THE JOURNAL OF PHYSICAL CHEMISTRY A

Article

pubs.acs.org/JPCA

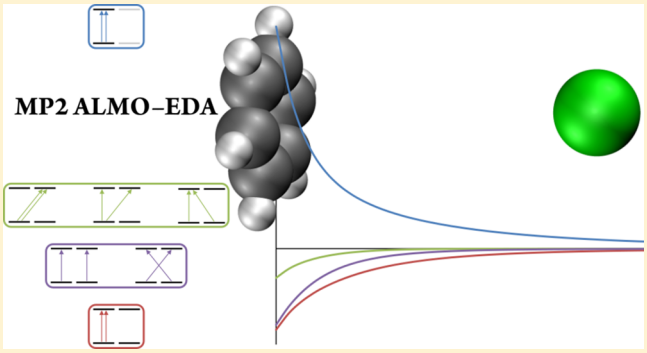
Efficient Implementation of Energy Decomposition Analysis for Second-Order Møller–Plesset Perturbation Theory and Application to Anion–π Interactions

Published as part of The Journal of Physical Chemistry virtual special issue "Mark S. Gordon Festschrift".

Jonathan Thirman* and Martin Head-Gordon*

Kenneth S. Pitzer Center for Theoretical Chemistry, Department of Chemistry, University of California, Berkeley, Berkeley, California 94720, United States

ABSTRACT: Energy decomposition analysis (EDA) is a widely used tool for extracting physical and chemical insights from electronic structure calculations of intermolecular interactions, as well as for the development of advanced force fields for describing those interactions. Recently, the absolutely localized molecular orbital (ALMO) EDA has been extended from the self-consistent field level to the second-order Møller–Plesset (MP2) theory level. This paper reports an efficient implementation of the MP2 ALMO-EDA that scales optimally, employs the resolution of the identity (RI) approximation for post-SCF matrix elements, and is shared-memory parallel. The algorithms necessary to achieve this implementation are described in detail. Performance tests using the aug-cc-pVTZ basis set for water clusters of up to 10 molecules are reported. The timings suggest that the MP2 ALMO-EDA is computationally feasible whenever MP2 energy calculations themselves are feasible, and the cost is dominated by the SCF itself in this size regime. The MP2 ALMO-EDA is applied to study the origin of substituent effects in anion–π interactions between chloride and benzene and mono- through hexafluorobenzene. The effect of fluoro substituents was primarily to change the frozen interaction. Detailed analysis supports the interpretation that anion–π interactions are favorable because of electrostatic interaction with the substituents.



1. INTRODUCTION

Second-order Møller–Plesset (MP2) theory[1] is perhaps the simplest wave function theory that includes the effects of dynamic electron correlation.[2] MP2 thus corrects the self-consistent field (SCF) Hartree–Fock (HF) molecular orbital model with missing physics that includes reduced ionic character in the electron density and thus reduced dipole moments,[3−6] and dispersion interactions[7,8] that provide a universal long-range attraction. Both of these effects are important for intermolecular interactions, and therefore MP2 theory provides a great improvement over HF theory in the accuracy obtained for ion–water interactions,[9] hydrogen bonds,[10] electrostatically bound complexes,[11,12] and dispersion-bound systems.[11] MP2 tends to overestimate dispersion interactions, particularly in stacking interactions of aromatic systems, as is well-known.[11,13,14] However, despite its simplicity, MP2 theory often matches more expensive methods such as coupled cluster singles and doubles (CCSD) for hydrogen-bonded systems such as water clusters.[15] Triples corrections, such as CCSD(T) are needed to reliably outperform MP2 theory on such systems, although spin-component scaling is a successful semiempirical alternative,[16−19] and even attenuated MP2[20,21] is valuable in smaller basis sets. MP2 theory also outperforms almost all density functionals for ion–water interactions, particularly for multiply charged ions.[22]

Because of their usefulness, as well as their low computational cost compared to that of CCSD, CCSD(T), etc., it is desirable to be able to analyze the results of MP2 calculations of intermolecular interactions. This is the task of an energy decomposition analysis (EDA). An EDA takes a calculated intermolecular binding energy and ascribes portions of it to different physical contributions such as permanent and induced electrostatics, dispersion, Pauli repulsions, and charge transfer.[23−26] Most EDAs that have been proposed and widely used to date apply to SCF wave functions, such as HF and/or Kohn–Sham DFT.[27] Although it is beyond our scope to review them in any detail, it is worth mentioning that the Kitaura–Morokuma (KM) decomposition[28] is perhaps the seminal method, followed soon afterward by the Ziegler–Rauk approach.[29] More recent significant developments have been the separation of polarization from charge transfer using

Received: November 15, 2016
Revised: December 21, 2016
Published: December 22, 2016

DOI: 10.1021/acs.jpca.6b11516
J. Phys. Chem. A 2017, 121, 717−728

fragment-blocked molecular orbital coefficients,[25,30−33] the development of density-based EDA approaches for DFT,[34−37] and very recent efforts to obtain stable basis set limits for polarization and charge transfer.[38−40] It is noteworthy that application of the KM EDA[41] has helped to inform the development of advanced force field approaches, such as the effective fragment potential method.[42−44]

There has been much less effort devoted to developing EDAs that unravel the contributions to post-HF binding energies, as given by MP2 or CC methods. The very simplest model is to ascribe the post-HF correlation binding in an intermolecular interaction to dispersion, as has sometimes been advocated.[45] However, because MP2 reduces the ionicity of HF charge distributions, the electrostatics are also affected, and thus the correlation contribution can even be net repulsive.[12] Local correlation models[46,47] are a good basis for identifying correlation contributions to specific intermolecular interactions,[48,49] with efforts along these lines continuing to the present day.[50] Local correlation models are quite naturally able to separate intermolecular correlation effects that are associated with dispersion from those that correspond to charge transfer. Difficulty arises in distinguishing post-HF contributions to polarization from charge transfer.

Nonetheless, we feel that an important criterion for a successful post-HF EDA is to organize the correlation contributions to an intermolecular interaction so as to correspond directly to the terms that are well-accepted at the SCF level. We recently presented[51] the theory and a pilot implementation for an MP2 EDA that meets this criterion for the first time, to our knowledge. This approach employs the fragment-blocked absolutely localized molecular orbitals (ALMOs) that were used in the earlier HF-level ALMO-EDA,[31,33] rather than the virtual orbitals derived from linear response[39] to electric fields and gradients that are used in the recent second generation ALMO-EDA[40] to ensure a valid complete basis set limit for polarization. The MP2 ALMO-EDA identifies the correlation contributions to the so-called frozen interaction (permanent electrostatics + Pauli repulsion), induced electrostatics, and charge transfer. In addition, using local correlation arguments,[47] a dispersion term is defined. The main purpose of this paper is to report an efficient production-level implementation of the MP2 ALMO-EDA, which will enable its application to any system for which MP2 calculations are themselves feasible.

The outline of the remainder of the paper is as follows. In section 2, the theory defining the MP2 ALMO-EDA is briefly reviewed, to clearly define all relevant terms. In section 3, we then address the critical issues needed to achieve an efficient implementation and discuss the way in which we have chosen to address them. To assess the success of the implementation, a series of benchmark timing tests are reported in section 4, including the question of how well our OpenMP parallelization performs. We then turn to an example of the application of the new MP2 ALMO-EDA implementation in section 5, which is the question of the origin of substituent effects in anion−π interactions. Anion−π interactions[52] have become a topic of considerable interest over recent years, as it has become evident that such interactions are not weak and are relevant in systems ranging from biomolecules to supramolecular complexes.[53−56] Wheeler and Houk have previously analyzed substituent effects in halide-aryl interactions,[57] and using high level calculations together with calculated electrostatic potentials, they concluded that direct through-space interactions of the anion with the substituents are the dominant factor controlling substituent effects.

2. BRIEF REVIEW OF THE MP2 EDA

The MP2 EDA is an extension of the HF version of the ALMO-EDA. At the HF level, the EDA uses ALMOs to create two constrained intermediate Slater determinants and defines its three components as the differences between these, the isolated system energies, and the full system energy. The first intermediate is the initial wave function. It is simply the HF solutions for the individual molecules as isolated systems joined into a single Slater determinant. The difference between its energy and the sum of the energies of isolated systems is the frozen term, describing the permanent electrostatics and Pauli repulsion. The second intermediate is the SCF for molecular interactions (SCF-MI) solution.[58−61] This is produced by finding the minimum energy solution subject to the constraint that molecular orbitals belonging to one molecule do not include any atomic orbital basis functions belonging to any other molecule. This produces the ALMOs that are key to the method. The difference between the SCF-MI energy and the initial energy is the polarization term, which describes induced electrostatics. Finally, the difference between the full system HF energy and the SCF-MI energy is the charge transfer term. Therefore, the three terms sum to the supermolecule binding energy by construction.

The MP2 version of the ALMO-EDA also decomposes the correlation contribution to the binding energy, by defining a MP2 to correction to each of the HF components and adding another one for dispersion. It does so through the definition an MP2 version of each energy, along with one additional intermediate to describe dispersion. The equivalent of the initial energy is referred to as E(frag/frz). It is complicated by the need for all virtual orbitals to be orthogonal to all occupied orbitals. This requires a projection scheme. The MP2 amplitudes from the isolated systems are frozen and used for the energy in the presence of the full system. The standard MP2 energy formula only applies to the MP2 solution. Because the presence of the other molecules is a one-electron perturbation, it is necessary to use the Hylleraas functional[62] for the MP2 energy. Additionally, the fact that the amplitudes of the MP2 solution are not variationally optimized parameters means that MP2 does not meet the criteria for the Hellmann−Feynman theorem,[63] so the relaxed density matrix[64] must be used in the Hylleraas functional in place of the ordinary "unrelaxed" MP2 response density matrix.

Both of the remaining intermediates are calculated by constrained MP2 using the ALMO basis. The ALMO basis allows each occupied and virtual orbital to be assigned to a particular molecule (though the ALMO virtuals must be projected into the complement of the occupied space). The first ALMO intermediate, E(frag/ALMO), is analogous to the SCF-MI energy. It is the MP2 solution when constrained to only allow double excitations entirely on one molecule. This allows the system to polarize at the correlation level but does not allow dispersion-like excitations. As a function of the number of fragments, N, this contains only $O(N)$ excitations. The second ALMO intermediate is E_{ccc}(sys/ALMO), or the charge conserving correlation (CCC) energy. This is the MP2 solution when constrained to allow double excitations involving two molecules, but only if they do not change the number of electrons associated with each molecule. CCCMP2 allows dispersion-like excitations, but not transfer of charge. It is a

DOI: 10.1021/acs.jpca.6b11516
J. Phys. Chem. A 2017, 121, 717−728

fragment analog of the "Diatomics-in-Molecules" (DIM) local correlation model[47,65] and includes $O(N^2)$ excitations. The difference between the two ALMO intermediates is the dispersion term of the EDA. The other three terms are also the differences between successive energies, the same as their HF counterparts. For the final energy decomposition, the HF and MP2 contributions to each term are summed together to produce four terms that decompose the MP2 binding energy.

3. IMPLEMENTATION

The terms of the EDA to be calculated are as follows.

$$\Delta E_{\mathrm{CT}} = E(\mathrm{sys}) - E_{\mathrm{ccc}}(\mathrm{sys/ALMO}) - (\mathrm{BSSE} - \mathrm{ABSSE}) \tag{1a}$$

$$\Delta E_{\mathrm{disp}} = E_{\mathrm{ccc}}(\mathrm{sys/ALMO}) - \sum_{\mathrm{frag}} E(\mathrm{frag/ALMO}) \tag{1b}$$

$$\Delta E_{\mathrm{pol}} = \sum_{\mathrm{frag}} E(\mathrm{frag/ALMO}) - E(\mathrm{frag/frz}) \tag{1c}$$

$$\Delta E_{\mathrm{frz}} = \sum_{\mathrm{frag}} E(\mathrm{frag/frz}) - E(\mathrm{frag}) - \mathrm{ABSSE} \tag{1d}$$

Several terms, E(sys), E(frag), and BSSE are calculated exactly as in a standard counterpoise corrected binding energy job and need not be discussed here. (If decomposing a binding energy that is not counterpoise corrected, the BSSE term is simply omitted, but ABSSE must still be included.) The remaining terms can be broken into two groups based on which molecular orbital basis is used. The first uses the modified ALMO basis and includes E_{ccc}(sys/ALMO) and E(frag/ALMO), whereas the second uses the isolated fragment basis and includes E(frag/frz) and ABSSE. ABSSE is the difference between E(frag) calculated with the full system's auxiliary functions and calculated with only the particular fragment's auxiliary functions. The latter is computed as part of a normal binding energy calculation, but the former must also be computed for the EDA.

3.1. ALMO Terms. Calculation begins by preparing the common basis for the two ALMO terms. As input, the calculation requires the polarized ALMO solution for the system as well as the corresponding Fock matrix. However, the ALMO basis is modified to be suitable for MP2. The full system occupied space is symmetrically orthogonalized. The new occupied orbitals are used to create an orthogonal projector and each virtual orbital is projected into the orthogonal complement of the occupied space. Then, the resulting virtual orbitals are orthonormalized within each fragment's virtual space, but not between fragments. The Fock matrix is transformed to this basis. As a final transformation, the system is canonicalized on each fragment. To do so, on each fragment, both the occupied−occupied and virtual−virtual blocks of the Fock matrix are diagonalized and the resulting eigenvectors are used to transform the corresponding ALMOs. This results in a new basis where the Fock matrix is diagonal within the fragment o−o and v−v blocks. These preparation steps take a negligible amount of time (both practically and asymptotically) compared to later steps, so further detail is not necessary. The new ALMO basis has the required property of strong orthogonality between the full system occupied and virtual spaces as well as a desirable amount of sparsity in both the overlap and Fock matrices. The overlap matrix is identity in the o−o block, zero in the o−v block, and identity in each fragment's v−v block, but it has no special properties in the cross-fragment parts of the v−v block. The Fock matrix is diagonal on each fragment's o−o block and v−v block but has no special properties in the cross-fragment parts of the blocks. The o−v part is not relevant to MP2. This sparsity allows faster calculation of E_{ccc}(sys/ALMO).

The three-center resolution-of-the-identity integrals must also be transformed. This step is the only one that scales as N^4 with an increasing number of identical fragments, so it is asymptotically dominant, though for most practical systems the actual construction of the four center electron repulsion integrals will take longer despite scaling as N^3. Calculation proceeds the same as in standard RIMP2, through the creation of a **B** matrix such that $\sum_P B^P_{ia} B^P_{jb} \approx (ia|jb)$. Unfortunately, the locality of CCCMP2 (MP2 with only charge conserving correlation) is not enough to allow creation of a smaller B matrix. Integrals of the form $(ia|jb)$ and $(ib|ja)$ are needed where i and a share a fragment as do j and b. For the first type of integral, only the entries in the on-fragment blocks of the B matrix are required. However, the second type has no such restriction and requires the construction of the entire B matrix. There are two steps to transforming the AO basis three center integrals into the B matrix. The first is transforming from the atomic orbital basis $(mn|P)$ to the ALMO basis $(ia|P)$. This can be done in two BLAS calls per auxiliary function. Without taking advantage of sparsity, this requires time $O(n^2oxN^4)$ where n in the number of basis functions, o is the number of occupied orbitals, v is the number of virtual orbitals, and x is the number of auxiliary functions on a fragment and N is the number of fragments. However, for a sufficiently large system, many integrals are zero. The next step is the multiplication by the $(P|Q)^{-1/2}$ matrix, which can be done in a single BLAS call and takes time $O(ovx^2N^4)$. As $x > n$ for the RI approximation to be accurate, this is the slower of the two steps. In addition to the normal **B** matrix, a fragment blocked **B** matrix is created where only entries $(ia|P)$ with i and a on the same fragment are included.

With this, the preliminary work is complete. The first thing to calculate is E(frag/ALMO) for each fragment. Because the basis is orthonormal and canonical on each fragment, this proceeds exactly as in normal RIMP2 with the only exception being that the t amplitudes are saved. The required integrals can be constructed easily from the fragment blocked B matrix.

The most complex step is the calculation of E_{ccc}(sys/ALMO). The basis is neither globally orthonormal nor canonical, so the standard MP2 methods cannot be used. Though we may use the energy equation,

$$E = -\sum_{ijab} t^{ab}_{ij} (ia|jb) \tag{2}$$

to get t^{ab}_{ij}, we must solve the linear system

$$\sum_{ijab} \Delta^{aba'b'}_{iji'j'} \cdot t^{ab}_{ij} = 2(i'a'|j'b') - (i'b'|j'a') \tag{3}$$

with Δ as an eighth-rank tensor defined as

$$\Delta^{aba'b'}_{iji'j'} = -f_{ii'} g_{aa'} g_{jj'} g_{bb'} + g_{ii'} f_{aa'} g_{jj'} g_{bb'} - g_{ii'} g_{aa'} f_{jj'} g_{bb'} + g_{ii'} g_{aa'} g_{jj'} f_{bb'} \tag{4}$$

Equation 2 is simple to compute as $(ia|jb)$ can be formed from the fragment blocked **B** matrix with a single BLAS call in time $O(o^2v^2xN^3)$. Equation 3 is more interesting and has two

DOI: 10.1021/acs.jpca.6b11516
J. Phys. Chem. A 2017, 121, 717−728

challenging pieces. One is calculating the right-hand side and the other is solving the equation.

3.1.1. Electron Repulsion Integrals. Calculating the right-hand side is challenging because of the $(ib|ja)$ term. This presents more of a challenge here than in ordinary MP2 because the fragment locality does not directly help. In ordinary MP2, every $(ib|ja)$ integral is also a $(ia|jb)$ integral, so they do not need to be calculated separately. However, when i/a are on a different fragment from j/b, $(ib|ja)$ does not appear as a $(ia|jb)$ integral in CCCMP2. Though calculating them has the same $O(o^2v^2xN^3)$ scaling as the other integrals, poor locality means that they cannot be computed as efficiently. Calculating the $(ib|ja)$ integrals involves all ovN^2 entries of the **B** matrix each used in ov different integrals, whereas the $(ia|jb)$ integrals involve only ovN entries of the B matrix used in each ovN different integrals.

The simplest way to compute these integrals is best, with some modifications to prevent duplicated work. A procedure to compute $\mathbb{I}_{ij}^{ab} = 2(ia|jb) - (ib|ja)$ appears in Figure 1 (this is specific to the restricted, closed-shell case). OpenMP parallelism can be exploited to speed up the process considerably.

Allocate space for a temporary $\mathbb{I}$ for each thread and set all entries to 0
Compute the upper triangle:
Loop over all auxiliary basis functions P
using OpenMP to divide the work among different threads
  Loop over all occupied orbitals i
    Loop over virtual orbitals a that share a fragment with i
      $\mathbb{I}_{ii}^{aa} \mathrel{+}= B_{ia}^P B_{ia}^P$
      Loop over virtual orbitals $b > a$ that share a fragment with i
        $\mathbb{I}_{ii}^{ba} \mathrel{+}= B_{ib}^P B_{ia}^P$
      Loop over occupied orbitals $j > i$
        Loop over virtual orbitals b that share a fragment with j
          $\mathbb{I}_{ji}^{ba} \mathrel{+}= 2B_{jb}^P B_{ia}^P - B_{ib}^P B_{ja}^P$
Combine the temporary copies of $\mathbb{I}$*:*
Loop over the temporary copies of $\mathbb{I}$ for each thread beyond the first
  Loop over all entries in $\mathbb{I}$
    Add the thread's temporary $\mathbb{I}$ to the main $\mathbb{I}$
    Free the temporary $\mathbb{I}$
Copy to lower triangle:
Loop over all occupied orbitals i
  Loop over virtual orbitals a that share a fragment with i
    Loop over virtual orbitals $b > a$ that share a fragment with i
      $\mathbb{I}_{ii}^{ab} = \mathbb{I}_{ii}^{ba}$
    Loop over occupied orbitals $j > i$
      Loop over virtual orbitals b that share a fragment with j
        $\mathbb{I}_{ij}^{ab} = \mathbb{I}_{ji}^{ba}$

Figure 1. Algorithm for $\mathbb{I}$.

The algorithm ensures that only one out of each pair of equivalent entries in $\mathbb{I}$ need be calculated. Each duplicate value can then be copied at the end. The looping order P, i, a is the same as the order of the indices in B, taking advantage of as much locality as possible. The choice of which of $\mathbb{I}_{ji}^{ba}$ or $\mathbb{I}_{ij}^{ab}$ to assign to in the main loop is also significant, as the right indices of $\mathbb{I}$ are the fast ones, so this helps increase locality of references to $\mathbb{I}$.

3.1.2. Solving the Linear Equation. With the right-hand side of eq 3, the task becomes solving the linear system. Because Δ is an fourth-rank tensor over occupied-virtual pairs with side length ovN, the system can be solved in time $O(o^4v^4N^4)$ without taking advantage of any of its internal structure. However, this has unacceptable eighth-order scaling with respect to the size of the fragments and undesirable fourth-order scaling with respect to the number of fragments. Taking advantage of some internal structure allows the application of an explicit inverse of Δ in time $O(o^3v^3N^3)$. This achieves the desired third-order scaling with respect to the number of fragments but still has less-than-ideal sixth-order scaling with respect to the size of the fragments, when MP2 scales as fifth order. Although Δ cannot easily be inverted, forward application is possible in time $O(o^2v^3N^3)$, matching MP2's scaling. Because this is much faster than inversion and Δ is a symmetric, positive definite operator, the iterative conjugate gradients method can be used to solve the system. The one requirement for conjugate gradients is a fast forward application of the operator. It can be significantly sped up with a good initial guess for the solution to the system and a simple approximate inverse of the operator as a preconditioner.

Although the zero vector may be used as the initial guess for iteration, convergence can be significantly faster with a better guess. A reasonable guess is available from the calculation of E(frag/ALMO). The t amplitudes can be combined into an initial guess for the CCC t where all double excitations involving two fragments are set to 0 (the CCC constraint excludes double excitations involving more than two fragments). As the correlation between fragments is generally small compared to the correlation on fragments (though more significant when it comes to binding), this will be a reasonably close first guess.

As the convergence of conjugate gradients depends on the condition number of the system, preconditioning can drastically improve convergence. This can be thought of as an approximate inverse, as it brings the operator closer to the identity. The preconditioner must be fast to apply, as it is used each iteration of the method. The fastest preconditioner is a diagonal one. We will use the inverse of the diagonal elements of Δ, as this can be applied with one division per double excitation, which is negligible time compared to the full forward application of Δ. This is the same as the inverse of Δ in the limit where overlap between virtual orbitals on different fragments is 0. In large or weakly interacting systems, most of these overlaps will be 0 and most of the remaining ones will be small, making this is a good preconditioner.

The only remaining piece is an efficient application of Δ. Remember that Δ is defined by eq 4

$$\Delta_{iji'j'}^{aba'b'} = -f_{ii'}g_{aa'}g_{jj'}g_{bb'} + g_{ii'}f_{aa'}g_{jj'}g_{bb'} - g_{ii'}g_{aa'}f_{jj'}g_{bb'} + g_{ii'}g_{aa'}g_{jj'}f_{bb'}$$

subject to the constraint that each pair i/a, j/b, i'/a', and j'/b' shares a fragment. Because this is a linear operator, the four pieces can be applied separately. Furthermore, the last two are the same as the first two, but with different indices. This leaves two pieces to consider. The second piece, $g_{ii'}f_{aa'}g_{jj'}g_{bb'}$ is the simpler one. Because of the global orthogonalization of the occupied orbitals, $g_{jj'} = \delta_{jj'}$. This along with the fragment constraint ensures that b and b' share a fragment. As each fragment's virtual space is orthonormal, $g_{bb'} = \delta_{bb'}$. A similar argument applies to $g_{ii'}$ and $f_{aa'}$, so taking advantage of the sparsity placed in f an g during the basis setup, the second piece of Δ becomes $f_a\delta_{ii'}\delta_{aa'}\delta_{jj'}\delta_{bb'}$. As this is diagonal, it is trivial to apply.

The first piece, $f_{ii'}g_{aa'}g_{jj'}g_{bb'}$, is more complicated. It is the same, except f and g switch places. The argument for $g_{jj'}g_{bb'} = \delta_{jj'}\delta_{bb'}$ still holds, but the argument for the first two terms does

DOI: 10.1021/acs.jpca.6b11516
J. Phys. Chem. A 2017, 121, 717−728

not. The virtual−virtual overlap matrix $g_{aa'}$ does not have any special properties, so it does not require that i and i' share a fragment. The between fragments part of $f_{ii'}$ also lacks any special properties. Therefore, the two matrices will actually have to be applied. Keeping the fragment constraint is tricky and requires two separate matrix multiplications per pair of fragments. Luckily, these are standard matrix multiplications and can be performed by BLAS. The first and third pieces can also be combined, cutting in half the number of matrix multiplications. The procedure for computing $r = \Delta t$ appears in Figure 2. It also parallelizable with OpenMP. The structure of f and g allow considerable time savings. The bulk of the time will be spent on the BLAS calls. The slower of the two requires time $O(ov^2)$ and is done N^2 times for each of the ovN j/b pairs, for a total time complexity of $O(o^2v^3N^3)$. This structure minimizes the number of writes to the r matrix, which cannot be local for both r_{ij}^{ab} and r_{ji}^{ba}. Within the final loop, adding to r_{ij}^{ab} takes care of the $-f_{ii'}g_{aa'}g_{jj'}g_{bb'}$ and $g_{ii'}f_{aa'}g_{jj'}g_{bb'}$ pieces of Δ, and adding to r_{ji}^{ba} takes care of $-g_{ii'}g_{aa'}f_{jj'}g_{bb'}$ and $g_{ii'}g_{aa'}g_{jj'}f_{bb'}$ without recomputing the identical blocks of the update matrix.

Loop over all fragments n
  Loop over all pairs j and b such that j and b share a fragment
  using OpenMP to divide the work among different threads
    Loop over all fragments n'
      Set the temporary matrix "block" to zero
      If $n = n'$
        On a fragment, f and g are diagonal
        Loop over pairs of i and a on fragment n
          $\text{block}_i^a \mathrel{-}= f_{ii}t_{ij}^{ab}$
          $\text{block}_i^a \mathrel{+}= f_{aa}t_{ij}^{ab}$
      Otherwise
        With BLAS, compute $\text{temp}_{i'}^a = g_{aa'}t_{i'j}^{a'b}$
        (with a constrained to fragment n and a' and i' constrained to fragment n')
        With BLAS, compute $\text{block}_i^a \mathrel{-}= f_{ii'}\text{temp}_{i'}^a$
        (with a and i constrained to fragment n and i' constrained to fragment n')
    Loop over pairs of i and a on fragment n
      Atomically, $r_{ij}^{ab} \mathrel{+}= \text{block}_i^a$
      Atomically, $r_{ji}^{ba} \mathrel{+}= \text{block}_i^a$

Figure 2. Algorithm for multiplication by Δ.

With this, all the ingredients for conjugate gradients are in place and the CCCMP2 equation can be set up and solved.

3.2. Isolated Terms. The remaining terms use the isolated fragment orbital basis. Because they do not involve any global calculations like E_{ccc}(sys/ALMO), they are simpler to compute. The total amount of work associated with them scales only linearly with the number of fragments. The first step is E(frag). This is the same as the canonical MP2 energy of each isolated fragment, but with the auxiliary functions from all fragments included. However, the **z** vector must be computed to be used for E(frag/frz). In addition, the t amplitudes are saved. For computing E(frag/frz), projected basis versions of the Fock matrix and the **B** matrix are required. The Hylleraas functional is used is for energy. This includes the normal MP2 energy evaluation and one using the relaxed response density matrix of the isolated fragment t amplitudes and z vector. These steps are not substantially different in implementation than ordinary MP2 and do not need to be described in further detail.

4. PERFORMANCE

The algorithms described above have been implemented in a development version of the Q-Chem quantum chemistry program package.[66] The only two subcalculations that scale as both fifth order in the size of the fragments and third order in the number of fragments are the calculation of the right-hand side of eq 3 and finding its solution. Therefore, these are the most important to benchmark

4.1. Benchmarking Solving the Linear Equation. Because of the iterative nature of the conjugate gradients algorithm, the two factors determining the speed of solving the system are the number of iterations and the time taken for each iteration. The number of iterations should grow very slowly or not at all with the size of the system. This was tested by performing test calculations with water clusters from two to ten molecules. The main and auxiliary basis sets used were aug-cc-pVTZ and rimp2-aug-cc-pVTZ, which have 92 main and 198 auxiliary basis functions per molecule. This system was also used as a test of the initial guess and the preconditioning for reducing the number of iterations. As can be seen in Table 1, with any of the algorithms, the number of iterations roughly stabilizes after about 6 water molecules. The preconditioner leads to a greater than factor of 10 reduction in the number of iterations, reducing conjugate gradients from taking the large majority of overall computational time to being merely one of several roughly equally sized steps. The initial guess is less important, but is worthwhile if it can save even one iteration as it is trivial to compute. Between the two of them, conjugate gradients takes at most 11 iterations for up to 10 water molecules. This is with fairly stringent convergence criteria ($\|r\|_\infty \leq 10^{-15}$, where r is the residual). A few more iterations could be saved with looser criteria, but this would come at the expense of precision.

Table 1. Number of Conjugate Gradient Iterations vs Number of Water Molecules, For Four Different Algorithms[a]

n	GP	P	G	⌀
2	10	10	117	130
3	10	10	117	130
4	11	11	117	128
5	11	11	119	129
6	11	12	125	133
7	11	12	122	135
8	11	12	123	131
9	11	12	123	133
10	11	12	122	132

[a]G denotes using the isolated fragment excitations as an initial guess, P denotes preconditioning with the diagonal elements of Δ, GP both, and ⌀ neither.

The other factor is the time taken per iteration. There should be roughly cubic scaling in the number of identical fragments. This behavior is seen is in Figure 3, as predicted.

The quality of the parallelization can also be benchmarked. Ideally, speed would grow in direct proportion to the number of threads, but this is not feasible for many problems. In any case, speed should still increase significantly with the number of threads. The behavior is graphed in Figure 4 for a varying number of threads. The graph shows an increase of speed with every additional thread, but with a slowing increase after 4 threads.

4.2. Benchmarking Integral Calculation. The same timing concerns apply to the calculation of 𝕀. It should also be cubic for an increasing number of fragments. This is verified by the plot in Figure 5. Behavior when parallelized is also important, as this calculation is primarily memory bound and cannot rely on BLAS calls, unlike the multiplication by Δ.

DOI: 10.1021/acs.jpca.6b11516
J. Phys. Chem. A 2017, 121, 717−728

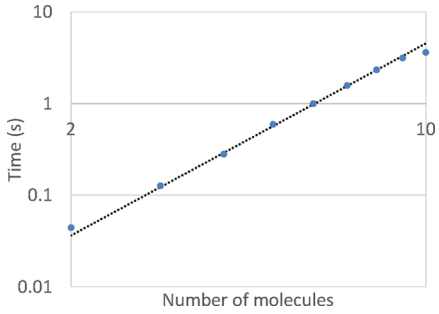


Figure 3. Log–log plot of time per conjugate gradient iteration vs number of water molecules, with a cubic fit line.

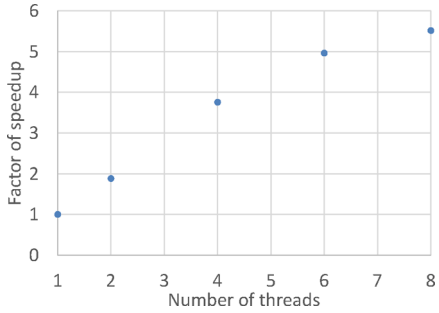


Figure 4. Plot of time per conjugate gradient iteration vs number of threads.

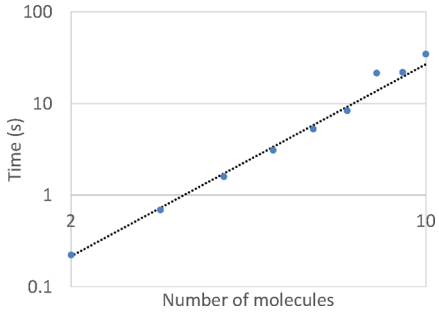


Figure 5. A log–log plot of time to compute 𝕀 vs number of water molecules, with a cubic fit line.

Because of this, calculation of 𝕀 shows a more modest improvement with additional threads when compared to the conjugate gradient iterations. This is shown in Figure 6.

4.3. Overall Benchmarks. Although those are the two most important steps for most systems, the transformation of the RI integrals into the molecular orbital basis can be significant when there is a large number of fragments. It is worth comparing the overall time for all of the EDA to the necessary standard calculations: HF plus RIMP2 energy. The times given for the EDA do not include calculating the three-center, atomic orbital basis RI integrals, as those are identical between the EDA and standard RIMP2. For the EDA, the two described steps are the most important. However, the transformation of the RI integrals to the molecular orbital basis B matrix is also significant. As expected, it grows in importance for greater numbers of fragments. For small numbers of molecules, the transformation (including the two fragment blocked B matrices needed for the frozen energy) takes about one-quarter of the overall running time. With 10 molecules, it is about one-third. Everything other than transformation of the B matrices and the CCC energy takes less than 2% of the overall time with 10 molecules. This includes all basis transformations and the single-fragment polarized, frozen, and isolated energies for all 10 molecules.

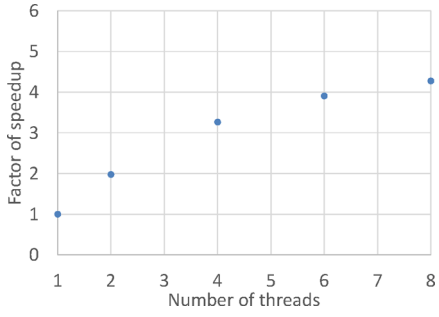


Figure 6. Plot of time to compute 𝕀 vs number of threads.

In comparing the EDA with standard RIMP2,[67] Q-Chem 4.3 was used[66] with eight threads on the same computer as the EDA benchmarks. The comparison is shown in Figure 7. The graph compares the running time of the EDA to the HF calculation as well as the additional time for the RIMP2 portion of the calculation. In all cases with three or more molecules, the SCF time was 40 or 50 times larger than the RIMP2 time. As can be seen, the standard RIMP2 calculation is slightly faster for four or fewer molecules. However, thanks to better scaling, the EDA takes less time for five or more molecules, despite also having to calculate the single fragment energies. With ten molecules, the EDA takes about 1.6 times less time. The CCC energy involves less interaction than standard MP2. However, it requires a less efficient formulation of the MP2 energy. This means that greater speed is not seen until a relatively large number of molecules. However, the benchmarks show that even for as few as two molecules, the EDA is not particularly expensive to do in addition to a standard MP2 binding energy calculation. When counterpoise correction is applied, this will be true to an even greater extent, as the EDA requires no change in algorithm to work with the counterpoise corrected binding energy.

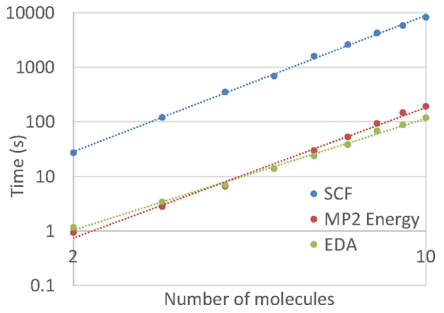


Figure 7. Q-Chem RIMP2 time (excluding SCF) vs EDA time.

This trend continues in systems with larger numbers of atoms. To demonstrate this, the same benchmark was performed on ammonia–water clusters with a single ammonia molecule surrounded by 5–30 water molecules, using the cc-pVDZ basis set. The result is shown in Figure 8. The efficiency gains of fragment blocking can be seen more strongly here, as there is a factor of 4 speedup in the largest system. As the system becomes larger, the additional computing time required for the EDA becomes negligible compared to the that for binding energy.

DOI: 10.1021/acs.jpca.6b11516
J. Phys. Chem. A 2017, 121, 717–728

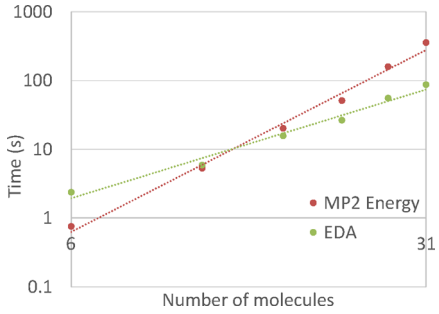


Figure 8. Q-Chem RIMP2 time (excluding SCF) vs EDA time.

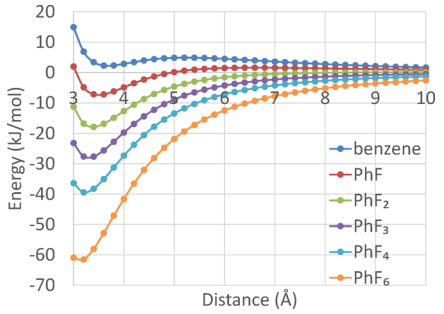


Figure 9. Binding energy of six different anion−π interactions.

5. APPLICATION: ANION−π INTERACTIONS

To demonstrate the use of the efficient implementation of the EDA, we will apply it to larger systems than we have previously done. The system is the anion−π interaction between the chloride anion and the π system of substituted benzene rings. Anion−π interactions have become the subject of much interest recently in part because of potential uses in biological systems or for molecule recognition.[54] They are similar to the better known cation−π interactions. However, the negative charge of the ion means that quadrupole moment of an unsubstituted benzene ring is repulsive rather than attractive. Therefore, attractive anion−π interactions are found only in systems where the aromatic ring has strongly electron withdrawing substituents. Therefore, there has been considerable interest in understanding the nature and origin of anion−π bonds. One such explanation comes from Wheeler and Houk, who argue[57] that the anion−π interaction can be understood as weak, repulsive anion−benzene interaction plus a purely electrostatic interaction between the ion and the dipole moments of the substituents. They say that higher order effects, such as substituents modifying the interaction between the ion and the π system are not important and that the interactions with each of the individual substituents can simply be summed together. Others, meanwhile, have claimed that charge transfer[68] or "multicenter covalency"[69] is the key force in the bonds. These claims are evaluated here through the EDA. Structures for the substituted benzene rings were obtained from Wheeler and Houk's Supporting Information.

5.1. Perpendicular Path. The test system is a chloride ion moved along a perpendicular path from the center of a benzene ring with a varying number of substituted fluorine atoms. PhF_2 is in the para isomer, PhF_3 is in the sym (1,3,5) isomer, and PhF_4 is in the 1,2,4,5 isomer. Therefore, other than PhF, no molecule studied here has a net dipole moment. The results support a picture where the influence of the additional fluorine atoms on the anion−π interaction is an electrostatic effect and the overall interaction only become attractive when there are enough fluorine atoms to reverse the overall quadrupole moment of a benzene molecule. As can be seen in Figure 9, the interaction between benzene and a chloride ion is not attractive, though there is a shallow local minimum. Every additional fluorine atom to the benzene ring makes the interaction more favorable. The long-range behavior of each curve is the R^{-3} decay characteristic of ion−quadrupole interactions (though even at 10 Å, not all curves are yet close to their asymptotic behavior). For two or fewer fluorine atoms, the long-range interaction is repulsive, and for three or more, it is attractive. Each additional fluorine atom adds nearly the same amount of favorable effect to the overall interaction.

Applying the EDA to these binding energy curves shows that the difference between the different π systems is an electrostatic one. Figure 10a shows the frozen interaction for the different systems, and Figure 10b shows the sum of the other components. As can be seen, in the frozen interaction, each fluorine atom makes the interaction more favorable. However, the other components show almost no change between the systems. The difference is barely visible on the graph. This is not a cancellation of larger differences—each individual component shows a similarly small change.

The frozen term can be understood in terms of Pauli repulsion and electrostatics. The close range part is dominated by Pauli repulsion, which should be similarly sized in each of these systems. Every additional fluorine adds a favorable interaction, which is similarly sized and shaped each time. This is the electrostatic interaction between the dipole of the carbon−fluorine bonds and the charge of the chloride ion. Other than PhF, the molecules have no net dipole, so the interaction can be understood in terms of the higher multipole moments. Even in the case of PhF, the path of the ion is perpendicular to the plane of the benzene ring and therefore also to the dipole. This means the leading term of the interaction decays as a quadrupole's characteristic R^{-3} in the long range rather than R^{-2}. Because the different in-plane components of the dipoles do not effect the energy, so there is no cancellation effect as the second fluorine atom added, even though it cancels the net dipole of the first. This makes it possible to closely approximate the effect of an additional fluorine atom by simply adding the scalar energy of the interaction.

The insignificant changes in the nonfrozen components of the EDA strong supports Wheeler and Houk's argument that the substituents primarily affect the interaction through electrostatics. We see that despite the electron withdrawing effect of the fluorine atoms, there is no significant increase in charge transfer from the chloride ion to the π system. Neither is there any noticeable change in polarization. Even dispersion changes little with the increasing number of electrons as more fluorine atoms are added.

Therefore, the nonelectrostatic parts of the interaction are not changed significantly with the addition of substituents. Accordingly, studying the interaction with benzene is sufficient to understand these effects for all of the systems considered here. The full EDA for that system is shown in Figure 11. At the distance of the interaction, charge transfer is not significant, as expected from the fact that both systems are electron rich. The local minimum that this interaction displays is primarily due to polarization and dispersion. The local minimum is a result of its faster R^{-4} decay compared to the dominant frozen interaction. In the case of PhF_3, where the quadrupole moment

DOI: 10.1021/acs.jpca.6b11516
J. Phys. Chem. A 2017, 121, 717−728

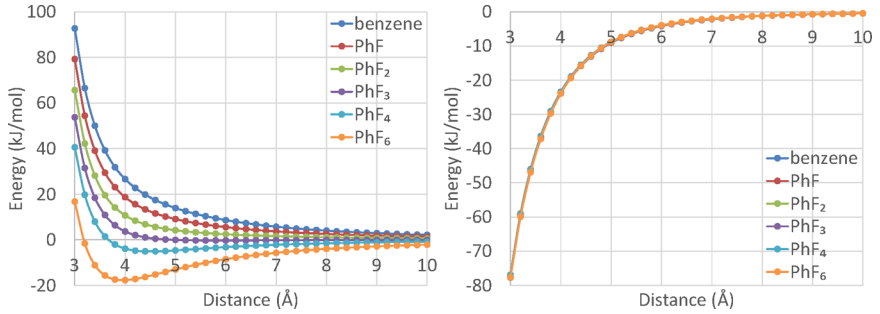


Figure 10. (a) Frozen interaction and (b) sum of polarization, charge transfer, and dispersion for six different anion−π interactions.

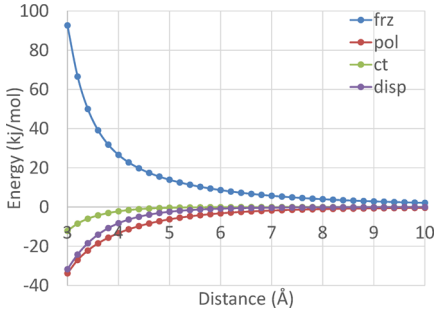


Figure 11. EDA for chloride−benzene interaction.

and thus the electrostatic interaction are nearly zero, polarization and dispersion are the two significant forces and do make a fairly strong favorable interaction. It is worth noting the importance of electron correlation for this interaction. Figure 12 shows the EDA for only Hartree−Fock. The major

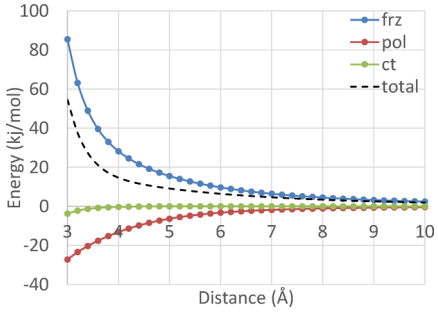


Figure 12. Hartree−Fock EDA for chloride−benzene interaction.

difference is the lack of dispersion, which is significant in the interacting region. Also, the HF graph shows the characteristic overestimation of multipole moments, and the repulsive frozen interaction is is significantly overestimated in the medium to long range. These two errors prevent HF from describing the local minimum in the binding energy. It also because of this that HF describes the PhF system as unbound and gets the long-range behavior of all systems wrong.

The frozen term can be analyzed further by looking at the multipole moments of the substituted benzene molecules. Beyond the range of Pauli repulsion, the frozen term is dominated by electrostatic interactions. The changes in multipole moments are due to the greater electronegativity of fluorine relative to hydrogen, so this could be viewed as interaction with the substituents even though the moments belong the overall molecule. Because we are only concerned with the z axis, the traceless Cartesian multipole moment Q_{z^l} and the real and pure spherical multipole moments Q_{l0} and Q_l^0 are all equal and we do not have to be concerned with a choice of convention. For every molecule except PhF and PhF_3, the odd moments are zero as they have even dihedral symmetry with a horizontal mirror plane about the chosen origin (which causes them to have even symmetry with respect to reflection about all three coordinate planes, whereas each odd multipole moment has terms that have odd symmetry with respect to at least one coordinate). For PhF and PhF_3, these moments are nonzero, but they do not interact with a charge moving along the z axis, as Y_l^0 is the only spherical harmonics that is nonzero there. As for Q_l^0, it must be zero for odd l in the presence of a horizontal plane of symmetry through the origin (as for all terms in Y_l^0, z and l have the same parity).

In many situations, the long-range electrostatic interaction can be closely approximated by its leading term—the lowest nonzero multipole moment. However, that is not sufficiently accurate in the case of the chloride−benzene interaction. As can be seen on the quadrupole line in Figure 13, the quadrupole

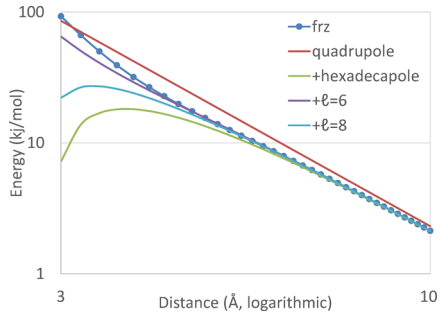


Figure 13. Frozen interaction of chloride and benzene vs truncated multipole expansions of the classical electrostatic interaction, log−log plot.

does not match the frozen term particularly well, even at 10 Å of separation. Adding in the hexadecapole moment makes the curves match much more closely but are still visibly different in the region closer than 6 Å. Evidently, even higher multipole moments are needed. Figure 13 also shows the addition of the sixth (hexacontatetrapole) and eighth (hecatonicosaoctopole) moments. With all of those multipole moments the expression for the approximation becomes $V(z) = -\frac{Q_2^0}{z^{-3}} - \frac{Q_4^0}{z^{-5}} - \frac{Q_6^0}{z^{-7}} - \frac{Q_8^0}{z^{-9}}$. The sixth-order approximation is closer to the frozen curve than the eighth-order one is. However, the frozen term also includes Pauli repulsion at close range, so the electrostatic curve should not necessarily match the frozen one at distances of a few angstroms. We can confirm the correct order of truncation by looking at the difference

DOI: 10.1021/acs.jpca.6b11516
J. Phys. Chem. A 2017, 121, 717−728

between the two curves. This is plotted in a lin−log plot in Figure 14. This curve shows Pauli repulsion's characteristic

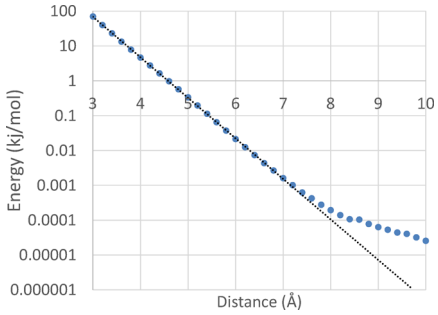


Figure 14. Lin−log plot of difference between frozen interaction and multipole approximation to classical electrostatic interaction.

exponential decay until about 7 Å and then it slows down. However, due to limitations with finite basis sets and subtractive calculation of Pauli repulsion, this is about as far as we should expect an accurate decomposition to go. With truncation at sixth order, deviation from exponential decay is clear, and the multipole approximation becomes less repulsive than the frozen energy beyond about 5 Å. Therefore, eighth order is the correct truncation for analyzing the electrostatics.

Therefore, the effect of the additional fluorine substituents on the frozen interaction can be analyzed by looking at changes in the first several multipoles of the substituted benzene molecule. They are shown in Table 2. To show that these accurately describe the interaction, the difference for all of them is plotted in Figure 15. The plot only goes to 7 Å for reasons described above. Even at this distance, it can be seen that the PhF_6 and PhF_3 curves decay too quickly. However, in the short range, all curves show exponential decay. Differences are small, but not negligible. The remainder at 3 Å ranges from 70 kJ/mol for benzene to 95 kJ/mol for PhF_6. The size of repulsion is the order of the increasing amount of fluorine atoms, but rather than being evenly spaced, differences get larger as more atoms are added. Regardless, this shows that the electrostatic portion of the frozen interaction is well described by the truncated multipole expansion.

The table of multipole moments reveals that changes are not quite the same for each additional fluorine atom. The quadrupole moment increases by 2.93 D·Å for the first atom, but only an average of 2.47 D·Å for the fifth and sixth. The hexadecapole moment changes are similarly inconsistent. However, the sixth-order moment changes vary by a factor of 1.5 between the smallest and the largest. The eighth-order moment is much more erratic. Adding the first fluorine atom decreases it by 320 D·$Å^7$, and the next few also lead to decreases. However, the fifth and the sixth actually increase it by an average of 84 D·$Å^7$. These differences are reflected in the frozen energy and therefore also the overall interaction energy. Accordingly, the interactions between the chlorine atom and

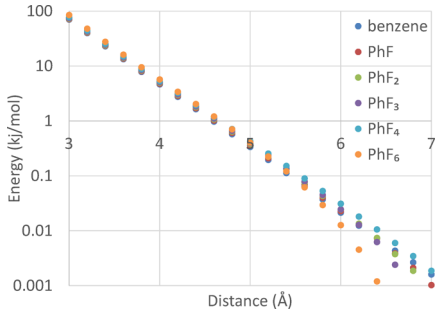


Figure 15. Lin−log plot of differences between frozen interactions and multipoles approximation to classical electrostatic interaction.

each of the fluorine substituents is not quite the same. Adding the same interaction potential is a good first-order approximation but can get the successive differences wrong by as much as 15%.

The EDA has shown that the favorability of the anion−π interaction when the benzene ring is highly substituted is a result of electrostatic interactions with the anion, agreeing with the argument of Wheeler and Houk. However, the effects are not quite additive, as different fluorine atoms cause somewhat different changes to the multipole moments of the system. Furthermore, as was described earlier, even this level of additivity is a result of perpendicular path over which the ion was translated.

5.2. Angled Path. To demonstrate this, the same calculations can be repeated where the anion is instead moved at a 30° angle. Only benzene, PhF, and PhF_2 are included here, so as to show the difference between a polar and nonpolar system when the ion's path is not perpendicular to the dipole. The binding energy is shown in Figure 16. As can be

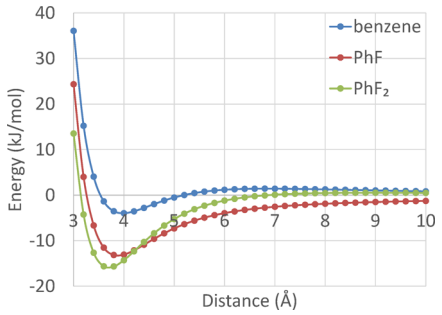


Figure 16. Binding energy of three different anion−π interactions, with an angled path.

seen, the first fluorine atom makes the interaction more favorable, but the second one cancels the dipole effect and makes the overall interaction less favorable than the unsubstituted benzene. This shows that the simple additive nature of the interaction with the substituents depends on that path of the ion being perpendicular to the dipole. Otherwise,

Table 2. Significant Multipole Moments for Substituted Benzenes (Units D·$Å^{l-1}$)

	benzene	PhF	PhF_2	PhF_3	PhF_4	PhF_6
Q_2^0	−7.9807	−5.0545	−2.1191	0.4191	3.2045	8.1383
Q_4^0	65.7708	40.0664	14.5271	−6.7322	−31.0507	−72.0979
Q_6^0	−435.886	−295.645	−159.343	−49.5916	68.7468	253.124
Q_8^0	2907.83	2587.72	2312.83	2118.02	2101.59	2269.89

DOI: 10.1021/acs.jpca.6b11516
J. Phys. Chem. A 2017, 121, 717−728

the in-plane components affect the energy and the interactions must be considered as vectors. Once again, the main change in the EDA components between systems is found in the frozen term, though PhF_2 does have a somewhat larger dispersion interaction in the short range due to the proximity of the ion and the second fluorine atom.

To make this clearer, the differences in the frozen energy between benzene and PhF as well as between PhF and PhF_2 are shown in Figure 17. As can be seen, the two curves do not resemble each other. The benzene to PhF difference is dominated by the interaction with the molecular dipole. The long range of the PhF to PhF_2 is opposite and roughly equal as the dipole is canceled by the second fluorine atom. However, in the close range, the added interaction is attractive. This happens because the second fluorine atom (like the first) increases the favorability of the quadrupole and of some higher moments. Therefore, simply adding the interaction energy with each atom is not sufficient, as the quadrupole is reinforced and the dipole is canceled. However, the situation is somewhat more complicated than this. The dipole moment is not the only one that is canceled with the addition of the second fluorine atom. At the quadrupole level, in the $x-z$ plane, Q_{z^2} is not the only important moment. $Q_{x^2-y^2}$ also contributes. As shown earlier, Q_{z^2} is approximately doubled from PhF to PhF_2 because the two C−F bonds reinforce each other. However, $Q_{x^2-y^2}$ is 0 for benzene and PhF_2 by symmetry, but −6.6 D·Å for PhF. It is another example of a mulitpole moment that is canceled rather than reinforced, so even at the quadrupole level, the interaction energies cannot simply be added.

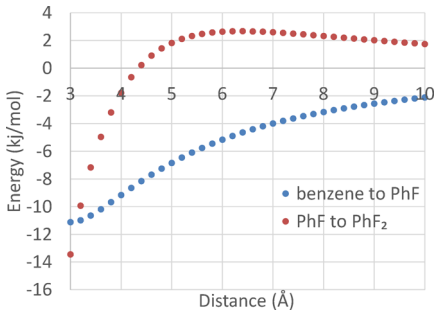


Figure 17. Frozen energy differences between three different anion−π interactions, with an angled path.

This continues for higher moments. Symmetry also requires benzene and PhF_2 to have no octopole moment, but PhF has nonzero Q_{xz^2} and $Q_{x(x^2-3y^2)}$ moments. At the hexadecapole level, all moments reinforce. Beyond that, the pattern continues. This could be ignored for the perpendicular path, because none of the canceling moments contribute to the energy along the z axis. This left only reinforcing moments, allowing the total interaction energy with several fluorine atoms to be approximated as the sum of several identical interactions. However, along any other path, the interaction must be considered as described here. If the second fluorine atom were not in the para position, then the interactions would not exactly cancel or double but would be somewhere in between. Therefore, the simple, linear changes in interaction energy with additional substituents is a result of the path examined and not a property of anion−π interactions in general.

6. CONCLUSIONS

The MP2 ALMO EDA has been implemented at a production quality level. Performance is comparable to full RIMP2 for a small number of molecules and scales better with respect to an increasing number of molecules. The implementation supports BLAS for matrix operations and OpenMP for parallelism. With this level of performance, it is practical to apply to the EDA to systems where RIMP2 is useful. Performance is aided by the problem being naturally amenable to preconditioned conjugate gradients. This implementation can also be readily applied to modifications of MP2 theory such as spin-component scaling[16,17] and attenuated MP2[20,21] that improve over standard MP2 theory. We are also very interested in extending this MP2 EDA to double hybrid density functionals that offer even more significant improvements in accuracy.[70−74]

The MP2 ALMO-EDA was then applied to an anion−π system. Results show that polarization, charge transfer, and dispersion were nearly unaffected by the presence of substituents, and that all of change was in the frozen interaction. This supports the idea that favorable anion−π interactions are favorable because of electrostatic interaction with the substituents, rather than because the substituents cause the π system to become electron deficient, or any other more complicated explanation.

AUTHOR INFORMATION

Corresponding Authors

*J. Thirman. E-mail: thirman@berkeley.edu.
*M. Head-Gordon. E-mail: mhg@cchem.berkeley.edu.

ORCID

Martin Head-Gordon: 0000-0002-4309-6669

Notes

The authors declare the following competing financial interest(s): M.H-G. is a part-owner of Q-Chem Inc.

ACKNOWLEDGMENTS

This work was supported by a grant (CHE-1363342) from the U.S. National Science Foundation, with additional support from Q-Chem Inc. through an NIH Subcontract from Grant 2 R44 GM096678-02. M.H-G. is a part-owner of Q-Chem Inc.

REFERENCES

(1) Binkley, J. S.; Pople, J. A. Møller−Plesset Theory for Atomic Ground State Energies. *Int. J. Quantum Chem.* **1975**, *9*, 229−236.
(2) Helgaker, T.; Klopper, W.; Tew, D. Quantitative Quantum Chemistry. *Mol. Phys.* **2008**, *106*, 2107−2143.
(3) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986.
(4) Thomas, J. R.; DeLeeuw, B. J.; Vacek, G.; Crawford, T. D.; Yamaguchi, Y.; Schaefer, H. F. The Balance Between Theoretical Method and Basis Set Quality: A Systematic Study of Equilibrium Geometries, Dipole Moments, Harmonic Vibrational Frequencies, and Infrared Intensities. *J. Chem. Phys.* **1993**, *99*, 403.
(5) Cohen, A. J.; Tantirungrotechai, Y. Molecular Electric Properties: An Assessment of Recently Developed Functionals. *Chem. Phys. Lett.* **1999**, *299*, 465−472.
(6) Hehre, W. *A Guide to Molecular Mechanics and Quantum Chemical Calculations*; Wavefunction: Irvine, CA, 2003.
(7) Stone, A. *The Theory of Intermolecular Forces*; Oxford University Press: Oxford, U.K., 1997.
(8) 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.

DOI: 10.1021/acs.jpca.6b11516
J. Phys. Chem. A 2017, 121, 717−728

(9) Mardirossian, N.; Lambrecht, D. S.; McCaslin, L.; Xantheas, S. S.; Head-Gordon, M. The Performance of Density Functionals for Sulfate-Water Clusters. *J. Chem. Theory Comput.* **2013**, *9*, 1368−1380.
(10) Yoo, S.; Aprà, E.; Zeng, X. C.; Xantheas, S. S. High-Level Ab Initio Electronic Structure Calculations of Water Clusters $(H_2O)_{16}$ and $(H_2O)_{17}$: A New Global Minimum for $(H_2O)_{16}$. *J. Phys. Chem. Lett.* **2010**, *1*, 3122−3127.
(11) Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. Benchmark Database of Accurate (MP2 and CCSD(T) Complete Basis Set Limit) Interaction Energies of Small Model Complexes, DNA Base Pairs, and Amino Acid Pairs. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985.
(12) Thirman, J.; Head-Gordon, M. Electrostatic Domination of the Effect of Electron Correlation in Intermolecular Interactions. *J. Phys. Chem. Lett.* **2014**, *5*, 1380−1385.
(13) Sherrill, C. D.; Takatani, T.; Hohenstein, E. G. An Assessment of Theoretical Methods for Nonbonded Interactions: Comparison to Complete Basis Set Limit Coupled-Cluster Potential Energy Curves for the Benzene Dimer, the Methane Dimer, Benzene-Methane, and Benzene-H_2S. *J. Phys. Chem. A* **2009**, *113*, 10146−10159.
(14) Janowski, T.; Ford, A. R.; Pulay, P. Accurate Correlated Calculation of the Intermolecular Potential Surface in the Coronene Dimer. *Mol. Phys.* **2010**, *108*, 249−257.
(15) Rezac, J.; Riley, K. E.; Hobza, P. Evaluation of the Performance of Post-Hartree-Fock Methods in Terms of Intermolecular Distance in Noncovalent Complexes. *J. Comput. Chem.* **2012**, *33*, 691−694.
(16) Distasio, R.; Head-Gordon, M. Optimized Spin-Component Scaled Second-Order Møller-Plesset Perturbation Theory for Intermolecular Interaction Energies. *Mol. Phys.* **2007**, *105*, 1073−1083.
(17) Hill, J. G.; Platts, J. A. Spin-Component Scaling Methods for Weak and Stacking Interactions. *J. Chem. Theory Comput.* **2007**, *3*, 80−85.
(18) Takatani, T.; Hohenstein, E. G.; Sherrill, C. D. Improvement of the Coupled-Cluster Singles and Doubles Method via Scaling Same- and Opposite-Spin Components of the Double Excitation Correlation Energy. *J. Chem. Phys.* **2008**, *128*, 124111.
(19) Grimme, S.; Goerigk, L.; Fink, R. F. Spin-Component-Scaled Electron Correlation Methods. *WIREs Comput. Mol. Sci.* **2012**, *2*, 886−906.
(20) Goldey, M.; Head-Gordon, M. Attenuating Away the Errors in Inter- and Intramolecular Interactions from Second-Order Møller-Plesset Calculations in the Small Aug-cc-pVDZ Basis Set. *J. Phys. Chem. Lett.* **2012**, *3*, 3592−3598.
(21) Goldey, M.; Dutoi, A.; Head-Gordon, M. Attenuated Second-Order Møller-Plesset Perturbation Theory: Performance Within the aug-cc-pVTZ Basis. *Phys. Chem. Chem. Phys.* **2013**, *15*, 15869−15875.
(22) Lao, K.; Schaffer, R.; Jansen, G.; Herbert, J. Accurate Description of Intermolecular Interactions Involving Ions Using Symmetry-Adapted Perturbation Theory. *J. Chem. Theory Comput.* **2015**, *11*, 2473−2486.
(23) Reed, A. E.; Curtiss, L. A.; Weinhold, F. Intermolecular Interactions from a Natural Bond Orbital, Donor-Acceptor Viewpoint. *Chem. Rev.* **1988**, *88*, 899−926.
(24) von Hopffgarten, M.; Frenking, G. Energy Decomposition Analysis. *WIRES Comput. Mol. Sci.* **2012**, *2*, 43−62.
(25) Mo, Y. R.; Bao, P.; Gao, J. L. Energy Decomposition Analysis Based on a Block-Localized Wavefunction and Multistate Density Functional Theory. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6760−6775.
(26) Phipps, M. J. S.; Fox, T.; Tautermann, C. S.; Skylaris, C.-K. Energy Decomposition Analysis Approaches and Their Evaluation on Prototypical Protein−Drug Interaction Patterns. *Chem. Soc. Rev.* **2015**, *44*, 3177−3211.
(27) Bickelhaupt, F.; Baerends, E. Kohn-Sham Density Functional Theory: Predicting and Understanding Chemistry. *Rev. Comput. Chem.* **2000**, *15*, 1−86.
(28) Kitaura, K.; Morokuma, K. A New Energy Decomposition Scheme for Molecular Interactions Within the Hartree-Fock Approximation. *Int. J. Quantum Chem.* **1976**, *10*, 325−340.
(29) Ziegler, T.; Rauk, A. On the Calculation of Bonding Energies by the Hartree Fock Slater Method − I. The Transition State Method. *Theor. Chim. Acta* **1977**, *46*, 1−10.
(30) Mo, Y. R.; Gao, J. L.; Peyerimhoff, S. D. Energy Decomposition Analysis of Intermolecular Interactions Using a Block-Localized Wave Function Approach. *J. Chem. Phys.* **2000**, *112*, 5530−5538.
(31) Khaliullin, R. Z.; Cobar, E. A.; Lochan, R. C.; Bell, A. T.; Head-Gordon, M. Unravelling the Origin of Intermolecular Interactions Using Absolutely Localized Molecular Orbitals. *J. Phys. Chem. A* **2007**, *111*, 8753−8765.
(32) Khaliullin, R.; Bell, A.; Head-Gordon, M. Analysis of Charge Transfer Effects in Molecular Complexes Based on Absolutely Localized Molecular Orbitals. *J. Chem. Phys.* **2008**, *128*, 184112.
(33) Horn, P. R.; Sundstrom, E. J.; Baker, T. A.; Head-Gordon, M. Unrestricted Absolutely Localized Molecular Orbitals for Energy Decomposition Analysis: Theory and Applications to Intermolecular Interactions Involving Radicals. *J. Chem. Phys.* **2013**, *138*, 134119.
(34) Wu, Q.; Ayers, P. W.; Zhang, Y. K. Density-Based Energy Decomposition Analysis for Intermolecular Interactions with Variationally Determined Intermediate State Energies. *J. Chem. Phys.* **2009**, *131*, 164112.
(35) Lu, Z.; Zhou, N.; Wu, Q.; Zhang, Y. Directional Dependence of Hydrogen Bonds: a Density-based Energy Decomposition Analysis and Its Implications on Force Field Development. *J. Chem. Theory Comput.* **2011**, *7*, 4038−4049.
(36) Wu, Q. Variational Nature of the Frozen Density Energy in Density-Based Energy Decomposition Analysis and its Application to Torsional Potentials. *J. Chem. Phys.* **2014**, *140*, 244109.
(37) Horn, P. R.; Head-Gordon, M. Alternative Definitions of the Frozen Energy in Energy Decomposition Analysis of Density Functional Theory Calculations. *J. Chem. Phys.* **2016**, *144*, 084118.
(38) 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.
(39) Horn, P. R.; Head-Gordon, M. Polarization Contributions to Intermolecular Interactions Revisited with Fragment Electric-Field Response Functions. *J. Chem. Phys.* **2015**, *143*, 114111.
(40) Horn, P. R.; Mao, Y.; Head-Gordon, M. Probing Non-Covalent Interactions With a Second Generation Energy Decomposition Analysis Using Absolutely Localized Molecular Orbitals. *Phys. Chem. Chem. Phys.* **2016**, *18*, 23067−23079.
(41) Chen, W.; Gordon, M. S. Energy Decomposition Analyses for Many-Body Interaction and Applications to Water Complexes. *J. Phys. Chem.* **1996**, *100*, 14316−14328.
(42) Merrill, G. N.; Gordon, M. S. Study of Small Water Clusters Using the Effective Fragment Potential Model. *J. Phys. Chem. A* **1998**, *102*, 2650−2657.
(43) Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. The Effective Fragment Potential Method: A QM-Based MM Approach to Modeling Environmental Effects in Chemistry. *J. Phys. Chem. A* **2001**, *105*, 293−307.
(44) Gordon, M. S.; Mullin, J. M.; Pruitt, S. R.; Roskop, L. B.; Slipchenko, L. V.; Boatz, J. A. Accurate Methods for Large Molecular Systems. *J. Phys. Chem. B* **2009**, *113*, 9646−9663.
(45) Su, P.; Li, H. Energy Decomposition Analysis of Covalent Bonds and Intermolecular Interactions. *J. Chem. Phys.* **2009**, *131*, 014102.
(46) Saebo, S.; Pulay, P. Local Treatment of Electron Correlation. *Annu. Rev. Phys. Chem.* **1993**, *44*, 213−236.
(47) Lee, M. S.; Maslen, P. E.; Head-Gordon, M. Closely Approximating Second-Order Møller-Plesset Perturbation Theory with a Local Triatomics in Molecules Model. *J. Chem. Phys.* **2000**, *112*, 3592−3601.
(48) Schütz, M.; Rauhut, G.; Werner, H.-J. Local Treatment of Electron Correlation in Molecular Clusters: Structures and Stabilities of $(H_2O)_n$, n = 2−4. *J. Phys. Chem. A* **1998**, *102*, 5997−6003.

DOI: 10.1021/acs.jpca.6b11516
J. Phys. Chem. A 2017, 121, 717−728

(49) Grimme, S.; Mück-Lichtenfeld, C.; Antony, J. Analysis of Non-Covalent Interactions in (Bio)organic Molecules Using Orbital-Partitioned Localized MP2. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3327.
(50) Schneider, W. B.; Bistoni, G.; Sparta, M.; Saitow, M.; Riplinger, C.; Auer, A. A.; Neese, F. Decomposition of Intermolecular Interaction Energies within the Local Pair Natural Orbital Coupled Cluster Framework. *J. Chem. Theory Comput.* **2016**, *12*, 4778−4792.
(51) Thirman, J.; Head-Gordon, M. An Energy Decomposition Analysis for Second-Order Møller−Plesset Perturbation Theory Based on Absolutely Localized Molecular Orbitals. *J. Chem. Phys.* **2015**, *143*, 084124.
(52) Quiñonero, D.; Garau, C.; Rotger, C.; Frontera, A.; Ballester, P.; Costa, A.; Deyà, P. M. Anion−π Interactions: Do They Exist? *Angew. Chem., Int. Ed.* **2002**, *41*, 3389−3392.
(53) Mooibroek, T. J.; Black, C. A.; Gamez, P.; Reedijk, J. What's New in the Realm of Anion−π Binding Interactions? Putting the Anion−π Interaction in Perspective. *Cryst. Growth Des.* **2008**, *8*, 1082−1093.
(54) Salonen, L. M.; Ellermann, M.; Diederich, F. Aromatic Rings in Chemical and Biological Recognition: Energetics and Structures. *Angew. Chem., Int. Ed.* **2011**, *50*, 4808−4842.
(55) Wang, D.-X.; Wang, M.-X. Anion−π Interactions: Generality, Binding Strength, and Structure. *J. Am. Chem. Soc.* **2013**, *135*, 892−897.
(56) Chifotides, H. T.; Dunbar, K. R. Anion−π Interactions in Supramolecular Architectures. *Acc. Chem. Res.* **2013**, *46*, 894−906.
(57) Wheeler, S. E.; Houk, K. N. Are Anion/π Interactions Actually a Case of Simple Charge-Dipole Interactions? *J. Phys. Chem. A* **2010**, *114*, 8658−8664.
(58) Stoll, H.; Wagenblast, G.; Preuß, H. On The Use of Local Basis Sets for Localized Molecular Orbitals. *Theoret. Chim. Acta* **1980**, *57*, 169−178.
(59) Gianinetti, E.; Raimondi, M.; Tornaghi, E. Modification of the Roothaan Equations to Exclude BSSE From Molecular Interaction Calculations. *Int. J. Quantum Chem.* **1996**, *60*, 157−166.
(60) Nagata, T.; Takahashi, O.; Saito, K.; Iwata, S. Basis Set Superposition Error Free Self-Consistent Field Method for Molecular Interaction in Multi-Component Systems: Projection Operator Formalism. *J. Chem. Phys.* **2001**, *115*, 3553−3560.
(61) Khaliullin, R. Z.; Head-Gordon, M.; Bell, A. T. An Efficient Self-Consistent Field Method for Large Systems of Weakly Interacting Components. *J. Chem. Phys.* **2006**, *124*, 204105.
(62) Hylleraas, E. A. Über den Grundterm der Zweielektronenprobleme von H^-, He, Li^+, Be^{+2} usw. *Eur. Phys. J. A* **1930**, *65*, 209−225.
(63) Feynman, R. Forces in Molecules. *Phys. Rev.* **1939**, *56*, 340−343.
(64) Wiberg, K. B.; Hadad, C. M.; LePage, T. J.; Breneman, C. M.; Frisch, M. J. Analysis of the Effect of Electron Correlation on Charge Density Distributions. *J. Phys. Chem.* **1992**, *96*, 671−679.
(65) DiStasio, R. A.; Jung, Y. S.; Head-Gordon, M. A resolution-of-the-identity implementation of the local triatomics-in-molecules model for second-order Møller-Plesset perturbation theory with application to alanine tetrapeptide conformational energies. *J. Chem. Theory Comput.* **2005**, *1*, 862−876.
(66) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; et al. Advances in Molecular Quantum Chemistry Contained in the Q-Chem 4 Program Package. *Mol. Phys.* **2015**, *113*, 184−215.
(67) Goldey, M.; DiStasio, R. A.; Shao, Y. H.; Head-Gordon, M. Shared Memory Multiprocessing Implementation of Resolution-of-the-Identity Second-Order Møller−Plesset Perturbation Theory with Attenuated and Unattenuated Results for Intermolecular Interactions Between Large Molecules. *Mol. Phys.* **2014**, *112*, 836−843.
(68) Kim, D.; Tarakeshwar, P.; Kim, K. S. Theoretical Investigations of Anion−π Interactions: The Role of Anions and the Nature of π Systems. *J. Phys. Chem. A* **2004**, *108*, 1250−1258.
(69) Foroutan-Nejad, C.; Badri, Z.; Marek, R. Multi-Center Covalency: Revisiting the Nature of Anion−π Interactions. *Phys. Chem. Chem. Phys.* **2015**, *17*, 30670−30679.
(70) Schwabe, T.; Grimme, S. Double-Hybrid Density Functionals with Long-Range Dispersion Corrections: Higher Accuracy and Extended Applicability. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3397−3406.
(71) Zhang, Y.; Xu, X.; Goddard, W. A. Doubly Hybrid Density Functional for Accurate Descriptions of Nonbond Interactions, Thermochemistry, and Thermochemical Kinetics. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 4963−4968.
(72) Chai, J. D.; Head-Gordon, M. Long-Range Corrected Double-Hybrid Density Functionals. *J. Chem. Phys.* **2009**, *131*, 174105.
(73) Zhang, I. Y.; Xu, X. Doubly Hybrid Density Functional for Accurate Description of Thermochemistry, Thermochemical Kinetics and Nonbonded Interactions. *Int. Rev. Phys. Chem.* **2011**, *30*, 115−160.
(74) Kozuch, S.; Martin, J. Spin-Component-Scaled Double Hybrids: An Extensive Search for the Best Fifth-Rung Functionals Blending DFT and Perturbation Theory. *J. Comput. Chem.* **2013**, *34*, 2327−2344.

DOI: 10.1021/acs.jpca.6b11516
J. Phys. Chem. A 2017, 121, 717−728