JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

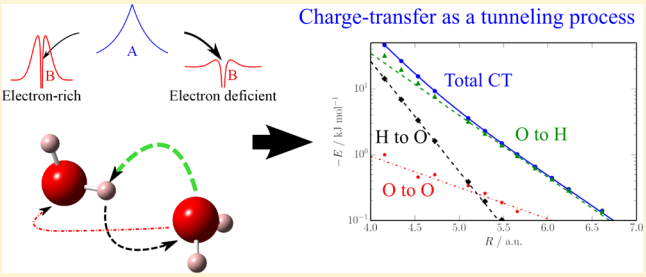
Charge Transfer from Regularized Symmetry-Adapted Perturbation Theory

Alston J. Misquitta*

School of Physics and Astronomy, Queen Mary, University of London, London E1 4NS, U.K.

Supporting Information

ABSTRACT: The charge-transfer (CT) together with the polarization energy appears at second and higher orders in symmetry-adapted perturbation theory (SAPT). At present there is no theoretically compelling way of isolating the charge-transfer energy that is simultaneously basis-set independent and applicable for arbitrary intermolecular separation. We argue that the charge-transfer can be interpreted as a tunneling phenomenon and can therefore be defined via regularized SAPT. This leads to a physically convincing, basis-independent definition of the charge-transfer energy that captures subtleties of the process, such as the asymmetry in the forward and backward charge transfer, as well as secondary transfer effects. With this definition of the charge-transfer the damping needed for polarization models can be determined with a level of confidence hitherto not possible.



I. INTRODUCTION

Symmetry-adapted perturbation theory (SAPT) remains one of the most accurate and versatile methods for calculating intermolecular interaction energies. Within SAPT the interaction energy is decomposed into physical components such as the electrostatic, exchange−repulsion, induction, and dispersion energies. This decomposition is useful both as an interpretative tool and as a means for developing models for the interaction energy. The latter is made possible as, with the exception of the short-ranged exponentially decaying exchange-repulsion energy, each of these components is associated with a well-defined multipole expansion, the coefficients of which can be calculated from monomer properties alone.

There is, however, an important exception to the list of physical components of the interaction energy as described by SAPT: the charge-transfer energy is not defined as a separate component but is included, together with the polarization energy, in the induction energy at second and higher orders in the interaction operator. While these components are not usually separated in a SAPT calculation of the interaction energy, there are approximate methods, such as that introduced by Stone[1] and used by Stone and Misquitta[2] in the context of a variant of SAPT based on a density-functional description of the monomers (termed SAPT(DFT)),[2] which enable us to make a reasonable partitioning of the induction energy—at least at second-order—into its charge-transfer and polarization constituents. Though these methods and their supermolecular counterparts[3] provide sensible charge-transfer energies, particularly for complexes at their equilibrium configurations, they exhibit serious drawbacks, all related to their basis-set dependence for small intermolecular separations. Furthermore, as we shall demonstrate in this paper, the Stone and Misquitta procedure leads to charge-transfer energies that do not seem to reflect the physical nature of the process.

There are several reasons why we require a rigorous definition of the charge-transfer energy.

- *Polarization models*: The polarization energy can be unambiguously defined through molecular polarizabilities and multipole moments, but only for large molecular separations where charge-density overlap is negligible. At shorter separations the expansion must be damped, and to determine the damping we require a knowledge of the polarization energy, and hence the charge-transfer energy.
- *Validation of density functionals*: One of the remaining issues with the exchange functionals used in density functional theory (DFT) is their inability to describe the charge transfer energy accurately.[4] A considerable effort is being made to attempt to remedy this problem, but for this to work accurate benchmark charge-transfer energies are needed.
- *Energy decomposition analysis (EDA)*: EDA techniques are a powerful means of developing an understanding of the nature of molecular interactions. However, with the exception of perturbative techniques like SAPT, these methods contain an element of ambiguity: energy components, particularly the charge-transfer and polarization energies, obtained from different EDAs can be significantly different. To benchmark EDA methods we need a physical and basis-independent definition of the interaction energy components. SAPT already provides

Received: August 7, 2013
Published: October 21, 2013

   dx.doi.org/10.1021/ct400704a | J. Chem. Theory Comput. 2013, 9, 5313−5326

such a decomposition for the electrostatic, exchange-repulsion, and dispersion energies, but not the charge-transfer and polarization. This disadvantage needs to be overcome.

The second and third points are linked: EDA techniques are used to determine the physical content of DFT interaction energies, but to do so reliably, the EDA methods themselves need to be calibrated against more rigorous decomposition methods, such as SAPT. In this paper we will focus on the first and second: the role the charge transfer plays in determining accurate polarization models and the provision of benchmark charge-transfer energies.

II. WHAT IS THE CHARGE-TRANSFER ENERGY?

Intermolecular charge-transfer is a ground-state phenomenon. When two molecules are in proximity, their electronic charge density can be delocalized, that is, shared between them, leading to a stabilization by an amount that is the charge-transfer energy. However, from the point of view of the interacting monomers, the delocalization can be viewed in terms of excitations partly localized on the partner monomer: these are the so-called intermolecular charge-transfer excitations. This is why the charge-transfer energy first appears at second order in intermolecular perturbation theory. Common to both views is the idea of charge delocalization leading to stability. One may even suggest that the term "charge-transfer" is misleading, and rather, "charge-delocalization" may be a more appropriate term for this phenomenon. We will return to this issue of nomenclature later in the paper.

This simple physical picture of the charge transfer process lies at the heart of current attempts to define the charge-transfer energy: if we can calculate interaction energies both by allowing and suppressing the delocalization process, the charge-transfer energy can be defined as an energy difference.

II.1. Basis-Space Definitions. Currently, definitions of the charge-transfer energy rely on basis-space localization methods. There are many flavours of this kind of approach, all of which are based on the idea that with a suitably localized basis set we suppress the charge-transfer-type excitations and consequently isolate the charge-transfer from the polarization energy. In the Stone and Misquitta[2] technique the charge-transfer energy is defined as the difference in the second-order induction energy, $E^{(2)}_{\mathrm{ind,tot}}$, calculated in a dimer centered (DC) and monomer centered (MC) basis set. That is,

$$\mathrm{CT}^{(2)}(\mathrm{SM09}) = E^{(2)}_{\mathrm{ind,tot}}[\mathrm{DC}] - E^{(2)}_{\mathrm{ind,tot}}[\mathrm{MC}] \quad (1)$$

Here, unless otherwise specified, the induction energy will be assumed to be the sum of the polarization and exchange components as[5] $E^{(n)}_{\mathrm{ind,tot}} = E^{(n)}_{\mathrm{ind,pol}} + E^{(n)}_{\mathrm{ind,exch}}$. In the MC basis the molecule is described using basis functions located on its own nuclei only, therefore suppressing any excitation that could give rise to an increased electronic density on the partner monomer. That is, charge-transfer would be suppressed. By contrast, in the DC type of basis, the monomers are additionally described using basis functions located on their partners, therefore allowing charge-transfer type excitations. Hence we are led to the above definition of the CT energy which we will henceforth term CT-SM09.

Note that in practical calculations, instead of the DC type of basis we often use an equivalent, but smaller, MC+ type of basis which includes midbond and the so-called far-bond functions.[6] The far-bond functions are a subset of the basis of the partner monomer; usually the s- and p-functions only. Extensive tests have shown that these two types of basis sets result in nearly identical induction energies.[6] Consequently we will use the DC and MC+ types interchangeably in eq 1 and will consider induction energies calculated in either basis type to include all CT effects.

There are, however, a few objections that one could raise with the CT-SM09 definition:

- Equation 1 leads to a basis-dependent definition of the CT energy. As the monomer basis gets larger and more complete, the distinction between the MC and DC basis types grows smaller. This would lead to smaller apparent CT energies. This issue has been acknowledged by Stone and Misquitta who have argued that this effect is indeed present but is small enough that for practical basis sets this is not and issue.
- A more subtle but related shortcoming of eq 1 is that it inevitably leads to a separation-dependent definition of the charge transfer. As the intermolecular separation reduces, the MC-type basis set is better able to describe CT-type excitations; that is, the definition of the CT gets progressively worse.
- The CT-SM09 definition is really only the second-order charge-transfer energy. There are contributions to the CT from third and higher orders in perturbation theory that, as we shall see, are comparable to the second-order CT.
- This definition relies on the use of localized atomic basis sets and offers no clear route to extension to plane-wave basis sets. Consequently, eq 1 is largely limited to programs that use Gaussian-type orbitals. This is certainly not a significant issue, but it would be advantageous to develop a method that was independent of the type of basis set.

The first two points are illustrated in Figure 1 for the water dimer in its minimum-energy, H-bonded conformation.[7,8] The variation of the CT-SM09 energy with basis set is indeed small at the equilibrium separation, particularly for the two larger basis sets. However, as the intermolecular separation reduces,

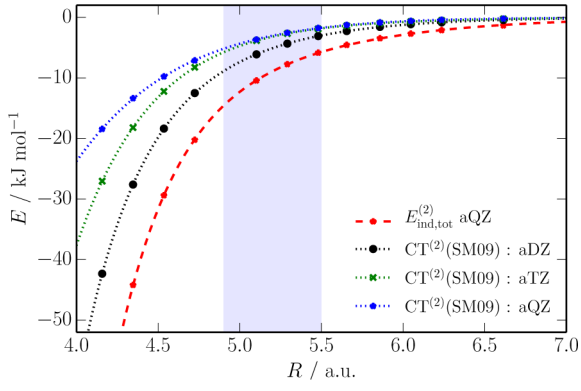


Figure 1. Second-order charge-transfer for the water dimer calculated using the Stone−Misquitta procedure (CT-SM09) with three basis sets (dotted lines). Also included is that total second-order induction energy $E^{(2)}_{\mathrm{ind,tot}}$ calculated using the aug-cc-pVQZ basis of the MC+ type. $E^{(2)}_{\mathrm{ind,tot}}$ calculated with the aug-cc-pVDZ and aug-cc-pVTZ MC+ basis types are visually identical and are not shown. R is the O···O separation. The shaded region indicates the range of R_{OO} values attained in the water dimer (largest value) and clusters of water through to the water hexamer (smallest value).

dx.doi.org/10.1021/ct400704a | J. Chem. Theory Comput. 2013, 9, 5313−5326

the differences between all basis sets grows, with the aug-cc-pVQZ basis resulting in the least amount of (second-order) charge-transfer energy at short separations.

These short separations are important for several reasons. The separation of water molecules can decrease significantly in clusters due to a combination of many-body polarization and charge-transfer effects. For example, the oxygen−oxygen separation in the water dimer at its global energy minimum configuration is 2.9 Å, but it is less than 2.5 Å in the cage hexamer. This range is indicated in Figure 1. Presumably the intermolecular separation could be even smaller for larger clusters with stronger many-body effects, or for water under pressure.

The differences in CT-SM09 calculated using the two larger basis sets may seem small even at the short end of the R_{OO} range, but the variation in CT-SM09 with basis at even shorter separations is large enough that the polarization damping model—which is fitted largely to these energies—becomes basis-set dependent with damping parameters varying from 1.4, to 1.7, through 1.9 au for the aDZ, aTZ, and aQZ basis sets. Not only is this variation large, but the damping parameters tend to depend strongly on the range of data used—a possible indication that the energies we are fitting to are not entirely polarization and might be contaminated with charge transfer. This variation in the damping parameter has no effect on the two-body interaction energy but causes the many-body polarization energy for clusters of water molecules to vary considerably. For example, with the above damping parameters, the many-body polarization energy of the cyclic water pentamer varies from −59, to −73, to −79 kJ mol^{-1}. This variation is large enough to make these polarization models unreliable and almost unusable without explicitly fitting them to reproduce the many-body energies of such clusters.[9] While this can be done, this approach is tedious and unsatisfactory as there is always the possibility that a damping model that works for one set of clusters may not work for another.

The Stone and Misquitta technique is but one of many methods that attempt to define the charge-transfer energy. Before continuing we mention two of these: Schenter and Glendening[10] have proposed a decomposition based on natural bond orbitals, which has the merit that the charge-transfer energies appear to converge with basis set, but these energies are unreasonably large. More recently Khaliullin et al.[3] developed a method that operates on a principle similar to that of Stone and Misquitta, only this time with absolutely localized molecular orbitals (ALMOs) used to suppress charge-transfer-type excitations. In a recent analysis by Azar et al.[11] the ALMO method has been demonstrated to exhibit shortcomings similar to those of the Stone and Misquitta definition: there is no complete basis set limit to the charge-transfer energy defined in this procedure.

III. CHARGE TRANSFER VIA REGULARIZED SAPT

The Stone−Misquitta and related approaches are all basis-space methods; that is, these methods attempt to isolate the charge-transfer energy by manipulating the atomic or molecular basis sets to try to suppress CT-type excitations. To see how we may find an alternative approach consider the Rayleigh−Schrödinger perturbation expression for the polarization component of the second-order induction energy of molecule A (the total induction energy is obtained by the inclusion of exchange effects contained in the exchange-induction energy[12]):

$$E^{(2)}_{\mathrm{ind,pol}}(A) = \sum_{r\neq 0} \frac{|\langle \Phi^{\mathrm{A}}_0 \Phi^{\mathrm{B}}_0 | \hat{V} | \Phi^{\mathrm{A}}_r \Phi^{\mathrm{B}}_0 \rangle|^2}{E^{\mathrm{A}}_0 - E^{\mathrm{A}}_r} = \sum_{r\neq 0} \frac{|\langle \Phi^{\mathrm{A}}_0 | \hat{\Omega}^{\mathrm{B}} | \Phi^{\mathrm{A}}_r \rangle|^2}{E^{\mathrm{A}}_0 - E^{\mathrm{A}}_r} = \langle \Phi^{\mathrm{A}}_0 | \hat{\Omega}^{\mathrm{B}} | \Phi^{\mathrm{A}}_0(1) \rangle \quad (2)$$

where the Φ^{X}_n and E^{X}_n are the exact eigenstates and eigenvalues of monomer X = A, B and many-electron operator electrostatic operator $\hat{\Omega}^{\mathrm{B}} = \sum_{i\in \mathrm{B}} \omega^{\mathrm{B}}(\mathbf{r}_i)$ where $\omega^{\mathrm{B}}(\mathbf{r})$ is the total electrostatic potential of monomer B, that is, it arises from both electrons and nuclei of B and can be written as

$$\omega^{\mathrm{B}}(\mathbf{r}) = -\sum_{\beta} \frac{Z_\beta}{|\mathbf{r} - \mathbf{R}_\beta|} + \int \frac{\rho^{\mathrm{B}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\, \mathrm{d}\mathbf{r}' \quad (3)$$

where Z_β and $\mathbf{R}_\beta$ are the nuclear charges and position vectors of monomer B and ρ^{B} is the unperturbed electronic charge density of B. In the last step in eq 2 we have defined the first-order induction wave function correction $\Phi^{\mathrm{A}}_0(1)$:

$$\Phi^{\mathrm{A}}_0(1) = \sum_{r\neq 0} \frac{\Phi^{\mathrm{A}}_r \langle \Phi^{\mathrm{A}}_r | \hat{\Omega}^{\mathrm{B}} | \Phi^{\mathrm{A}}_0 \rangle}{E^{\mathrm{A}}_0 - E^{\mathrm{A}}_r} \quad (4)$$

Similar expressions exist for the second-order induction energy of monomer B. The subscript "pol" in eq 2 indicates that this is the energy obtained from the so-called polarization expansion,[13] that is, without the inclusion of exchange effects. This notation is unfortunate as it can lead to the erroneous conclusion that $E^{(2)}_{\mathrm{ind,pol}}(A)$ is the classical polarization energy, when, as discussed earlier, this term contains parts of both the classical polarization and charge-transfer energies.

The third form of $E^{(2)}_{\mathrm{ind,pol}}(A)$ in eq 2 allows us to make the useful interpretation of the second-order induction as the second-order energy response to the static potential of the partner monomer (or, more generally, the potential arising from the entire environment of monomer A). This potential, eq 3, consists of two parts: an attractive, but singular, term arising from the nuclei of B, and a finite, repulsive term arising from the electrons of B. At long-range, monomer A sees the net effect of these two terms, but at short-range, when charge-densities overlap, monomer A additionally responds to the singularity arising from the nuclear potential of B. The electronic density of B screens the singularity of the nuclear potential, but the extent of the screening depends on whether B is electron-rich or electron-deficient. For an electron-rich atom B (such as the electronegative oxygen atom in water) the nuclear potential is heavily screened, leading to little tunneling of the electronic charge density of A into the nuclear potential of B. However, for an electron-deficient atom (such as the electropositive hydrogen atom in a water molecule) the screening is weak, leading to a significant degree of tunneling.

This interpretation of the charge-transfer offers us a natural way of suppressing the charge-transfer process altogether: if the potential well developed by the nuclear potential can be suitably eliminated, there would be no possibility of tunneling and, consequently, no charge transfer. This can be achieved by *regularizing* the nuclear potential that appears in eq 3 to suppress the nuclear well.

III.1. Regularization. Regularization is a technique originally developed to accelerate the convergence of the

class of SAPT theories. Here we will be concerned mainly with the version of SAPT based on symmetrized Rayleigh−Schrödinger (SRS) perturbation theory. The subject of regularization and the role it plays in convergence of SAPT is comprehensively discussed in refs 14 and 15. Although a full discussion of convergence issues would be out of place in this article, we will discuss some of the important issues here as these will serve to place the concept of regularization in context.

Possibly the earliest major attempt to understand the convergence properties of intermolecular perturbation theories was made by Claverie[16] who argued that the polarization expansion[13] (essentially many-body Rayleigh−Schrödinger perturbation theory) on which SRS theory is based either diverged, or if it did converge, for systems with more than two electrons, it would converge to a Pauli-forbidden state which would be more strongly bound than the physical ground state. These ideas were expanded by Kutzelnigg[17] who showed that the polarization expansion indeed diverged. Simultaneously, Morgan and Simon[18] proved that if any of the interacting atoms had atomic number greater than two, the physical ground state of the system could be buried in a continuum of Pauli-forbidden states.

The cause of the divergence of the polarization expansion, and consequently SRS theory,[19] lies in the presence of the unphysical states. Adams explored this issue in a series of papers (see for example ref 19) and showed that the concept of regularization introduced by Herring[20] could be used to destabilize these unphysical tunneling states and hence to ensure the convergence of the theory. Simultaneously, Patkowski and collaborators[15,21] applied similar ideas to develop regularized perturbation theories that converged to the physical ground state of the dimer. The idea here is to write the singular electron−nuclear potential as a short-ranged, singular part and a long-ranged part that is well-behaved. In the notation used by Patkowski et al. this is expressed as

$$\frac{1}{r} = v_p(r) + v_t(r) \tag{5}$$

where v_t is the singular, short-ranged part and v_p the long-ranged, well-behaved part of the nuclear potential. Various schemes can be used to achieve this splitting; here we will use the Gaussian-based scheme:[21]

$$\begin{aligned} v_p(r) &= \frac{1}{r}(1 - e^{-\eta r^2}) \\ v_t(r) &= \frac{1}{r}e^{-\eta r^2} \end{aligned} \tag{6}$$

Patkowski et al. used SRS theory for the nonsingular part (v_p) and a strong symmetry-forcing theory (using a symmetrization technique completely suppressing the Pauli-forbidden terms) for the singular part,[21,22] leading to a unified theory proven to be convergent for the van der Waals and as well as chemical bonding separations.[15]

III.2. Charge-Transfer via Regularization. The issues discussed above are fundamental to foundation and development of perturbation theories but are not germane at low orders in perturbation theory where even weak symmetry-forcing theories like SRS theory are known to yield sensible results.[13,15,22] There is a tremendous body of work demonstrating the accuracy of low-order SAPT (truncated at second or third order with higher-order corrections estimated in a nonperturbative manner) for numerous systems (for example, see refs 12 and 23−26). For such a truncated theory regularization has a very different role to play: following on from our discussion of the charge transfer, we argue that, if used at second and higher orders in perturbation theory, the regularization procedure can be used to destabilize the charge-transfer states and consequently allow us to define a charge-transfer free interaction energy.

At second order in perturbation theory we calculate the regularized induction energy, $E^{(2)}_{\mathrm{ind,tot}}(\mathrm{Reg})$, that, with an appropriate amount of regularization that is yet to be defined, is free of charge-transfer contributions (details below). Then, in a manner analogous with eq 1, we may define the second-order charge-transfer energy as

$$\mathrm{CT}^{(2)}(\mathrm{Reg}) = E^{(2)}_{\mathrm{ind,tot}} - E^{(2)}_{\mathrm{ind,tot}}(\mathrm{Reg}) \tag{7}$$

Here $E^{(2)}_{\mathrm{ind,tot}}(\mathrm{Reg})$ is the regularized second-order induction energy that may be identified with the true polarization energy. There is no basis restriction on the above definition, except that the basis set used needs to be large enough to converge the total induction energy, which is the usual requirement for any energy calculation.

III.3. Implementation. We have implemented a version of regularized SRS theory (R-SRS) derived by Patkowski et al. in ref 27. In R-SRS theory the dimer wave function corrections are obtained in response to the interaction operator with regularized nuclear potentials, but the interaction energy corrections are calculated using the original, unregularized interaction operator. For the second-order induction energy this means that instead of using $\Phi^{\mathrm{A}}_0(1)$, we calculate the first-order induction wave function correction in response to the *regularized* electrostatic potential $\omega^{\mathrm{B}}_{\mathrm{Reg}}$:

$$\omega^{\mathrm{B}}_{\mathrm{Reg}}(\mathbf{r}) = -\sum_{\beta} Z_\beta v_p(\mathbf{r} - \mathbf{R}_\beta) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\,\mathrm{d}\mathbf{r}' \tag{8}$$

to give

$$\Phi^{\mathrm{A}}_0(1)[\mathrm{Reg}] = \sum_{r\neq 0} \frac{\Phi^{\mathrm{A}}_r \langle \Phi^{\mathrm{A}}_r | \hat{\Omega}^{\mathrm{B}}_{\mathrm{Reg}} | \Phi^{\mathrm{A}}_0 \rangle}{E^{\mathrm{A}}_0 - E^{\mathrm{A}}_r} \tag{9}$$

where the many-electron electrostatic operator is defined as $\hat{\Omega}^{\mathrm{B}} = \sum_{i\in \mathrm{B}} \omega^{\mathrm{B}}_{\mathrm{Reg}}(\mathbf{r}_i)$. The regularized second-order polarization component of the induction energy is then defined as

$$E^{(2)}_{\mathrm{ind,pol}}(\mathrm{Reg}) = \langle \Phi^{\mathrm{A}}_0 | \hat{\Omega}^{\mathrm{B}} | \Phi^{\mathrm{A}}_0(1)[\mathrm{Reg}] \rangle \tag{10}$$

To this, as always, we need to add the similarly regularized exchange-induction energy.[13,15,22] Notice that in this step we have used the full (unregularized) electrostatic potential of monomer B.

The SAPT(DFT) expression for the polarization part of the induction energy of monomer A in response to the field of B is

$$E^{(2)}_{\mathrm{ind,tot}}(\mathrm{A} \leftarrow \mathrm{B}) = 2 s^i_v (\omega^{\mathrm{B}})^v_i \tag{11}$$

where i and v label occupied and virtual states, $(\omega^{\mathrm{B}})^v_i$ are matrix elements of the unperturbed potential of monomer B given in eq 3, and the amplitudes s^i_v are obtained, in the case of SAPT(DFT), by solving the coupled Kohn−Sham equations

$$\mathbf{H}^{(1)}_{\mathrm{A}} \mathbf{s}^{\mathrm{A}} = -\omega^{\mathrm{B}} \tag{12}$$

where $\mathbf{H}_A^{(1)}$ is the electric Hessian[28,29] of Kohn–Sham linear-response theory that is given in full form for hybrid functionals in ref 5. The expression for $E_{ind,exch}^{(2)}(A \leftarrow B)$ also involves the amplitudes defined above, though the matrix elements multiplying it are more complex and are given in full form in refs 27 and 30. Analogous expressions exist for the induction energy of monomer B due to the field of A.

In regularized SAPT(DFT) (R-SAPT(DFT)) the amplitudes—the coefficients of the first-order induction wave function—are calculated using eq 12, but this time with the matrix elements on the RHS replaced by those calculated with the regularized electrostatic potential $\omega_{Reg}^{B}(\mathbf{r})$ given in eq 8. Otherwise, the expressions for $E_{ind,pol}^{(2)}$ and $E_{ind,exch}^{(2)}$ are identical with the nonregularized forms. These equations have been implemented in the CamCASP program[31] and are available as part of a regular SAPT(DFT) calculation of the interaction energy. Note that the expression for $E_{ind,exch}^{(2)}$ used in this paper does not involve scaling.[5] The scaling expression seems to work by cancellation of errors (of the second-order terms with corresponding terms in the δ_{int}^{HF} estimate of the higher-order induction effects[13]) and cannot be relied upon to give reasonable results for the regularized induction energy.

IV. NUMERICAL DETAILS

All calculations of SAPT(DFT) energies and molecular electrostatic and polarization models have been performed with the CamCASP[31] program. The Kohn–Sham orbitals and orbital energies used in by the CamCASP program were obtained using the DALTON 2.0[32] program with a patch distributed as part of the SAPT2008[33] suite of codes.

SAPT(DFT) induction energies were calculated using the PBE0[34,35] exchange-correlation functional asymptotically corrected (AC) using the Fermi–Amaldi[36] long-range form and the Tozer-Handy splicing function.[37] For details see ref 24. The second-order induction energies were evaluated using the hybrid form of the linear-response kernel.[5,24] $E_{ind,exch}^{(2)}$ was evaluated without scaling[5] using the expressions in refs 38 and 39. This was found to be necessary for the regularized induction energies.

Distributed multipoles have been calculated using the GDMA2[40] module that is part of the CamCASP suite. Unless otherwise stated, these include terms to rank 4 (hexadecapole moments) on all sites, including the hydrogen atoms. Distributed (anisotropic) polarizabilities have been calculated using the Williams–Stone–Misquitta (WSM) method[5,41] with terms to rank 3 on all sites except for the hydrogen atoms for which these were restricted to rank 1. Molecular multipole and polarizability models have been calculated using a d-aug-cc-pVTZ basis with the PBE0/AC density and density response functions. Model energy evaluations were performed using the Orient[42] program.

V. RESULTS

V.1. Determining the Regularization Parameter η.

Regularization involves the parameter η that has the units of L^{-2}. Equivalently, $\eta^{-1/2}$ has the dimensions of length. From Figure 2 and the discussion in Section III we know that this length scale will be associated with the width of the screened nuclear potential, which, in principle, will be dependent on the atom type and bonding environment. We are faced with two choices: either the parameter η is only weakly dependent on atom type, in which case a fixed value may be used for all calculations, or this parameter exhibits a strong atom-type dependence, in which case the regularization procedure will be potentially cumbersome to apply in practice. The first and most important question is how are we to determine the appropriate value of η?

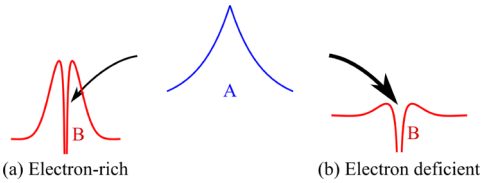


Figure 2. Interpreting the charge transfer as a tunneling process. We can identify two cases for atom B. (a) Electron-rich: The nuclear potential of B is heavily screened by its electronic density. (b) Electron-deficient: The nuclear potential of B is poorly screened. Charge-transfer from the electronic density of A (center, blue) into the potential of B is minimal in case (a) but it is significant for the case (b). This is illustrated by the width of the arrows.

In principle, η could be obtained by examining the screened nuclear potential and determining the appropriate regularization needed to "fill" it in to as to prevent all tunneling states. But it is at present unclear what would constitute sufficient "filling"; consequently, we have instead opted for two alternate procedures:

- We could determine η by observing that at the optimal regularization there will be no charge transfer, but the polarization component of the induction energy will be left unchanged. Consequently, all basis sets large enough to describe the pure polarization effect would result in identical regularized induction energies.
- Alternatively, we could determine η by requiring that all the induction energy in rare gas dimers is charge-transfer.

Consider the first proposal: Our premise here is that the true polarization energy can be described by a reasonably large monomer basis, that is, without the need of basis functions located on the partner monomer. Consequently, with the correct regularization, i.e., the value of η that completely suppresses all charge-transfer tunneling effects without altering the long-range part of the electrostatic potential responsible for the polarization, all basis sets capable of describing the true polarization energy will yield the same, regularized, induction energy. If, on the other hand, the regularization is insufficient, then some degree of charge-transfer will be allowed. This will lead to a spread in energies as the larger basis sets will be better able to describe the residual tunneling. Finally, if the regularization is excessive, the long-range part of the potential will be affected. This will also cause a variation in energies calculated with different basis sets, with the larger, more diffuse basis sets being more affected.

In Figure 3 we display regularized second-order induction energies for the water dimer calculated with various basis sets and three values of η = 1.0, 3.0, and 5.0 au. For the regularization parameter η = 3.0 au all but the smallest basis set results in the same $E_{ind,tot}^{(2)}$(Reg) energy over the entire range of intermolecular separations. The one exception is the aug-cc-pVDZ basis in the monomer-centered type which is unable to adequately describe even the polarization component of the induction energy. For larger and smaller values of η the spread in the results from the six basis sets is seen to increase. To ensure that this observation is not specific to the O···H contact in the water dimer, in Figure 4 we compare regularized

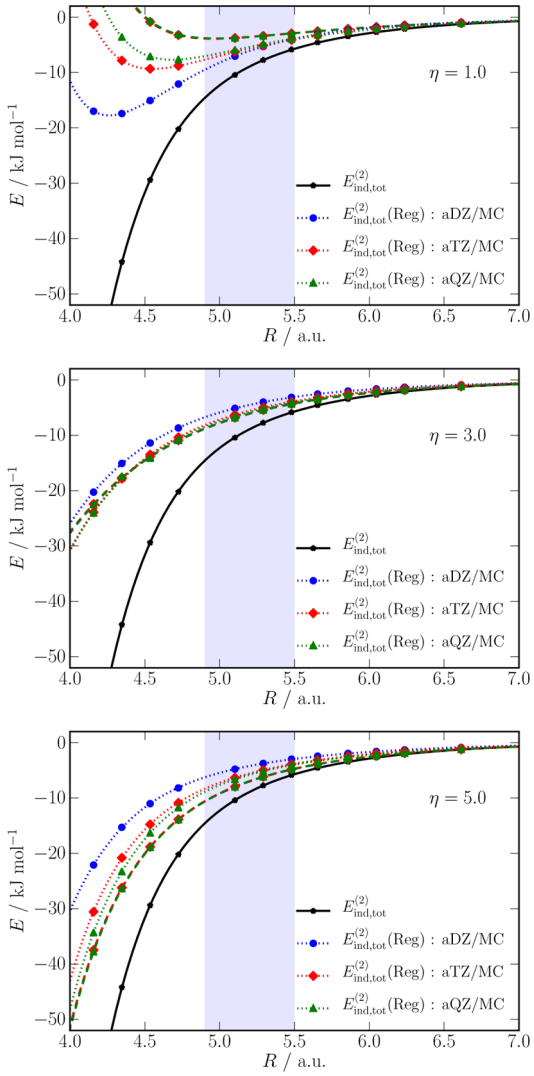


Figure 3. Regularized and unregularized total induction energies for the water dimer in its hydrogen-bonded orientation. $E^{(2)}_{ind,tot}(Reg)$ energies are shown for the aug-cc-pV*x*Z, *x* = D, T, Q, basis sets in the MC (dotted lines) and MC+ (dashed lines) types. The $E^{(2)}_{ind,tot}(Reg)$ energies calculated with the MC+ basis types are nearly indistinguishable. Only the total induction energy (solid black line) calculated with the aug-cc-pVQZ MC+ basis is shown. The shaded area is as described in the caption to Figure 1

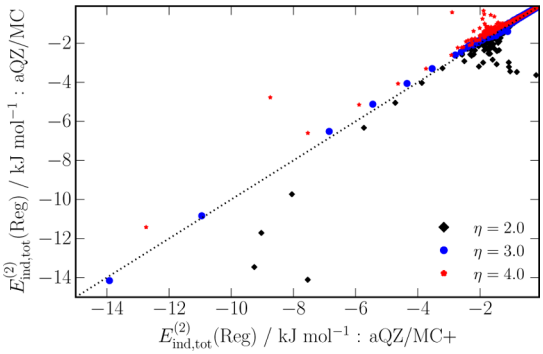


Figure 4. Second-order regularized total induction energies for the water dimer in 400 pseudorandom dimer configurations and 14 hydrogen-bonded configurations. Energies calculated with the aug-cc-pVQZ MC basis set are plotted against those calculated with the aug-cc-pVQZ MC+ basis set. $E^{(2)}_{ind,tot}(Reg)$ is calculated using three values of the regularization parameter η. The 14 H-bonded dimers have energies (with η = 3.0) that are along the diagonal line.

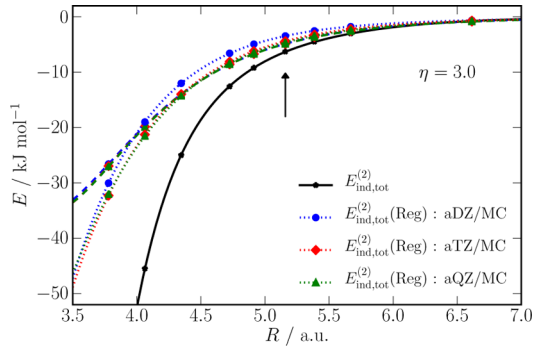


Figure 5. Induction and charge-transfer energies for the HF dimer at its global minimum hydrogen-bonded configuration. See the caption to Figure 3 for a description. The arrow indicates the equilibrium separation.

induction energies for 414 water dimers: 400 of these were chosen using the pseudorandom algorithm described in ref 43 and the remaining 14 are those already shown in the above figures. It is remarkable that with η = 3.0 au the regularized induction energies calculated with the aug-cc-pVQZ MC basis and aug-cc-pVQZ MC+ basis are nearly perfectly in agreement for all 414 dimers. Varying η to 2.0 or 4.0 au results in significantly less correlation between the two sets of energies, not just for the hydrogen-bonded dimer configurations with strong charge-transfer but also for configurations with O···O contacts which exhibit small charge-transfer energies. We point out here that with this procedure we are unable to distinguish between η in the range 3.0 ± 0.2 au. However, it seems fairly clear that the optimum value is very close to 3.0 au

In Figure 5 we display charge-transfer energies for the HF dimer in its minimum energy hydrogen-bonded configuration[44] calculated with various basis sets. Only results for η = 3.0 au are presented as the behavior of this system is largely similar to that of the water dimer: here too, η = 3.0 au is a good choice for the regularization parameter, though the regularized induction energies from the larger basis sets are not in as good agreement at short separations at which a larger value of η may be more appropriate. This may indicate that η should depend on the type of atom. Although this discrepancy shows up at separations not easily accessed, this needs further investigation. In Figure 6 we compare regularized induction energies calculated using a few values of η for 120 pseudorandom dimers and 39 dimers at specific orientations. We see a good agreement of regularized induction energies calculated with the aug-cc-pVQZ MC and MC+ basis sets for a regularization of η = 3.0 au

Now consider the second proposal: that we determine the optimum value of the regularization parameter by requiring that all the induction energy in a rare gas dimer is charge-transfer. There are a few issues with this proposal. The total induction energy of the rare gas dimers is nearly zero at and around the equilibrium separation; consequently, we need to use very short separations to obtain an appreciable total induction energy. This leads to a conundrum: while we can argue that the multipole expansion should not result in any polarization energy for these dimers, at short separations, the expansion is not valid as the charge densities penetrate. It is worth bearing this in mind in the following discussion.

dx.doi.org/10.1021/ct400704a | J. Chem. Theory Comput. 2013, 9, 5313−5326

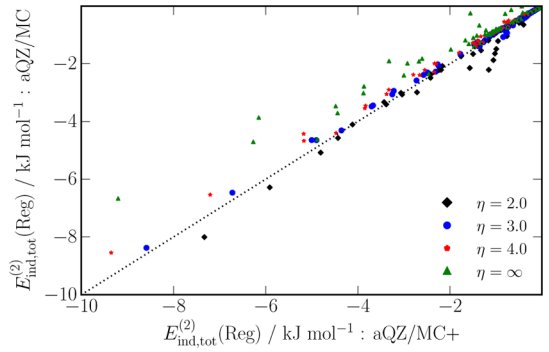


Figure 6. Second-order regularized induction energies for the HF dimer in 120 pseudorandom dimer configurations and three set of 13 configurations at specific geometries: (1) the global minimum, (2) the linear H···H, and (3) linear O···O orientations. $E^{(2)}_{\mathrm{ind,tot}}$(Reg) is calculated using three values of the regularization parameter η with the aug-cc-pVQZ MC basis set values plotted against those calculated with the aug-cc-pVQZ MC+ basis set.

The induction energy of the argon dimer, like that of all rare-gas dimers, is almost zero at the equilibrium separation[45] of 7.13 au, with the polarization and exchange components almost completely canceling. However, as shown in Figure 7, for very

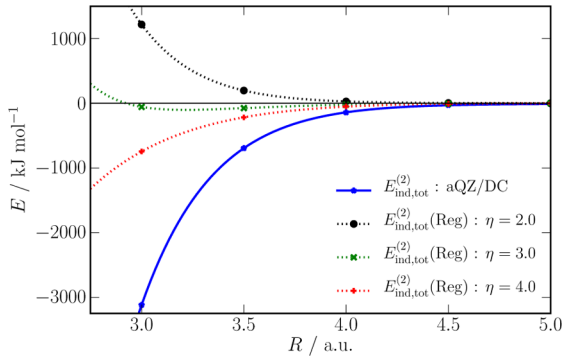


Figure 7. Argon dimer second-order induction energies without regularization and regularized energies with three values of the regularization parameter η. All calculations used the aQZ DC-type of basis. These are very short interatomic separations: the minimum energy configuration is at roughly 7.13 au.

small interatomic separations this is no longer the case, and we get an exponentially varying induction energy. However, with regularization parameter η = 3.0 au the regularized induction energy is close to zero for all separations. Using this value of η we have calculated second-order charge-transfer energies using the aug-cc-pVxZ, x = D, T, Q basis sets. From Figure 8 we see that CT$^{(2)}$(Reg) is insensitive to the basis set, but, as with the water dimer example, CT$^{(2)}$(SM09) shows considerable basis variation, with the charge-transfer energy tending to zero as the basis set increases. Unlike the first proposal, the optimum value of η obtained in this manner is dependent on the type of system: it is 3.2 au for the Ar···Ne complex and 2.9 au for the Ar···He complex. Furthermore, for the neon and helium dimers it varies from 4.0 to 5.0 au, though, in these cases, the total induction energies are considerably smaller even at very short separations, so there is a associated ambiguity in the choice of η.

Even if we set aside the second proposal, the first method provides compelling evidence that the regularization parameter η = 3.0 au is appropriate, at least for the lighter atoms. This

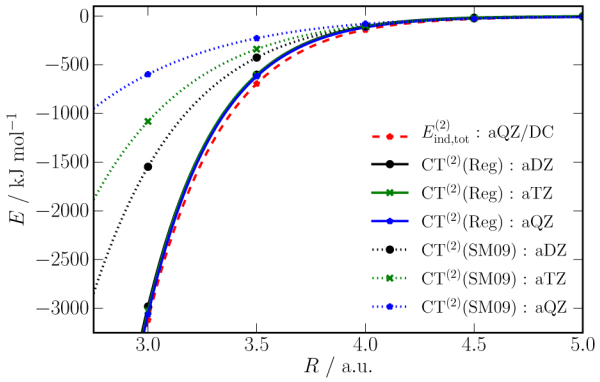


Figure 8. Argon dimer second-order induction energies (solid line) and charge-transfer energies calculated using various basis sets. The second-order charge-transfer energies are calculated using the regularization procedure with η = 3.0 au (dashed lines) and the Stone−Misquitta procedure (dotted lines).

value corresponds to a regularization length-scale of 0.577 Bohr. As mentioned above, we see some indication that η should vary with the type of atom, but numerical evidence suggests that this variation is likely to be small and is probably manifest at very short separations only. In the rest of this article CT$^{(2)}$(Reg) will be computed through eq 7 with the regularization parameter η = 3.0 au

V.2. Analysis of the Second-Order Charge-Transfer Energy. Having determined the manner in which the regularization is to be performed, we will now analyze the CT$^{(2)}$(Reg) energies in detail. In Figure 9 we display

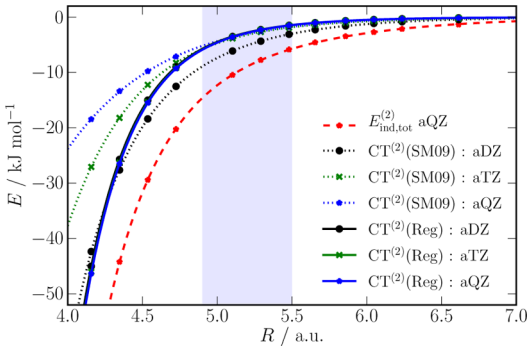


Figure 9. Second-order charge-transfer energies for the water dimer in its hydrogen-bonded orientation. CT$^{(2)}$(Reg) (solid lines) and CT$^{(2)}$(SM09) (dotted lines) energies are shown for the aug-cc-pVxZ, x = D, T, Q, basis sets. All CT$^{(2)}$(Reg) calculations were performed with the MC+ basis type and show virtually no variation with basis. The total (unregularized) second-order induction energy (solid black line) calculated with the aug-cc-pVQZ MC+ basis is shown. The shaded area is as described in the caption to Figure 1.

CT$^{(2)}$(Reg) and CT$^{(2)}$(SM09) energies for the water dimer. The CT$^{(2)}$(Reg) energies have been calculated using eq 7 with both terms on the RHS calculated using the MC+ type of basis and therefore show very little sensitivity to the choice of basis set (as long as augmented double-ζ or better in quality). Contrast this with the strong basis sensitivity of the CT$^{(2)}$(SM09) energies. There are a few features of these energies worth highlighting: at long-range, the CT$^{(2)}$(Reg) energies are similar to those of CT$^{(2)}$(SM09) with the aug-cc-pVQZ basis set. But at short-range, CT$^{(2)}$(Reg) is closer to CT$^{(2)}$(SM09) with the much smaller aug-cc-pVDZ basis set.

dx.doi.org/10.1021/ct400704a | J. Chem. Theory Comput. 2013, 9, 5313−5326

This behavior supports the discussion in Section II.1: The $CT^{(2)}$(SM09) definition does result in accurate second-order charge-transfer energies with the aug-cc-pVQZ basis set at long-range, but as the intermolecular separation decreases, this large basis set starts to pick up charge-transfer excitations leading to an underestimation of the charge transfer defined via eq 1. At short separations, we should expect a smaller basis set, in this case, the aug-cc-pVDZ basis, to yield more accurate $CT^{(2)}$(SM09) energies. This is indeed what our $CT^{(2)}$(Reg) definition demonstrates.

In Figure 10 we present the charge-transfer energies calculated in the aug-cc-pVQZ basis set on a semilog scale.

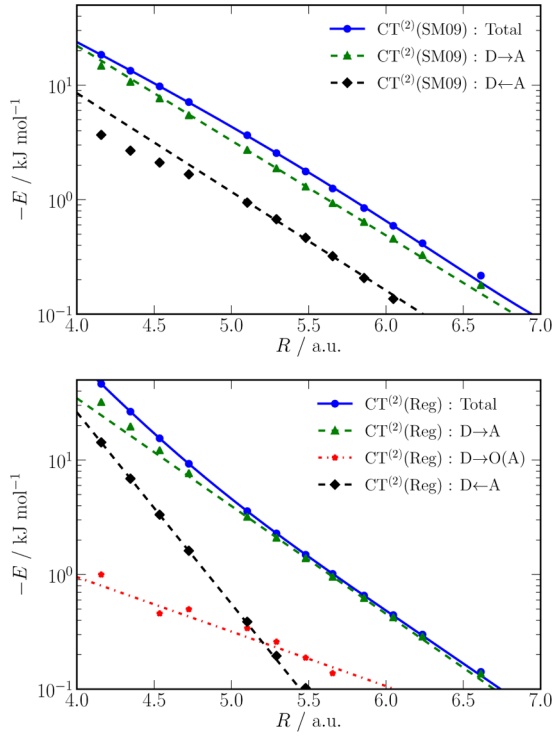


Figure 10. Decomposition of the second-order charge-transfer energies for the water dimer in its hydrogen-bonded orientation. All calculations used the aug-cc-pVQZ basis set. The second-order charge-transfer (solid blue lines) can be decomposed into a contribution from the electron donor (oxygen) to the electron acceptor (hydrogen) (green, dashed) and another, weaker contribution, from the acceptor (hydrogen) to the donor (oxygen) (black, dashed). Furthermore, the regularization procedure allows us to isolate a contribution from the donor to the oxygen of the acceptor (red, dotted).

The charge transfer is usually thought of as being decaying exponentially with separation, so on this plot it should appear as a straight line. However, this is not the case for either of the methods. The $CT^{(2)}$(SM09) energy does exhibit an exponential behavior at large intermolecular separations but becomes subexponential at short separations. This is a consequence of the problem discussed in the above paragraph: the $CT^{(2)}$(SM09) will always result in too little CT at short-range, and this effect is larger for the larger basis sets. By contrast, $CT^{(2)}$(Reg) is superexponential at short separations: it exhibits a clear double (possibly even multiple) exponential behavior. This should be expected. If we accept that the charge-transfer process is a tunneling phenomenon, then we should also expect to see contributions from tunneling into each of the (screened) nuclear wells. The dominant process is expected to be the charge density of the electron donor (oxygen) tunneling into the poorly screened nuclear potential of the electron acceptor (hydrogen). (In this paper, we use the terms "donor" and "acceptor" to refer to electrons and not protons.) There will also be tunneling of the acceptor density into the well-screened nuclear potential of the donor oxygen. However, we may additionally expect weaker processes such as the donor density tunneling into the acceptor oxygen. Each process will be approximately exponential, with the barrier height and width determining the exponent. Perturbation theory allows us make the donor to acceptor and acceptor to donor decomposition. These results are displayed in Figure 10. The decomposition of $CT^{(2)}$(SM09) is qualitatively different from that of $CT^{(2)}$(Reg) . In the former we see two processes both apparently with the same exponent (except at short range). These do not reflect the tunneling processes described above. However, $CT^{(2)}$(Reg) exhibits at least two exponential processes: The acceptor to donor energy has a larger exponent and decays rapidly with separation. This is what we would expect as tunneling into the well-screened donor oxygen potential must proceed through a large barrier, leading to a large exponent.

The regularization procedure allows us to analyze the charge-transfer process in even more detail as we can selectively regularize the nuclear potential of the acceptor hydrogen atoms and isolate the charge-transfer contribution from the donor to the oxygen of the acceptor. From Figure 10 we see that, as might be expected, this is a much weaker process. What is unusual is the relatively small exponent associated with this process. The reason for this is as yet unclear.

In Figures 11 and 12 we present a similar analysis of the second-order charge-transfer energy for the HF dimer in its

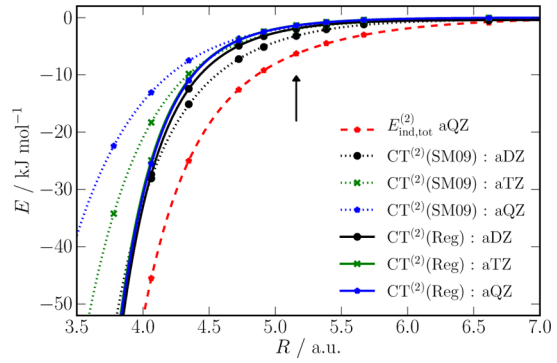


Figure 11. Second-order charge-transfer energies for the HF dimer in its hydrogen-bonded orientation. See the caption to Figure 9 for a description. The arrow indicates the equilibrium separation.

hydrogen-bonded orientation. The features we have highlighted for the water dimer are also seen here. As with the water dimer, $CT^{(2)}$(Reg) is largely basis-independent (the aug-cc-pVTZ/MC + and aug-cc-pVQZ/MC+ results are nearly identical, but the aug-cc-pVDZ/MC+ values differ) and interpolates between the $CT^{(2)}$(SM09) values calculated using the aug-cc-pVDZ basis (at short-range) and the aug-cc-pVQZ basis (at long-range). Also, as for the water dimer, the acceptor to donor second-order charge-transfer energy becomes appreciable only for short dimer separations (less than 5 au). This is quite different from the $CT^{(2)}$(SM09) energy for which we see acceptor to donor contributions at all separations.

Table 1 shows numerical values of the second-order induction and charge-transfer energies for various dimers.

dx.doi.org/10.1021/ct400704a | J. Chem. Theory Comput. 2013, 9, 5313−5326

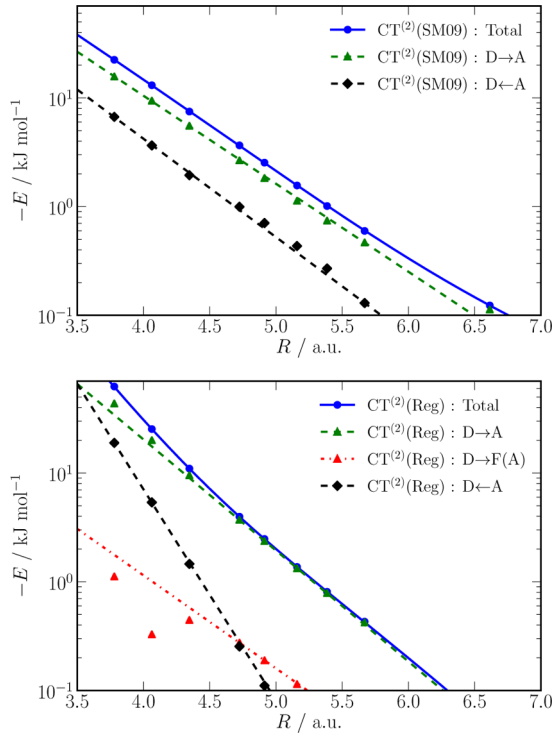


Figure 12. Decomposition of the second-order charge-transfer energies for the HF dimer in its hydrogen-bonded orientation. See the caption to Figure 10 for a description.

The results for the water and HF dimers quantify what is already displayed in the above figures: at the selected geometries, $CT^{(2)}$(Reg) and $CT^{(2)}$(SM09) agree reasonably well, but the differences get larger at shorter separations. With the exception of the H_3B complexes, the charge transfer is mainly from the donor to the acceptor; the reverse process (acceptor to donor) is much weaker. This is particularly true for $CT^{(2)}$(Reg); although the donor to acceptor energy is dominant for $CT^{(2)}$(SM09), the acceptor to donor energies can be significantly larger than those from $CT^{(2)}$(Reg). For the mixed HF and CO system, charge-transfer from the C to H in FH···CO is nearly 3 times as large as that from O to H in FH···OC, consistent with the strong σ-donor property of C in CO.

The H_3B complexes with CO and NH_3 are interesting as both of these have very short separations. These separations are short enough that one may question the use of an intermolecular perturbation theory like SAPT(DFT). Perhaps surprisingly, SAPT(DFT) interaction energies are within 5% of CCSD(T) energies for both complexes, with the agreement between the two best for the H_3B···NH_3 complex. The differences between $CT^{(2)}$(Reg) and $CT^{(2)}$(SM09) are quite large for both complexes. As discussed in Section II.1, the CT-SM09 definition should be expected to recover an ever smaller fraction of the true second-order CT as the intermolecular separation decreases. This seems to be the case here. While $CT^{(2)}$(SM09) correctly identifies the donor (NH_3, CO) to acceptor (H_3B) charge transfer as the larger quantity, the actual amount of CT from this method is substantially smaller than both $CT^{(2)}$(Reg) and the infinite-order results of Khaliullin et al.[3] Interestingly, while $CT^{(2)}$(Reg) is consistent with the Khaliullin results for the H_3B···NH_3 complex, the two sets of results differ qualitatively for the H_3B···CO complex. Here we would expect transfer from the CO to the H_3B to dominate. This is given by $CT^{(2)}$(Reg), but the ALMO method shows the opposite result.

Finally, the pyridine dimer is an interesting case as it exhibits a double hyhrogen bond between each of the N···H pairs in the D_{2h} symmetry, planar complex. The donor (N) to acceptor (H) bond length is the longest of the complexes discussed above. This leads to relatively small total induction energies and even smaller charge-transfer energies. The latter are just over 1/10 of the total induction energy—nearly an order of magnitude smaller than the charge-transfer energies in the other

Table 1. Second-Order Induction Energies and Charge-Transfer Energies for Several Systems at Specified Geometries[a]

	$E^{(2)}_{ind,tot}$		$CT^{(2)}$(Reg)		$CT^{(2)}$(SM09)		Khaliullin et al.[3]	
R, Bohr	D → A	D ← A	D → A	D ← A	D → A	D ← A	D → A	D ← A
Water dimer in hydrogen-bonded orientation								
R_{OH} = 3.67	−4.59	−1.25	−1.39	−0.10	−1.30	−0.46	–	–
R_{OH} = 3.29	−8.64	−1.79	−3.19	−0.39	−2.71	−0.94	–	–
HF dimer in hydrogen-bonded orientation								
R_{FH} = 3.44	−5.90	−0.37	−1.33	−0.04	−1.13	−0.43	–	–
FH···CO (linear)								
R_{HC} = 4.0	−6.24	−0.23	−1.39	−0.05	−1.28	−0.17	–	–
FH···OC (linear)								
R_{HO} = 4.0	−3.04	−0.02	−0.37	−0.02	−0.37	−0.13	–	–
H_3B···CO: aug-cc-PVTZ/DC: B3LYP optimized (linear)								
R_{BC} = 2.89	−296.24	−50.45	−139.16	−31.85	−20.16	−11.73	−123	−128
H_3B···NH_3: aug-cc-PVTZ/DC: B3LYP optimized								
R_{BN} = 3.21	−160.77	−14.83	−61.65	−9.84	−14.86	−7.23	−130	−11
Pyridine dimer: aug-cc-PVTZ/MC+: D_{2h} double H-bonded geometry								
R_{NH} = 4.82	−1.63	−1.63	−0.18	−0.18	−0.41	−0.41	–	–

[a]Unless otherwise stated, all calculations use the aug-cc-pVQZ basis set. The (electron) donor to (electron) acceptor (D → A) and acceptor to donor (D ← A) contributions are shown. For the water dimer, the longer R_{OH} distance of 3.67 au is representative of the O···H separation in the dimer and the shorter distance of 3.29 au is representative of the separation in clusters of water molecules. The HF dimer is at its equilibrium geometry, while the HF···CO dimers are both linear structures close to their radial equilibria. Both H_3B structures are optimized, and the pyridine dimer is in its doubly hydrogen-bonded planar, D_{2h} symmetry configuration. The results from Khaliullin et al. (ref 3) are infinite-order charge-transfer energies calculated using the 6-31(2+,2+)*G*(*df*,*pd*) basis set. All energies are reported in kJ mol^{-1} and distances in Bohr.

dx.doi.org/10.1021/ct400704a | J. Chem. Theory Comput. 2013, 9, 5313−5326

complexes. In this case the electron delocalization process works symmetrically in both directions; consequently there is no net charge transferred between the two pyridine molecules.

V.3. Polarization Models. Now that we have demonstrated both the numerical stability and the physical nature of $CT^{(2)}$(Reg), we are in a position to use this definition to determine the damping needed in polarization models. The basic idea is to fit the damping parameters so that the (noniterated) classical polarization model matches the true second-order polarization energies for a large number of dimers. In the following examples the classical polarization model was constructed using a rank 3 (3 on the heavy atoms and 1 on the hydrogen atoms) WSM distributed polarizability model[5,41] and a rank 4 (all atoms) distributed multipole model.[40] The models are damped using the Tang−Toennies damping functions[46] with isotropic (possibly site-pair-dependent) damping parameters. For numerical details see Section IV. Note that the damping coefficients obtained in this section depend on the details of the model; consequently, a direct comparison to other attempts to determine the damping coefficients[9] cannot be made.

Figure 13 shows three polarization models for the water dimer in its hydrogen-bonded orientation. These models differ

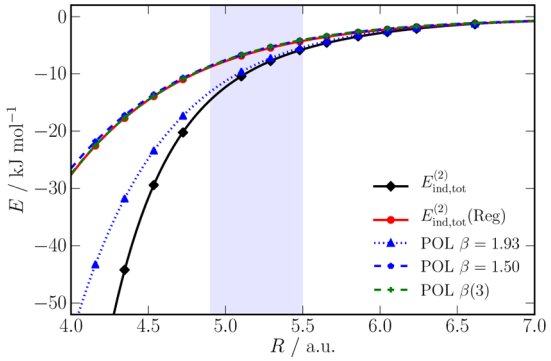


Figure 13. Second-order polarization energies for the water dimer in its hydrogen-bonded orientation. Reference polarization energies are $E^{(2)}_{ind,tot}$(Reg) with η = 3.0 au. Polarization models are as described in the text. Damping parameters are specified in atomic units. Damping models either use a single damping parameter for all site pairs or, for the $\beta(3)$ model, use parameters that vary with site-pair.

in their damping parameters only. The first uses a damping parameter derived from the expression

$$\beta = \sqrt{2I_A} + \sqrt{2I_B} \quad (13)$$

which was given by Misquitta and Stone[5] and is derived from the molecular vertical ionization energies I_A and I_B. In a later paper on dispersion models,[47] these authors demonstrated that this simple expression resulted in accurate damping models for a variety of systems, ranging from hydrogen-bonded complexes to van der Waals systems in a wide range of orientations. For the water dimer, using a vertical ionization energy of 0.4638 au,[48] using eq 13 we get β = 1.93 au. From Figure 13 we see that, as demonstrated in ref 5, the resulting polarization model agrees well with the *total, unregularized* energy $E^{(2)}_{ind,tot}$. The agreement is particularly good at the dimer equilibrium separation, though at shorter separations this model tends to *underestimate* $E^{(2)}_{ind,tot}$. Notwithstanding this seemingly good performance, there is considerable evidence from Sebetci and Beran[9] and also from our own numerical experiments that polarization models derived using eq 13 significantly overestimate the *many-body* polarization energy. By fitting to the many-body energies of clusters of water molecules, Sebetci and Beran obtain an optimized parameter of β_{opt} = 1.45 au; i.e., the damping needs to be considerably enhanced.

The reason for this is that the polarization model should reproduce the true polarization energy and not $E^{(2)}_{ind,tot}$. If the model is derived to match $E^{(2)}_{ind,tot}$, it will implicitly include, via the damping, some amount of the two-body charge-transfer energy. While this is, in itself, not a significant problem for the two-body energy, it can lead to large errors in the many-body polarization energy which will then include spurious two-body charge-transfer effects. This would lead to the polarization overbinding of clusters of molecules seen by Sebetci and Beran and discussed in Sections I and II of this article.

As discussed above, the solution to this problem is to determine the damping of the polarization model by fitting it to reproduce the true polarization energy only. If we do this using a single damping parameter, that is, the damping parameter is independent of type of sites, we get β = 1.50 au. From Figure 13 we see that this model results in a very good match with $E^{(2)}_{ind,tot}$(Reg), at least for the dimer in its hydrogen-bonded orientation. However, as can be seen in Figure 14, the

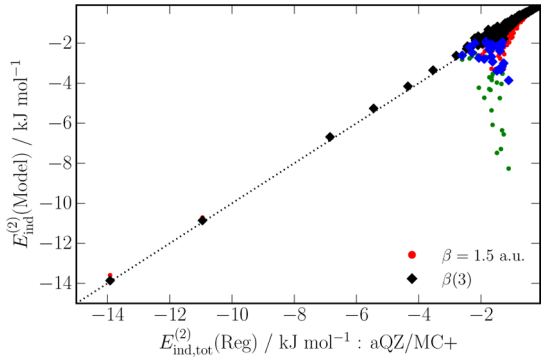


Figure 14. Second-order polarization energies for the water dimer in the 414 orientations described in the caption to Figure 4. Model second-order polarization energies are plotted against reference true polarization energies defined to be $E^{(2)}_{ind,tot}$(Reg) with η = 3.0 au calculated using the aug-cc-pVQZ/MC+ basis set. Polarization models are as described in the text. Damping parameters are specified in atomic units. Two damping models are used: the β = 1.5 au (red and green dots) model uses the same parameter for all site pairs and the $\beta(3)$ model (black and blue diamonds) uses parameters that vary with site-pair. Each set of points is divided into those with $E^{(2)}_{exch}$ > 40 kJ mol^{-1} (blue diamonds and green squares) and those with $E^{(2)}_{exch} \leq$ 40 kJ mol^{-1} (black diamonds and red squares). The former set includes mainly those configurations with close O···O contacts.

agreement is not as good for other orientations; in particular, at those with close O···O separations the polarization model tends to overestimate $E^{(2)}_{ind,tot}$(Reg) . A detailed examination of the water dimer polarization energies at orientations with close H···H and O···O contacts suggests that the damping parameter is strongly dependent on the types of the sites in the interacting pair. A much better fit to $E^{(2)}_{ind,tot}$(Reg) is obtained with a three parameter damping model, $\beta(3)$, in which β_{OH} = 1.61, β_{HH} = 1.80, and β_{OO} = 1.09 au. Of these, β_{HH} is not very well-defined and can be fixed to a relatively wide range of values. From Figures 13 and 14 we see that this model is significantly better than the simpler, one-parameter model with β = 1.5 au.

dx.doi.org/10.1021/ct400704a | J. Chem. Theory Comput. 2013, 9, 5313−5326

Even the three-parameter model exhibits somewhat larger errors for the dimers with close O···O contacts (highlighted in Figure 14). It is possible that a proper nonlinear optimization of the damping model will result in a model that removes these discrepancies, but it is also possible that at least some of the residual error is from the lack of angular dependence in the damping model. These issues are currently under investigation.

Although it may appear that the single-parameter model with $\beta = 1.50$ au matches the Sebetci and Beran value of $\beta_{opt} = 1.45$ au, matters are not as straightforward. First of all, Sebetci and Beran used a somewhat different multipole and polarizability models: their distributed multipole model had terms on the hydrogen atoms limited to rank 1, and in their WSM polarizability model the maximum rank was 2. Both of these, particularly the former, result in smaller polarization energies. Consequently, the damping need not be as large. Indeed, using a multipole model similar to theirs and by fitting to hydrogen-bonded dimer orientations only, we obtain a single parameter damping parameter of 1.6 au (a larger β means less damping). But these orientations are not fully representative of those found in water clusters such as those used by Sebetci and Beran. Here the contacts are more diverse, particularly in the more compact water clusters. We conjecture that the Sebetci and Beran value of $\beta_{opt} = 1.45$ au is a compromise that is the average of the site-pair-dependent parameters in model $\beta(3)$ described above. While this needs to be confirmed, we emphasize that all of these results are consistent: the polarization damping should be much smaller than what we would expect from eq 13.

The HF dimer exhibits many of these results. The polarization model based on eq 13 and a vertical ionization potential of 0.5669 au[48] results in a damping parameter of 2.2 au. From Figure 15 we see that this model closely matches

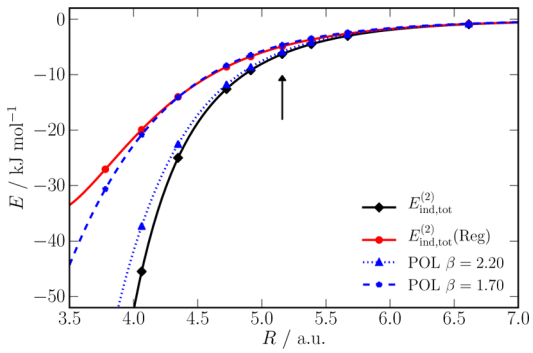


Figure 15. Second-order polarization energies for the HF dimer in its hydrogen-bonded orientation. For a description see the caption to Figure 13.

$E^{(2)}_{ind,tot}$; in this case the match is even better than for the water dimer. Our second model is obtained by fitting to the true polarization energy $E^{(2)}_{ind,tot}(Reg)$. This yields a single parameter model with $\beta = 1.70$ au, which, as we see from Figure 16, is able to match the $E^{(2)}_{ind,tot}(Reg)$ energies for other dimer orientations, albeit with a larger scatter than we had for the water dimer. At least some of this scatter is due to the unusual short-range mismatch in $E^{(2)}_{ind,tot}(Reg)$ and the polarization model seen in Figure 15. But there is also evidence that, as with the water dimer, an accurate polarization model would require three different damping parameters, with the F···F interaction

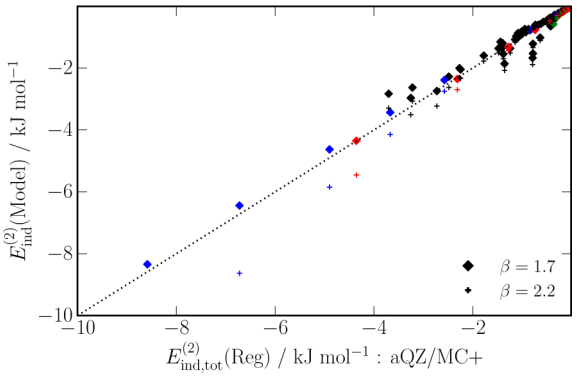


Figure 16. Second-order polarization energies for the HF dimer in the 159 orientations described in the caption to Figure 6. Results from two damping models are shown: $\beta = 2.2$ au (plus signs) and $\beta = 1.7$ au (diamonds). These data are color-coded as follows: dimers with F···H contacts (blue), H···H contacts (red), and F···F contacts (green).

needing a considerably stronger damping. We have yet to explore this possibility fully.

Curiously, Sebetci and Beran[9] found the original damping model with $\beta = 2.2$ au to be suitable for cyclic clusters of HF molecules. Exploration of this issue shows that with a multipole model similar to the one they used (with terms limited to rank 1 on the hydrogen atoms) we obtain an optimized damping parameter of 2.0 au by fitting to $E^{(2)}_{ind,tot}(Reg)$. This parameter should increase still more when the maximum rank of the polarization model is reduced to 2 from the maximum of 3 which we have used. Therefore, in this case too, we believe that our results are fully consistent with those from Sebetci and Beran and indicate that we are indeed now able to derive many-body polarization models from the dimer alone.

VI. DISCUSSION

We have presented a definition of the charge-transfer energy that is based on interpreting the transfer of charge between molecules through the process of tunneling. In this view, as illustrated in Figure 2, intermolecular charge transfer occurs by electron density tunneling into the screened nuclear potential of the partner monomer. This viewpoint leads to a simple way of defining the charge-transfer energy by *regularizing* the screened nuclear potential so as to suppress tunneling, while still allowing classical electromagnetic polarization.

The Gaussian-type regularization we have used involves a parameter η that has the dimensions L^{-2}, or, equivalently, the regularization introduces a nuclear screening length scale $\eta^{-1/2}$. This is the single most important parameter in this procedure. A priori the screening length scale is expected to depend on the atom type, but this dependence has been shown to be weak. Using a number of systems and two different procedures, we have demonstrated that the regularization parameter, η, is very close to 3.0 au, that is, the regularization occurs on a length-scale of 0.58 Bohr. This value of the parameter is suitable for systems involving lighter elements, but further work is needed to understand why this value is appropriate and how it may be expected to change for systems containing heavier elements.

Once the value of the regularization parameter has been fixed, the second-order charge-transfer energy, $CT^{(2)}(Reg)$, has a well-defined complete-basis-set limit for all intermolecular separations. This strongly contrasts with the definition put forward by Stone and Misquitta[2] which exhibits a strong basis dependence, particularly at short separations.

With the proposed definition of the charge transfer through regularization, the charge-transfer energy for the water dimer and HF dimer systems exhibits a clear double exponential character. This has been shown to stem from the strong asymmetry in the electron donor to acceptor and electron acceptor to donor components of the $CT^{(2)}$(Reg) energies: the donor to acceptor process is not only dominant but decays more slowly with separation, indicative of a tunneling process through a smaller barrier. Additionally, using a partial regularization we have been able to decompose the donor to acceptor charge-transfer energy into a primary process from the electron-rich donor O or F atom into the electron deficient acceptor H atom and a much smaller secondary process from the donor O or F into the acceptor O or F. The secondary process is an order of magnitude smaller than the primary donor to acceptor process.

This new definition of the second-order charge-transfer energy has been used on a variety of systems, including some with very strong hydrogen bonds ($H_3B\cdots CO$ and $H_3B\cdots NH_3$) and one doubly hydrogen-bonded system (the pyridine dimer in its D_{2h} planar configuration). In all cases the computed charge-transfer energy makes physical sense. The differences between $CT^{(2)}$(Reg) and other definitions such as the Stone and Misquitta and ALMO method from Khaliullin et al.[3] are particularly dramatic for the most strongly bound H_3B complexes for which only $CT^{(2)}$(Reg) results in physically meaningful energies. In the doubly hydrogen-bonded pyridine example, while there is a stabilization due to the tunneling, due to symmetry there is no net charge transferred between the two molecules: they remain neutral. So, in a sense, the term "charge-transfer" is a misnomer.

Finally, we have used regularization to suppress all charge-transfer contributions from the second-order induction energy thereby defining a pure second-order polarization energy against which we have fitted the damping in the WSM polarization models. Comparisons against data from Sebetci and Beran[9] indicate that the resulting models are able to describe the many-body polarization energies accurately. This is a major step forward in our ability to model the major part of the many-body nonadditive energy in polarizable systems from calculations on the dimer alone.

On the basis of the arguments put forward in the Introduction and the many examples provided herein we now propose a new definition of the charge-transfer energy:

The process of charge transfer can be thought of as a charge delocalization through electron tunneling into the screened nuclear potentials of the partner monomers. The resulting energy of stabilization is the intermolecular charge-transfer or delocalization energy.

We suggest that it may be conceptually advantageous to term this the *delocalization* energy and reserve the term "charge transfer" for the phenomenon of charge-transfer excitations studied by Mulliken.[49]

Though the interpretation of the CT as a tunneling process may appear different from the usual picture of the CT as an excitation from a donor orbital to an acceptor located on the partner monomer, we believe that these two viewpoints are consistent. The second-order Rayleigh−Schrödinger perturbation expression for the second-order induction energy given in eq 2 contains terms that arise from excited states that have a significant component at the sites of the partner monomer. For these states to make a significant contribution to the induction energy, the matrix elements in the numerator need to be significant (and the energy differences in the denominator small). One way this can happen is when the excited state has significant contributions in the regions of the nuclei of the partners where the screened potential ω^B is significant. This is just another way of describing tunneling states.

Charge transfer as incipient chemical bonding: if we accept the physical picture of charge transfer as a tunnelling of electrons into the screened nuclear potentials of the partner monomer, then we must also accept the view that charge transfer is a form of incipient covalent bonding.

There are several advantages to the proposed definition of the charge-transfer energy:

- It leads to a physically appealing definition of CT.
- The resulting (electron) acceptor to donor and (electron) donor to acceptor contributions make physical sense.
- It leads to polarization models that are applicable to many-body systems although they are based on dimer calculations alone.
- The method is implemented as part of regular SAPT-(DFT) calculation.
- The results are independent of basis set, and true basis-set convergence can be achieved.
- The method is independent of the type of basis set: it could be used, for example, with plane-wave basis sets. Therefore it should be possible to use this definition to calculate the charge-transfer, or delocalization, energy, in a variety of electronic structure programs.

Among the issues that need resolving is our incomplete understanding of the origin of the regularization length scale and how it depends on atom type and the manner in which the present procedure needs to be extended to calculate charge-transfer effects beyond second order in perturbation theory. Both of these issues are under current investigation.

■ ASSOCIATED CONTENT

Supporting Information

Additional data and details. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: a.j.misquitta@qmul.ac.uk.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The author would like to acknowledge Prof. Anthony J. Stone and Prof. Kenneth Jordan for many stimulating and fruitful discussions, and in particular, Prof. Stone for many useful comments and suggestions and a stimulating collaboration. Additionally, the author thanks the School of Physics at Queen Mary, University of London for support and the Thomas Young Centre for providing a stimulating research platform in the London area.

■ REFERENCES

(1) Stone, A. J. Computation of charge-transfer energies by perturbation theory. *Chem. Phys. Lett.* **1993**, *211*, 101−109.

(2) Stone, A. J.; Misquitta, A. J. Charge-transfer in symmetry-adapted perturbation theory. *Chem. Phys. Lett.* **2009**, *473*, 201−205.

(3) Khaliullin, R. Z.; Bell, A. T.; Head-Gordon, M. Analysis of charge transfer effects in molecular complexes based on absolutely localized molecular orbitals. *J. Chem. Phys.* **2008**, *128*, 184112−16.
(4) Cohen, A. J.; Mori-Sánchez, P.; Yang, W. Insights into current limitations of density functional theory. *Science* **2008**, *321*, 792−794.
(5) Misquitta, A. J.; Stone, A. J. Accurate induction energies for small organic molecules: I. theory. *J. Chem. Theory Comput.* **2008**, *4*, 7−18.
(6) Williams, H. L.; Mas, E. M.; Szalewicz, K.; Jeziorski, B. On the effectiveness of monomer-, dimer-, and bond-centered basis functions in calculations of intermolecular interaction energies. *J. Chem. Phys.* **1995**, *103*, 7374−7391.
(7) Mas, E. M.; Szalewicz, K. Effects of monomer geometry and basis set saturation on computed depth of water dimer potential. *J. Chem. Phys.* **1996**, *104*, 7606−7614.
(8) Mas, E. M.; Szalewicz, K.; Bukowski, R.; Jeziorski, B. Pair potential for water from symmetry-adapted perturbation theory. *J. Chem. Phys.* **1997**, *107*, 4207−4218.
(9) Sebetci, A.; Beran, G. J. O. Spatially homogeneous qm/mm for systems of interacting molecules with on-the-fly ab initio force-field parametrization. *J. Chem. Theory Comput.* **2010**, *6*, 155−167.
(10) Schenter, G. K.; Glendening, E. D. Natural energy decomposition analysis: The linear response electrical self energy. *J. Phys. Chem.* **1996**, *100*, 17152−17156.
(11) Azar, R. J.; Horn, P. R.; Sundstrom, E. J.; Head-Gordon, M. Useful lower limits to polarization contributions to intermolecular interactions using a minimal basis of localized orthogonal orbitals: Theory and analysis of the water dimer. *J. Chem. Phys.* **2013**, *138*, 084102−14.
(12) Jeziorski, B.; Szalewicz, K. Symmetry-adapted perturbation theory. In *Handbook of Molecular Physics and Quantum Chemistry*; Wilson, S., Ed.; Wiley: 2002; Vol. *8*, pp 37−83.
(13) Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation theory approach to intermolecular potential energy surfaces of van der waals complexes. *Chem. Rev.* **1994**, *94*, 1887−1930.
(14) Patkowski, K.; Korona, T.; Jeziorski, B. Convergence behavior of the symmetry-adapted perturbation theory for states submerged in pauli forbidden continuum. *J. Chem. Phys.* **2001**, *115*, 1137−1152.
(15) Patkowski, K.; Jeziorski, B.; Szalewicz, K. Unified treatment of chemical and van der waals forces via symmetry-adapted perturbation expansion. *J. Chem. Phys.* **2004**, *120*, 6849−6862.
(16) Claverie, P. Theory of intermolecular forces. i. on the inadequacy of the usual rayleigh-schrödinger perturbation method for the treatment of intermolecular forces. *Int. J. Quantum Chem.* **1971**, *5*, 273−296.
(17) Kutzelnigg, W. The primitive wave function in the theory of intermolecular interactions. *J. Chem. Phys.* **1980**, *73*, 343(17).
(18) Morgan, J. D., III; Simon, B. Behavior of molecular potential energy curves for large nuclear separations. *Int. J. Quantum Chem.* **1980**, *17*, 1143.
(19) Adams, W. H. Definition of eigenproblems suited to intermolecular perturbation theory. *Int. J. Quantum Chem.* **1999**, *72*, 393−404.
(20) Herring, C. Critique of the heitler−london method of calculating spin couplings at large distances. *Rev. Mod. Phys.* **1962**, *34*, 631−645.
(21) Patkowski, K.; Jeziorski, B.; Szalewicz, K. Symmetry-adapted perturbation theory with regularized coulomb potential. *J. Mol. Struct. (Theochem)* **2001**, *547*, 293−307.
(22) Adams, W. H. Two new symmetry-adapted perturbation theories for the calculation of intermolecular interaction energies. *Theor. Chim. Acta* **2002**, *108*, 225−231.
(23) Jeziorski, B.; Szalewicz, K. *Encyclopedia of Computational Chemistry*; Wiley: Chichester, U.K., 1998; Vol. 2, p 1376.
(24) Misquitta, A. J.; Podeszwa, R.; Jeziorski, B.; Szalewicz, K. Intermolecular potentials based on symmetry-adapted perturbation theory with dispersion energies from time-dependent density-functional theory. *J. Chem. Phys.* **2005**, *123*, 214103.
(25) Podeszwa, P.; Bukowski, R.; Szalewicz, K. Potential energy surface for the benzene dimer and perturbational analysis of $\pi - \pi$ interactions. *J. Phys. Chem. A* **2006**, *110*, 10345−10354.
(26) Bukowski, R.; Szalewicz, K.; Groenenboom, G.; van der Avoird, A. Interaction potential for water dimer from symmetry-adapted perturbation theory based on density functional description of monomers. *J. Chem. Phys.* **2006**, *125*, 044301.
(27) Patkowski, K.; Szalewicz, K.; Jeziorski, B. Induction and exchange-induction energy with a regularized coulomb potential. Private communication.
(28) Casida, M. E. Time-dependent density-functional response theory for molecules. In *Recent Advances in Density-Functional Theory*; Chong, D. P., Ed.; World Scientific: 1995; p 155.
(29) Colwell, S. M.; Handy, N. C.; Lee, A. M. Determination of frequency-dependent polarizabilities using current density-functional theory. *Phys. Rev. A* **1995**, *53*, 1316−1322.
(30) Jeziorski, B.; Moszynski, R.; Ratkiewicz, A.; Rybak, S.; Szalewicz, K.; Williams, H. *SAPT: A program for many-body symmetry-adapted perturbation theory calculations of intermolecular interaction energies*; STEF: Cagliari, 1993; Vol. *B*, p 79.
(31) Misquitta, A. J.; Stone, A. J. *CAMCASP: a program for studying intermolecular interactions and for the calculation of molecular properties in distributed form*; University of Cambridge: 2013. http://www-stone.ch.cam.ac.uk/programs.html#CamCASP (accessed Oct. 2013).
(32) Helgaker, T.; Jensen, H. J. A.; Joergensen, P.; Olsen, J.; Ruud, K.; Aagren, H.; Auer, A.; Bak, K.; Bakken, V.; Christiansen, O.; Coriani, S.; Dahle, P.; Dalskov, E. K.; Enevoldsen, T.; Fernandez, B.; Haettig, C.; Hald, K.; Halkier, A.; Heiberg, H.; Hettema, H.; Jonsson, D.; Kirpekar, S.; Kobayashi, R.; Koch, H.; Mikkelsen, K. V.; Norman, P.; Packer, M. J.; Pedersen, T. B.; Ruden, T. A.; Sanchez, A.; Saue, T.; Sauer, S. P. A.; Schimmelpfennig, B.; Sylvester-Hvid, K. O.; Taylor, P. R.; Vahtras, O. *Dalton, a molecular electronic structure program*, Release 2.0; 2005. See http://www.kjemi.uio.no/software/dalton/dalton.html (accessed Oct. 2013).
(33) Bukowski, R.; Cencek, W.; Jankowski, P.; Jeziorski, B.; Jeziorska, M.; Lotrich, V.; Kucharski, S.; Misquitta, A. J.; Moszynski, R.; Patkowski, K.; Podeszwa, R.; Rybak, S.; Szalewicz, K.; Williams, H.; Wheatley, R. J.; Wormer, P. E. S.; Zuchowski, P. S. *SAPT2008: an ab initio program for many-body symmetry-adapted perturbation theory calculations of intermolecular interaction energies*; University of Delaware and University of Warsaw: 2008 (accessed Oct. 2013).
(34) Perdew, J.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868. Also see erratum, *Phys. Rev. Lett.* **1996**, *78*, 1396.
(35) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The pbe0 model. *J. Chem. Phys.* **1999**, *110*, 6158−6170.
(36) Fermi, E.; Amaldi, E. Le orbite [infinito] s degli elementi. *Memorie della classe di scienze fisiche della Reale Accademia d'Italia* **1934**, *6*, 117−149.
(37) Tozer, D. J.; Handy, N. C. Improving virtual Kohn−Sham orbitals and eigenvalues: Application to excitation energies and static polarizabilities. *J. Chem. Phys.* **1998**, *109*, 10180−10189.
(38) Hesselmann, A.; Jansen, G.; Schütz, M. Density-functional theory-symmetry-adapted intermolecular perturbation theory with density fitting: A new efficient method to study intermolecular interaction energies. *J. Chem. Phys.* **2005**, *122*, 014103.
(39) Podeszwa, R.; Bukowski, R.; Szalewicz, K. Density-fitting method in symmetry-adapted perturbation theory based on kohn-sham description of monomers. *J. Chem. Theory Comput.* **2006**, *2*, 400−412.
(40) Stone, A. J. Distributed multipole analysis: Stability for large basis sets. *J. Chem. Theory Comput.* **2005**, *1*, 1128−1132.
(41) Misquitta, A. J.; Stone, A. J.; Price, S. L. Accurate induction energies for small organic molecules: Ii. models and numerical details. *J. Chem. Theory Comput.* **2008**, *4*, 19−32.
(42) Stone, A. J.; Dullweber, A.; Engkvist, O.; Fraschini, E.; Hodges, M. P.; Meredith, A. W.; Nutt, D. R.; Popelier, P. L. A.; Wales, D. J. *Orient: a program for studying interactions between molecules*, version 4.8;

University of Cambridge: 2013. http://www-stone.ch.cam.ac.uk/programs.html#Orient (accessed Oct. 2013).
(43) Misquitta, A. J.; Welch, G. W. A.; Stone, A. J.; Price, S. L. A first principles prediction of the crystal structure of $C_6Br_2ClFH_2$. *Chem. Phys. Lett.* **2008**, *456*, 105−109.
(44) Peterson, K. A.; T. H. Dunning, J. Benchmark calculations with correlated molecular wave functions vii. binding energy and structure of the hf dimer. *J. Chem. Phys.* **1995**, *102*, 2032.
(45) Patkowski, K.; Murdachaew, G.; Fou, C.-M.; Szalewicz, K. Accurate *ab initio* potential for argon dimer including highly repulsive region. *Mol. Phys.* **2005**, *103*, 2031−2045.
(46) Tang, K. T.; Toennies, J. P. An improved simple model for the van der waals potential based on universal damping functions for the dispersion coefficients. *J. Chem. Phys.* **1984**, *80*, 3726−3741.
(47) Misquitta, A. J.; Stone, A. J. Dispersion energies for small organic molecules: first row atoms. *Mol. Phys.* **2008**, *106*, 1631−1643.
(48) Lias, S. G. Ionization energy evaluation. In *NIST Chemistry Webbook, Nist Standard Reference Database Number 69*; Mallard, W. G.,Linstrom, P. J., Eds.; Gaithersburg, 2000. http://webbook.nist.gov (accessed Oct. 2013).
(49) Mulliken, R. S. Molecular compounds and their spectra. *J. Am. Chem. Soc.* **1952**, *74*, 811−824.