence.[54,61] Next we compare the coupling matrix elements obtained with the plane wave FODFT method with the results of high-level ab initio calculations at the NEVPT2 level of theory (see Figure 2). The latter method was shown to give excellent agreement with MRCI and FCI calculations on small dimer cations.[54] We find that after application of a uniform scaling factor of 1.348, the agreement between sFODFT (s for scaled) and NEVPT2 is very good with MUREs of only 14%, 27%, and 13% for the HAB7 training set and thiophene and imidazole test sets, respectively. These errors translate into errors in the nonadiabatic electron transfer (ET) rate ($k_{ET} \propto |H_{ab}|^2$ [57]) of factors of 1.3, 1.6, and 1.3, respectively, i.e. negligible for all practical purposes.

Table 1. Dependence of Electronic Coupling Matrix Elements (H_{ab}) on the Reciprocal Space Plane Wave Cutoff[a]

cutoff (Ry)	H_{ab} (meV)		
90	3.763	4.462	2.165
120	3.662	4.349	2.151
150	3.664	4.349	2.150
210	3.667	4.351	2.149

[a]Calculations are carried out for three orientations of a C_{60} dimer anion using the FODFT method.[21,54,59]

Table 2. Electronic Coupling Matrix Elements (H_{ab}) from sFODFT Calculations for the C_{60} Dimer Anion in Three Different Orientations and for the Stacked Porphin Dimer Anion (in meV)[a]

C_{60}				porphin	
d^b (Å)	H_{ab} (meV)			d^b (Å)	H_{ab} (meV)
9	26.4	23.3	23.2	3	858.3
10	9.9	6.7	5.8	4	182.7
11	3.2	1.9	1.4	5	43.4
12	1.0	0.5	0.3	6	11.6
13	0.3	0.1		8	1.4
14	0.1				

[a]The atomic coordinates of the dimers are given in the Supporting Information. [b]Distance between the center of mass of donor and acceptor molecule.

Correlation H_{ab} vs S_{ab}. The correlation between H_{ab} and the exact overlap between the SOMO orbitals, S_{ab}, eq 4, is shown in Figure 3. With "exact" we mean that the overlap is obtained from real space grid integration of the SOMO orbitals, not from the projection of the SOMO on the minimum Slater basis. Indeed, we find a rather good linear correlation over 4 orders of magnitude in H_{ab}, the range covered by the data. Minimizing the exponentiated root-mean-square logarithmic error (ERMSLE) for the training set (HAB7+C60, symbols in blue)

dx.doi.org/10.1021/ct500527v | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

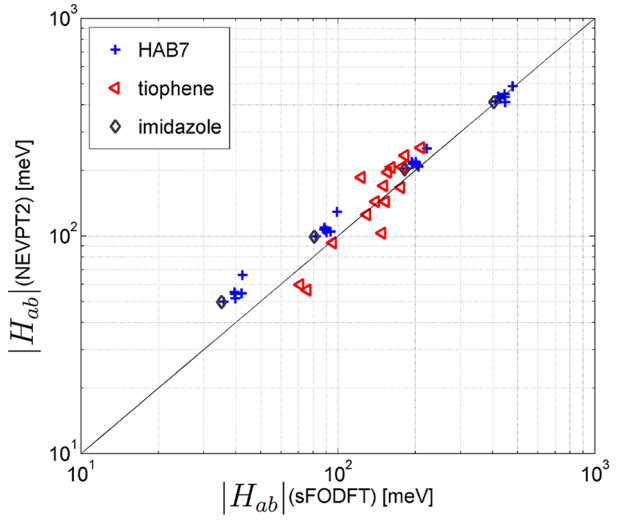


Figure 2. Correlation between electronic coupling matrix elements obtained from high-level ab initio calculations (NEVPT2) and sFODFT. Data are taken from ref 54.

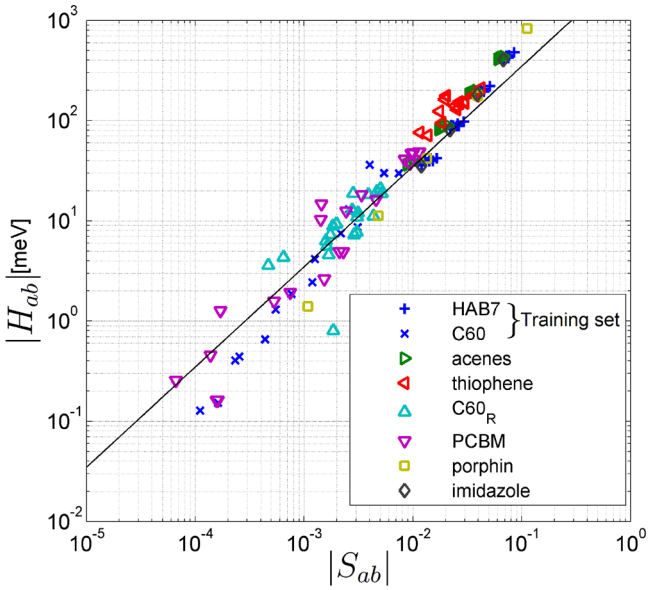


Figure 3. Correlation between electronic coupling matrix element (H_{ab}) from sFODFT and exact overlap (S_{ab}) between SOMO orbitals of donor and acceptor. The straight line is a fit to the training set (data points in blue) according to eq 13.

$$\mathrm{ERMSLE} = \exp\left[\left\langle\left(\ln\frac{|H_{ab}|}{C|S_{ab}|}\right)^2\right\rangle^{1/2}\right] \tag{13}$$

we obtain a straight line fit through the origin, $|H_{ab}| = C|S_{ab}|$, with a slope C = 3.498 eV (R^2 = 0.959, ERMSLE = 1.6).

Projection of SOMO on Minimum STO Basis. The SOMO was projected on a minimum valence shell STO basis, 1*s* for H; 2*s*, 2p_x, 2p_y, and 2p_z for C, N, and O; and 3*s*, 3p_x, 3p_y, and 3p_z for S. The Slater decay coefficients used are summarized in Table 3. They are the standard Clementi-Raimondi values[62] except for the 2*p* orbitals of C and N (only these two atom types exhibited significant SOMO amplitude in the molecules of the training and test sets). We optimized the Slater decay coefficient of the 2*p* orbital of C by maximizing the completeness of the projection of the SOMO of C_{60}. This gives a value of μ_{2p} = 1.3120 au^{-1}. For N we obtained μ_{2p} = 1.7000 au^{-1} by maximizing the completeness of the SOMO of imidazole using the above value for C. These optimized decay coefficients are somewhat smaller than the Clementi-Raimondi values, μ_{2p} = 1.5679 au^{-1} for C and 1.9170 au^{-1} for N.[62] The projection had a completeness of $p = \max_{\mu_n}\langle\tilde{\phi}'^{N}_{D}|\tilde{\phi}'^{N}_{D}\rangle$ = 0.978 for the molecules in the training and test sets, which did not significantly increase upon addition of further STO orbitals of higher principal quantum number. The projections on the *s* states are generally very small, $\sum_{i\in D}^{\text{atoms}} c^*_{s,i}c_{s,i} \le 10^{-6}$ (10^{-4} for C_{60}) and are neglected for the overlap calculation without loss of accuracy. Moreover, after separation of the *p* orbitals in π and σ contributions, the p_σ contributions to the SOMO orbital are very small, $\sum_{i\in D}^{\text{atoms}}\sum_{k,l=1}^{2} c^*_{p\sigma k,i}c_{p\sigma l,i} \le 0.02$, and are also neglected for the overlap calculation.

Table 3. Summary of Slater Decay Coefficients Used for the Valence Shell Minimum STO Basis (μ in au^{-1})[d]

	H	C	N	O	S
μ_{1s}[a]	1.0000				
μ_{2s}[a]		1.6083	1.9237	2.2458	
μ_{2p}[a]		1.3120[c]	1.7000[c]	2.2266	
μ_{3s}[a]					2.1223
μ_{3p}[a]					1.8273
$\bar{\mu}_{2p}$[b]		1.0000[c]	1.5000[c]		

[a]For projection of SOMO orbitals. [b]For calculation of $\bar{S}_{ab}$. [c]Optimized as explained in Section 4. [d]Data taken from ref 62 except where indicated.

Calculation of $\bar{S}_{ab}$. The overlap is calculated using only the projected p_π contribution of C and of N if present. The Slater coefficients used for the projection may not be the best choice to establish the correlation between H_{ab} and $\bar{S}_{ab}$. While for a (nearly) complete projection onto the STO orbitals it is important to reproduce the shape of the SOMO close to the nuclei, for the calculation of the overlap it is important to reproduce the shape of the SOMO far away from the nuclei, in the region between the donor and acceptor. Therefore, we defined a second Slater decay coefficient for C, $\bar{\mu}_{2p}$, that is chosen to maximize the correlation between H_{ab} (sFODFT) and $\bar{S}_{ab}$, while keeping the projection coefficients fixed to the values obtained with μ_{2p}. This is done by minimizing the ERMSLE (eq 13, S_{ab} replaced by $\bar{S}_{ab}$) with respect to $\bar{\mu}_{2p}$ for the dimers in the training set (HAB7+C_{60}, blue symbols in Figure 4). For N an optimized value $\bar{\mu}_{N,2p}$ = 1.500 au^{-1} was obtained by minimizing the ERMSLE for the imidazole and porphine test sets using the optimized value for C. The modified Slater coefficients $\bar{\mu}_{2p}$ are summarized in Table 3. They were used for the calculation of $\bar{S}_{ab}$ for all dimers in the training and test sets.

Correlation H_{ab} vs $\bar{S}_{ab}$. The correlation between H_{ab} and $\bar{S}_{ab}$ is shown in Figure 4. The linear correlation for the training set (symbols in blue, R^2 = 0.974, ERMSLE = 1.4) is even slightly better than for H_{ab} versus S_{ab} indicating that the approximate overlap is an even better descriptor than the "exact" overlap, provided that the Slater decay coefficients are optimized (see above). The linear straight line fit through the origin, $|H_{ab}| = \bar{C}|\bar{S}_{ab}|$ has now a smaller slope, $\bar{C}$ = 1.819 eV, than for the fit against S_{ab}. This is consistent with a slope of less than 1 for S_{ab} versus $\bar{S}_{ab}$ (Figure S, R^2 = 0.986, slope = 0.762). We would like to point out that $\bar{C}$ is very sensitive to the Slater decay coefficient used and strongly decreases with decreasing $\bar{\mu}_{2p}$. Hence, it is important to use the same set of Slater

dx.doi.org/10.1021/ct500527v | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

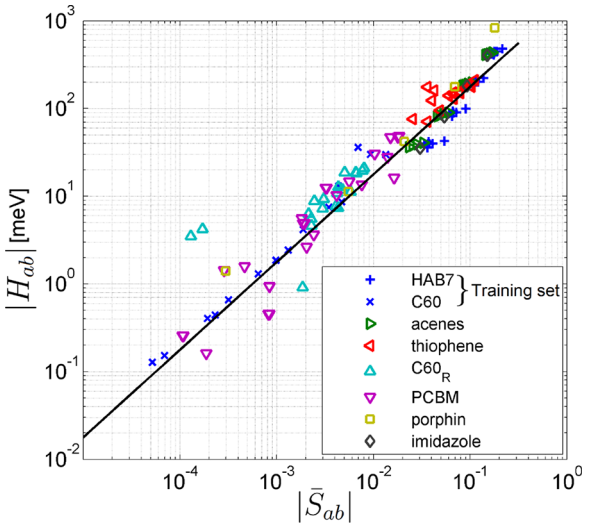


Figure 4. Correlation between electronic coupling matrix element (H_{ab}) from sFODFT and overlap between the SOMO orbitals of donor and acceptor represented in a minimum Slater valence basis ($\bar{S}_{ab}$). The straight line is a fit to the training set (data points in blue) according to eq 13, with S_{ab} replaced by $\bar{S}_{ab}$.

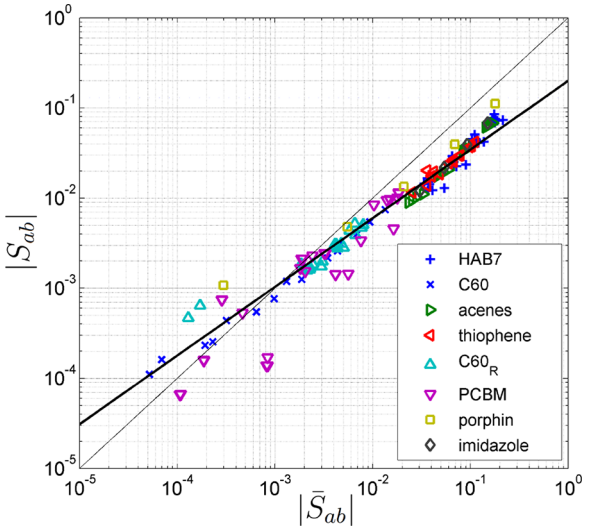


Figure 5. Correlation between exact and approximate orbital overlap. The straight line is a fit to the training set (data points in blue), similar to eq 13. The dashed line has a slope of 1 and is a guide to the eye.

coefficients for calibration of the linear relation and subsequent applications. In this respect we note that significantly larger slopes were reported previously, |C| = 19.5 eV for DNA bases[46] and |C| = 14.1 eV for PCBM.[20] Most likely, this is a consequence of the small basis set used in these calculations (3-21G[46]) compared to the present Slater type orbital parametrization of plane wave DFT calculations.

The transferability of the linear correlation from the training set to the π-conjugated organic molecules contained in the test sets (acenes, randomly orientated thiophenes and C_{60} (C^r_{60}), PCBM, porphin, and imidazole) is illustrated in Figure 4. We find that for larger overlaps ($|\bar{S}_{ab}| = 5 \times 10^{-3} - 10^{-1}$) the scatter is again reasonably small but for smaller overlaps ($|\bar{S}_{ab}| \leq 5 \times 10^{-3}$) somewhat larger. The ERMSLE and maximum errors (MAX) for these sets are summarized in Table 4. Overall, we find very good transferability, with ERMSLEs ranging from 1.3 for acenes to 2.7 for C^r_{60}.

Table 4. Exponentiated Root Mean Square Logarithmic Error (ERMSLE) and Maximum Error (MAX) for Electronic Couplings, H_{ab}, Obtained from the Linear Fit Eq 1 to the Training Set

	H_{ab} vs S_{ab}[a]		H_{ab} vs $\bar{S}_{ab}$[b]	
set	ERMSLE[c]	MAX[d]	ERMSLE[e]	MAX[f]
training set	1.6	3.7	1.4	2.9
test sets[g]	1.6	8.2	1.9	15.7
acenes	1.5	1.9	1.3	1.6
thiophene	1.7	2.5	1.5	2.8
C^r_{60}	1.8	8.2	2.7	15.7
PCBM	1.7	3.5	2.5	8.8
porphin	1.8	2.7	1.9	2.6
imidazole	1.4	1.7	1.4	1.5

[a]Exact overlap obtained from grid integration. [b]Overlap obtained after projection on a valence shell minimum STO basis. [c]Equation 13, data shown in Figure 3. [d]MAX = $\exp[\max_i\{|\ln|H_{ab,i}/(CS_{ab,i})|\|\}]$. [e]Equation 13, S_{ab} and C replaced by $\bar{S}_{ab}$ and $\bar{C}$, respectively. Data shown in Figure 4. [f]As in footnote *d* with S_{ab} and C replaced by $\bar{S}_{ab}$ and $\bar{C}$, respectively. [g]Average over all test sets (acenes, thiophene, C^r_{60}, PCBM, porphin, imidazole).

5. DISCUSSION

Accuracy of the Method. We would like to discuss the accuracy of the current method, that we denote "Analytic Overlap Method" (AOM), in the context of ET rates and NAMD simulations. We find that the error in H_{ab} (as defined in eq 13), averaged over all test sets spanning values of 0.1−500 meV, is small, a factor of 1.9. This corresponds to an error of a factor of approximately 3 in the nonadiabatic ET rate ($k_{ET} \propto |H_{ab}|^2$ [57]) . Hence, the electronic couplings obtained with AOM can be judged as *on average* chemically accurate, if the latter refers to an accuracy of 1 order of magnitude in chemical reaction rates. However, the relative error of the method is not uniform over the entire range of coupling values covered in this study, as we discuss in the following.

We find that the relative error of AOM is small for all data points in the large electronic coupling regime 5−500 meV, but the error can be substantial in the small electronic coupling regime <5 meV (see Figure 4). For instance, for a configuration of the C^r_{60} test set the coupling value obtained from sFODFT is 3 meV, whereas the coupling value obtained from AOM is 0.2 meV (see outlier in Figure 4). This maximum error of a factor of 15.7 (MAX, defined in Table 4) would correspond to an underestimation of the nonadiabatic ET rate by 2 orders of magnitude. While this error appears to be large, its impact on the results of ET/NAMD simulation for electron transport in real materials is expected to be relatively small. First, electron transport in the condensed phase (crystal, thin film) is determined by the strongest electronic coupling of an organic molecule with its nearest neighbors. Thus, it is most important to be able to compute the larger coupling values accurately rather than the smaller ones. Second, the strongest electronic couplings between a molecule and its neighbors in a given geometry of an OS material are typically a few 10 to a few 100 meV, which is the range where AOM is accurate. Considering for instance the C^r_{60} test set, which is comprised of configurations representative of pairs of C_{60} molecules in the

dx.doi.org/10.1021/ct500527v | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

fcc-C_{60} crystal, the fraction of configurations where the electronic coupling is less than 5 meV and the error of AOM potentially large is no more than about 15% (see Figure 4). Hence, it is highly improbable that all coupling values for a C_{60} molecule and its 12 nearest neighbors in the fcc lattice are small and the error in the transport determining rate large. Therefore, we conclude that AOM gives sufficiently accurate results in the regime that is relevant to electron transport in OS materials and hence should be well suited for ET/NAMD simulation in the condensed phase.

Interestingly, the trend that the relative error becomes larger for smaller coupling values is also observed for the correlation of H_{ab} with the exact overlap S_{ab}, see Figure 3. For the above example, the MAX error for the C_{60}^{r} test set is 8.2 compared to 15.7 with respect to $\overline{S}_{ab}$ (see Table 4). This indicates that the larger relative errors observed in the small coupling regime <5 meV is not primarily due to the approximation of the exact overlap S_{ab} by the minimum STO overlap $\overline{S}_{ab}$ but due to inherent deviations of the relation between H_{ab} and S_{ab} from linearity. In this respect we note that deviations from the linear relationship are also expected for the very large overlap/coupling regime >500 meV, as in fact it was shown that $H_{ab} \propto S_{ab}/(1-S_{ab}^2)$.[57,63,64] However, maximum coupling values in organic dimers at van der Waals distance and optimum π-stacking are typically <500 meV (see e.g. ref 54), and $S_{ab}/(1-S_{ab}^2) \approx S_{ab}$ to a good approximation. Therefore, we conclude that neither the very low nor the very high coupling regime are relevant for ET in organic materials, which again justifies the use of the linear relation eq 1.

Speedup of H_{ab} Calculation. In Figure 6 we compare typical CPU timings for the calculation of H_{ab} for small organic molecules using high-level ab initio methods (MRCI+Q, NEVPT2), scaled (s) DFT methods (constrained DFT with the PBE functional (sCDFT0) and sFODFT with the PBE functional), fragment orbital density functional tight binding (sFODFTB), and AOM. MRCI+Q calculations could only be carried out for ethylene in a reasonable amount of time (hence the gray arrow for benzene and naphthalene), and NEVPT2 calculations could only be carried out for ethylene and benzene. The speedup obtained using AOM is about 6 orders of magnitude compared to the most efficient DFT method (sFODFT) and about 1 order of magnitude compared to sFODFTB. We note that the DFT calculations were carried out with a plane wave code,[59] which for small molecules may be less efficient than codes employing atom centered basis functions, depending on the size of the localized basis set used. Hence the speedup of AOM may be somewhat smaller compared to the most efficient DFT codes. Intriguingly, the mean relative unsigned error (MRUE) of the AOM method, assessed by averaging over ethylene, cyclobutadiene, cylocpentadiene, and pyrrole using MRCI+Q values as a reference (20%), is not much greater than for the scaled DFT and DFTB methods (5−15%, solid bars in red).

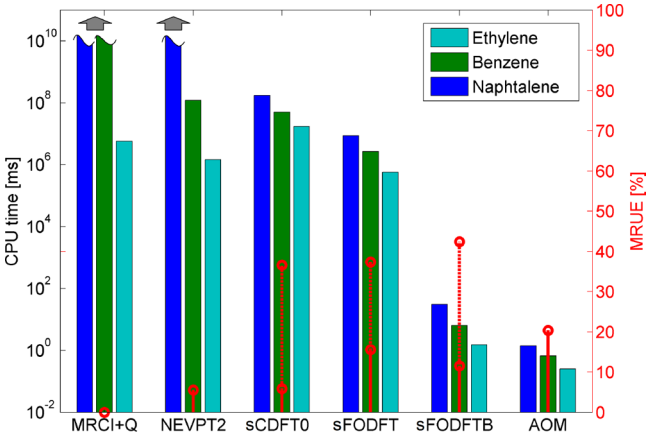


Figure 6. CPU time (bars) for the calculation of H_{ab} for homodimer cations of ethylene, benzene, and naphthalene and the mean relative unsigned error (MRUE, red circles) for dimers of small π-conjugated molecules relative to MRCI+Q. High-level ab initio methods (MRCI+Q, NEVPT2), scaled DFT methods (sCDFT0, sFODFT), a scaled density functional tight binding method (sFODFTB), and the atomic overlap method (AOM) developed in this work are compared. For the DFT and DFTB methods the error after scaling is given in solid lines, and the error before scaling is given in dashed lines.

6. CONCLUSION

We have shown that the linear relation eq 1 gives a good approximation to electronic coupling matrix elements between π-conjugated organic molecules in the range relevant to electron transport in the condensed phase (5−500 meV). While such relations have been established before for a *single* donor−acceptor pair, we find that the linear relation holds surprisingly well for a very diverse set of donor−acceptor pairs including arenes, arenes with S, N, and O heteroatoms, and buckyballs in different random orientations and distances. Most importantly, the good linear correlation is retained or can be even improved if the overlap is calculated in terms of a minimum STO Slater basis with optimized Slater decay coefficients. This allows us to estimate electronic couplings analytically leading to speedups of about 6 orders of magnitude compared to standard DFT calculations for molecules typically used in molecular OSs. Moreover, it is trivial to make the scheme linearly scaling through introduction of a cutoff for the calculation of atomic orbital pair overlaps. Together with a fast estimator for site energies using e.g. polarizable or non-polarizable force fields, this scheme provides the basis for a fast yet reasonably accurate estimation of the electronic Hamiltonian that can be used in NAMD simulations of charge transport in OS materials.

■ ASSOCIATED CONTENT

Supporting Information

The atomic coordinates of the dimer structures used for the calculations reported in Table 2. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: j.blumberger@ucl.ac.uk.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We would like to thank Alexander Heck, Karlsruhe Institute of Technology, for providing the timing of FODFTB calculations. F.G. was supported by an IMPACT Ph.D. studentship cosponsored by Pacific Northwest National Laboratory (PNNL) and University College London, J.S. was supported by an IMPACT PhD studentship cosponsored by University College London and the Department of Physics and

Astronomy, A.K. was supported by EPSRC grant EP/J015571/1 and M.D. by the U.S. Department of Energy (DOE), Office of Basic Energy Sciences (BES), Division of Chemical Sciences, Geosciences and Biosciences. PNNL is a multiprogram national laboratory operated for DOE by Battelle. J. B. acknowledges The Royal Society for a University Research Fellowship. Electronic structure calculations were carried out on HECToR (Edinburgh), access to which was granted through the Materials Chemistry Consortium (EPSRC grants EP/F067496 and EP/L000202). The authors also acknowledge the use of the UCL High Performance Computing Facility "Legion", and associated support services, for completion of this work.

REFERENCES

(1) Coropceanu, V.; Cornil, J.; da Silva, D. A.; Olivier, Y.; Silbey, R.; Bredas, J.-L. *Chem. Rev.* **2007**, *107*, 926.
(2) Nelson, J.; Kwiatkowski, J. J.; Kirkpatrick, J.; Frost, J. M. *Acc. Chem. Res.* **2009**, *42*, 1768.
(3) Troisi, A. *Chem. Soc. Rev.* **2011**, *40*, 2347.
(4) Fratini, S.; Ciuchi, S. *Phys. Rev. Lett.* **2009**, *103*, 266601.
(5) Ortmann, F.; Bechstedt, F.; Hannewald, K. *New J. Phys.* **2010**, *12*, 023011.
(6) Gajdos, F.; Oberhofer, H.; Dupuis, M.; Blumberger, J. *J. Phys. Chem. Lett.* **2013**, *4*, 1012.
(7) Raos, G.; Casalegno, M.; Ide, J. *J. Chem. Theory Comput.* **2014**, *10*, 364.
(8) Ho, P. K. H.; Kim, J. S.; Burroughes, J. H.; Becker, H.; Li, S. F. Y.; Brown, T. M.; Cacialli, F.; Friend, R. H. *Nature* **2000**, *404*, 481.
(9) Cacialli, F.; Wilson, J. S.; Michels, J. J.; Daniel, C.; Silva, C.; Friend, R. H.; Severin, N.; Samori, P.; Rabe, J. P.; O'Connell, M. J.; Taylor, P. N.; Anderson, H. L. *Nat. Mater.* **2002**, *1*, 160.
(10) Fenwick, O.; Credgington, D.; Hammiche, A.; Lazzerini, G. M.; Silberberg, Y.; Cacialli, F. *Nat. Nanotechnol.* **2009**, *4*, 664.
(11) Kippelen, B.; Bredas, J.-L. *Energy Environ. Sci.* **2009**, *2*, 251.
(12) Brabec, C. J.; Heeney, M.; McCulloch, I.; Nelson, J. *Chem. Soc. Rev.* **2011**, *40*, 1185.
(13) Beiley, Z. M.; McGehee, M. D. *Energy Environ. Sci.* **2012**, *5*, 9173.
(14) Ruiz, C.; Garcia-Frutos, E. M.; Hennrich, G.; Gomez-Lor, B. *J. Phys. Chem. Lett.* **2012**, *3*, 1428.
(15) Kirkpatrick, J. *Int. J. Quantum Chem.* **2008**, *108*, 51.
(16) Troisi, A.; Cheung, D. L.; Andrienko, D. *Phys. Rev. Lett.* **2009**, *102*, 116602.
(17) Kwiatkowski, J. J.; Frost, J. M.; Nelson, J. *Nano Lett.* **2009**, *9*, 1085.
(18) MacKenzie, R. C. I.; Frost, J. M.; Nelson, J. *J. Chem. Phys.* **2010**, *132*, 064904.
(19) Troisi, A.; Orlandi, G. *Phys. Rev. Lett.* **2006**, *96*, 086601.
(20) Cheung, D. L.; Troisi, A. *J. Phys. Chem. C* **2010**, *114*, 20479.
(21) Oberhofer, H.; Blumberger, J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 13846.
(22) Oberhofer, H.; Blumberger, J. *Angew. Chem., Int. Ed.* **2010**, *49*, 3631.
(23) Seidel, R.; Faubel, M.; Winter, B.; Blumberger, J. *J. Am. Chem. Soc.* **2009**, *131*, 16127.
(24) Moens, J.; Seidel, R.; Geerlings, P.; Faubel, M.; Winter, B.; Blumberger, J. *J. Phys. Chem. B* **2010**, *114*, 9173.
(25) Tateyama, Y.; Blumberger, J.; Ohno, T.; Sprik, M. *J. Chem. Phys.* **2007**, *126*, 204506.
(26) Breuer, M.; Rosso, K. M.; Blumberger, J. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 611.
(27) Tipmanee, V.; Blumberger, J. *J. Phys. Chem. B* **2012**, *116*, 1876.
(28) Gershenson, M. E.; Podzorov, V.; Morpurgo, A. F. *Rev. Mod. Phys.* **2006**, *78*, 973.
(29) Grozema, F. C.; Siebbeles, L. D. A. *Int. Rev. Phys. Chem.* **2008**, *27*, 87.
(30) Troisi, A. *J. Chem. Phys.* **2011**, *134*, 034702.
(31) Wang, L.; Beljonne, D. *J. Phys. Chem. Lett.* **2013**, *4*, 1888.
(32) Wang, L.; Beljonne, D. *J. Chem. Phys.* **2013**, *139*, 064316.
(33) Ren, J.; Vukmirovic, N.; Wang, L.-W. *Phys. Rev. B* **2013**, *87*, 205117.
(34) Tully, J. C. *J. Chem. Phys.* **1990**, *93*, 1061.
(35) Tully, J. C. In *Classical and Quantum Dynamics in Condensed Phase Simulations*; Berne, B. J., Ciccotti, G., Coker, D. F., Eds.; World Scientific: 1998; p 34.
(36) Tully, J. C. In *Modern Methods for Multidimensional Dynamics Computations in Chemistry*; Thompson, D. L., Eds.; 1998; p 493.
(37) Akimov, A. V.; Neukirch, A. J.; Prezhdo, O. V. *Chem. Rev.* **2013**, *113*, 4496.
(38) Niehaus, T. A.; Heringer, D.; Torralva, B.; Frauenheim, Th. *Eur. Phys. J. D* **2005**, *35*, 467.
(39) Nelson, T.; Fernandez-Alberti, S.; Roitberg, A. E.; Tretiak, S. *J. Chem. Phys.* **2013**, *138*, 224111.
(40) Akimov, A. V.; Prezhdo, O. V. *J. Am. Chem. Soc.* **2014**, *136*, 1599.
(41) Kubar, T.; Elstner, M. *J. Phys. Chem. B* **2010**, *114*, 11221.
(42) Woiczikowski, P. B.; Steinbrecher, T.; Kubar, T.; Elstner, M. *J. Phys. Chem. B* **2011**, *115*, 9846.
(43) Kubar, T.; Elstner, M. *Phys. Chem. Chem. Phys.* **2013**, *15*, 5794.
(44) Kubar, T.; Elstner, M. *J. R. Soc. Interface* **2013**, *10*, 20130415.
(45) Hoffmann, R. *J. Chem. Phys.* **1963**, *39*, 1397.
(46) Troisi, A.; Orlandi, G. *J. Phys. Chem. B* **2002**, *106*, 2093.
(47) Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. *J. Chem. Phys.* **2001**, *10252*, 114.
(48) Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. *Chem. Phys. Lett.* **2001**, *350*, 297.
(49) Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. *J. Chem. Phys.* **2002**, *117*, 9138.
(50) Wu, Q.; Van Voorhis, T. *J. Chem. Phys.* **2006**, *125*, 164105.
(51) de la Lande, A.; Salahub, D. R. *J. Mol. Struct.: THEOCHEM* **2010**, *943*, 115.
(52) Oberhofer, H.; Blumberger, J. *J. Chem. Phys.* **2009**, *131*, 064101.
(53) Oberhofer, H.; Blumberger, J. *J. Chem. Phys.* **2010**, *133*, 244105.
(54) Kubas, A.; Hoffmann, F.; Heck, A.; Oberhofer, H.; Elstner, M.; Blumberger, J. *J. Chem. Phys.* **2014**, *140*, 104105.
(55) Senthilkumar, K.; Grozema, F. C.; Bickelhaupt, F. M.; Siebbeles, L. D. A. *J. Chem. Phys.* **2003**, *9809*, 119.
(56) Mulliken, R. S.; Rieke, C. A.; Orloff, D.; Orloff, H. *J. Chem. Phys.* **1949**, *17*, 1248.
(57) Newton, M. D. *Chem. Rev.* **1991**, *91*, 767.
(58) Rispens, M. T.; Meetsma, A.; Rittberger, R.; Brabec, C. J.; Sariciftci, N. S.; Hummelen, J. C. *Chem. Commun.* **2003**, 2116.
(59) Developer's version of the CPMD 4.0, The CPMD Consortium; MPI für Festkörperforschung and the IBM Zurich Research Laboratory: 2013. http://www.cpmd.org (accessed Aug 21, 2014).
(60) Gajdos, F.; Oberhofer, H.; Dupuis, M.; Blumberger, J. *J. Phys. Chem. Lett.* **2014**, *5*, 2765.
(61) Pieniazek, P. A.; Arnstein, S. A.; Bradforth, S. E.; Krylov, A. I.; Sherrill, C. D. *J. Chem. Phys.* **2007**, *127* (16), 164110.
(62) Clementi, E.; Raimondi, D. L. *J. Chem. Phys.* **1963**, *38*, 2686.
(63) Farazdel, A.; Dupuis, M.; Clementi, E.; Aviram, A. *J. Am. Chem. Soc.* **1990**, *112*, 4206.
(64) Blumberger, J.; McKenna, K. *Phys. Chem. Chem. Phys.* **2013**, *15*, 2184.