

Communication: Practical intramolecular symmetry adapted perturbation theory via Hartree-Fock embedding

Robert M. Parrish, Jérôme F. Gonthier, Clémence Corminbœuf, and C. David Sherrill¹

Citation: *The Journal of Chemical Physics* **143**, 051103 (2015); doi: 10.1063/1.4927575

View online: <http://dx.doi.org/10.1063/1.4927575>

View Table of Contents: <http://aip.scitation.org/toc/jcp/143/5>

Published by the [American Institute of Physics](#)

Articles you may be interested in

[Spatial assignment of symmetry adapted perturbation theory interaction energy components: The atomic SAPT partition](#)

The Journal of Chemical Physics **141**, 044115044115 (2014); 10.1063/1.4889855

[Intramolecular symmetry-adapted perturbation theory with a single-determinant wavefunction](#)

The Journal of Chemical Physics **143**, 224107224107 (2015); 10.1063/1.4936830

[Density-fitted open-shell symmetry-adapted perturbation theory and application to \$\pi\$ -stacking in benzene dimer cation and ionized DNA base pair steps](#)

The Journal of Chemical Physics **145**, 134106134106 (2016); 10.1063/1.4963385

[Note: The performance of new density functionals for a recent blind test of non-covalent interactions](#)

The Journal of Chemical Physics **145**, 186101186101 (2016); 10.1063/1.4967424

[Long-range interactions from the many-pair expansion: A different avenue to dispersion in DFT](#)

The Journal of Chemical Physics **146**, 024111024111 (2017); 10.1063/1.4973728

**PHYSICS
TODAY**

**COMPLETELY
REDESIGNED!**

Physics Today Buyer's Guide
Search with a purpose.

Communication: Practical intramolecular symmetry adapted perturbation theory via Hartree-Fock embedding

Robert M. Parrish,¹ Jérôme F. Gonthier,¹ Clémence Corminbœuf,² and C. David Sherrill^{1,a)}

¹Center for Computational Molecular Science and Technology, School of Chemistry and Biochemistry, School of Computational Science and Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, USA

²Laboratory for Computational Molecular Design, Institut des Sciences et Ingénierie Chimiques, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

(Received 15 June 2015; accepted 17 July 2015; published online 4 August 2015)

We develop a simple methodology for the computation of symmetry-adapted perturbation theory (SAPT) interaction energy contributions for intramolecular noncovalent interactions. In this approach, the local occupied orbitals of the total Hartree-Fock (HF) wavefunction are used to partition the fully interacting system into three chemically identifiable units: **the noncovalent fragments A and B and a covalent linker C**. Once these units are identified, the noninteracting HF wavefunctions of the fragments A and B are separately optimized while embedded in the HF wavefunction of C, providing the dressed zeroth order wavefunctions for A and B in the presence of C. Standard two-body SAPT (particularly SAPT0) is then applied between the relaxed wavefunctions for A and B. This intramolecular SAPT procedure is found to be remarkably straightforward and efficient, as evidenced by example applications ranging from diols to hexaphenyl-ethane derivatives. © 2015 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4927575>]

The quantification and characterization of noncovalent interactions is an important task for theoretical chemistry. Such interactions may occur in an intermolecular sense, e.g., in the hydrogen bonding of two water molecules, wherein no covalent links exist between the two molecules of interest, or in an intramolecular sense, e.g., in the formation of hydrogen bonds between hydroxyl groups attached to the same molecule. In the former case, the intermolecular interaction energy and noninteracting zeroth-order wavefunctions of the monomers are easily defined quantities, leading to successful methodologies for intermolecular interactions. Example techniques include supermolecular interaction energies, energy decomposition analysis (EDA),¹ **block-localized wavefunction** (BLW)² or absolutely localized molecular orbital (ALMO)³ techniques, and symmetry-adapted perturbation theory (SAPT).^{4,5} The latter is particularly well developed and has the laudable property that it is formally completable to the full *ab initio* limit. We will therefore focus on SAPT in this work.

By contrast, methodology for the characterization of intramolecular interactions is virtually nonexistent. It has proven quite difficult to isolate sensible zeroth-order wavefunctions for the intramolecular fragments before the mutual interaction is activated. For a canonical example, when one looks at 2,4-pentandiol (Figure 1), it is fairly obvious that there should be an intramolecular hydrogen bond between the two hydroxyl groups (fragments A and B), and that this should be largely but not exactly like the intermolecular hydrogen bond between two water molecules. However, the presence of the carbohydrate backbone (the linker C) markedly complicates the task of choosing a noninteracting wavefunction for A and B. On the

one hand, it is difficult to determine which particles (electrons and protons) should be assigned to A, B, and C; an improper choice can yield zeroth-order wavefunctions for A and B with intrinsic charges that may overwhelm the subsequent $A \cdots B$ analysis. On the other hand, even if a chemically reasonable set of particles is used in the partition, the robust production of zeroth-order wavefunctions for A and B has been elusive. While the $A \cdots B$ interaction is weak and thus treatable by perturbation theory (e.g., SAPT), the $A \cdots C$ and $B \cdots C$ interactions are extremely strong, i.e., covalent bonds! Thus, to elucidate the noncovalent part of the interaction, the $A \cdots C$ and $B \cdots C$ interactions and orthogonality constraints must be maintained throughout the production of the zeroth-order wavefunctions. One approach to address these considerations has been recently reported,⁶ using the **chemical Hamiltonian approach** (CHA).⁷ While initial tests were promising, the CHA approach is based on assignment of the basis functions to atoms and thus is not well-defined at the complete basis set limit. Moreover, the current approach is mathematically simpler.

In this work, we demonstrate that a simple, chemically reasonable, and computationally robust procedure exists for the partitioning and determination of the zeroth-order wavefunctions of A and B, and that standard SAPT can be immediately adapted to this choice of reference. We then explore the strengths and limitations of this intramolecular SAPT (ISAPT) scheme in sample applications ranging from 2,4-pentandiol to hexaphenyl-ethane.

Note that the *intramolecular* developments considered in this manuscript stem from two previous installments of our work on *intermolecular* interactions: atomic SAPT (A-SAPT)⁸ and **functional group SAPT** (F-SAPT).⁹ The former developed an effective two-body partition of SAPT, allowing one to parti-

^{a)}Electronic mail: sherrill@gatech.edu

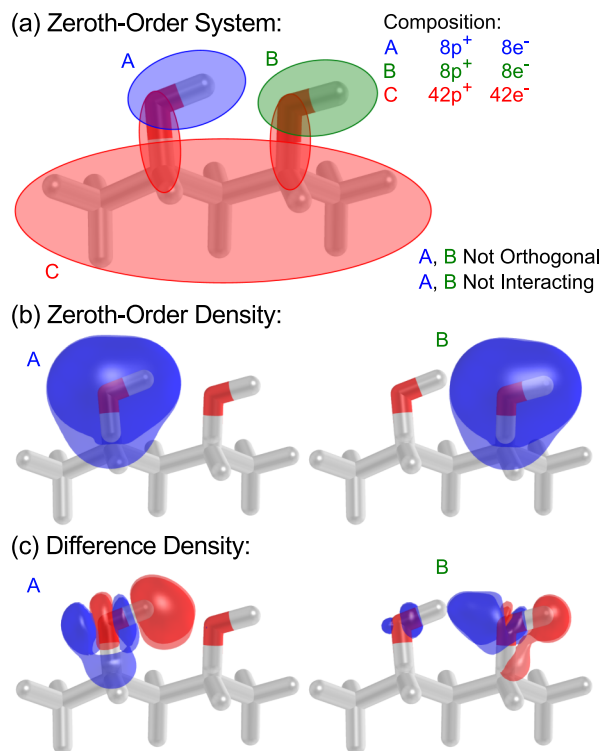


FIG. 1. (a) Schematic of the zeroth-order system for 2,4-pentandiol for the choice of linking σ bonds assigned to fragment C. (b) Zeroth-order density field for 2,4-pentandiol (links assigned to C) superimposed with the zeroth-order density field for the corresponding water dimer. Isosurfaces at 0.004 electron bohr⁻³. (c) Difference density fields superimposed as in (b) showing where density depletes (red) or accumulates (blue) when the HF interaction between A and B is activated. Isosurfaces at ± 0.0008 electron bohr⁻³.

tion SAPT to the level of atom-pairwise intermolecular interactions. The latter modified A-SAPT to provide a chemically sensible partition to the level of functional-group pairwise intermolecular interaction. Herein, we use the ideas from A-SAPT and F-SAPT to identify separate chemical subsystems in intramolecular interactions, allowing for an “ISAPT” analysis between fragments of a single molecule.

To frame the discussion, consider 2,4-pentandiol, depicted in Figure 1(a). We seek to develop a SAPT procedure to elucidate the strength and nature of the A \cdots B interaction between the two hydroxyl groups, within the linking pentane C.

The first task is to assign particles (electrons and protons) to the fragments A and B and the linker C. If the complete HF wavefunction is localized by a robust orbital localization methodology, the chemical outline of the system becomes immediately obvious. We then partition the system into neutral A, B, C, and link bond fragments (note that formally charged A, B, and C fragments are possible, but are not encountered in this work). For example, for 2,4-pentandiol, the hydroxyl groups A and B each consists of the O 1s core orbital, the O–H σ orbital, the two O lone pairs, and eight protons (one from H and seven from O). For the pentane group C, we initially identify the five C 1s core orbitals, the four C–C σ bonds, and the ten C–H σ bonds, along with 38 corresponding protons. Further, two linking C–O σ bonds are present (one between A

and C and one between B and C), each of which consists of a C–O σ orbital and one proton from each C and O. For the moment, let us assign these linking σ bonds to the linker C; other choices will be discussed below. This scheme of fragment identification is identical to that used in F-SAPT.⁹

At this point, the particle assignment to the fragments is complete (and, given a list of the atoms in A, B, and C, can be completely automated by examination of the sum of orbital atomic charges). However, all three fragments in the system are fully interacting at a HF level; in particular, the wavefunctions of fragments A and B are interacting and orthogonal. The next task is to provide appropriate zeroth-order wavefunctions with the A \cdots B interaction and orthogonality constraint removed. We might be tempted to find separately noninteracting wavefunctions for A, B, and C, but this does not make chemical sense; the infinite-order HF interactions and orthogonality constraints A \cdots C and B \cdots C are critical to provide the correct topology of the covalent links between these fragments and must be retained if one is to extract the noncovalent part of the interaction A \cdots B later. To obtain the zeroth-order wavefunctions for A and B, each with the interactions with C built in, we invoke HF-in-HF embedding as outlined by Manby and coworkers.¹⁰ That is, we perform a standard HF computation for the protons and electrons assigned to fragment A, but with the Coulomb and exchange potentials of the particles of C provided as a spin-specific external potential $\hat{W}_C^\alpha \equiv \hat{V}_C + \hat{J}_C - \hat{K}_C^\alpha$ and with the occupied orbitals of C deleted from the subspace in which the Fock matrix for A is diagonalized. Here, \hat{V}_C is the Coulomb potential of the protons of C, \hat{J}_C is the Coulomb potential of the electrons of C, and \hat{K}_C^α is the exchange potential of the α electrons of C.

At this point, one obtains a set of relaxed occupied and virtual orbitals and corresponding eigenvalues for A and B, each embedded in the HF wavefunction for the linker C, but noninteracting and nonorthogonal between A and B. Next, one may simply substitute the orbitals, eigenvalues, and electrostatic potentials of the A and B systems into a standard two-body SAPT methodology and obtain the SAPT analysis of the A \cdots B interaction, with C built in as a modification of the vacuum.

The overall ISAPT procedure is summarized as follows:

1. Compute the HF wavefunction of the complete chemical system.
2. Localize the occupied orbitals.
3. Assign local occupied orbitals and protons to A, B, and C by consideration of orbital atomic charges.
4. Relax the HF wavefunction for A to be noninteracting and nonorthogonal to B, but fully interacting and orthogonal to C (via embedding the HF wavefunction for A in the frozen HF wavefunction for C).
5. Repeat the previous step for B.
6. Perform a standard two-body SAPT analysis between the relaxed HF wavefunctions for A and B.

All computations have been performed in a development version of Psi4,¹¹ using density-fitted SAPT0 (DF-SAPT0) in the jun-cc-pVDZ basis set (to benefit from favorable relative error cancellation)^{4,12,13} and appropriate -JKFIT and -MP2FIT auxiliary basis sets. The full system’s basis set is

used throughout, analogous to a dimer-centered basis set in standard intermolecular SAPT. Localization to identify the occupied orbitals of C is accomplished by the robust intrinsic bond orbital (IBO) methodology of Knizia¹⁴ with the cc-pVTZ cores as the minimal AO basis set. To maintain orthogonality against C in the HF-in-HF embedding procedure, we simply diagonalize the Fock matrix in a reduced subspace with the columns spanning the occupied orbitals of C removed. The $\delta_{\text{HF},r}^{(2)}$ correction to induction is applied throughout this work, using the appropriate HF interaction energy for the zeroth-order wavefunctions for A and B embedded in C.

The chemical soundness of our approach is indicated by Figures 1(b) and 1(c), which compares the zeroth-order density fields and difference density fields upon activation of the HF interaction between the hydroxyls in 2,4-pentanediol and the corresponding water dimer. The qualitative shapes of the density fields and difference density fields are remarkably similar, with the only deviations occurring where the σ linkers would be assigned to C in this case. In particular, the key changes caused by the interaction are faithfully reproduced, e.g., the interstitial polarization and charge transfer in B are identical between ISAPT and standard supermolecular SAPT.

Figure 2 compares the ISAPT0/jun-cc-pVDZ interaction energy components for 2,4-pentanediol with SAPT0/jun-cc-pVDZ for the corresponding water dimer (this latter approach is often referred to as a “cut-and-cap” analysis). To explore the choices of linking σ bond assignment, we have developed a code that can assign the link bonds wholly to C, wholly to A or B, or evenly between C and A or B (in this last case, the link bonds are frozen in second-order terms and provide only electrostatic potentials to the $A \cdots B$ interaction). Overall, the interaction energy components are in excellent agreement with cut-and-cap analysis. The agreement is particularly good for exchange, induction, and dispersion. For these terms, assignment of the linking σ bonds to A or B results in a larger magnitude of each interaction term, as more occupied orbitals are involved in steric overlap or induction/dispersion response. The disappointing case is electrostatics with assignment of the linking σ bonds to A or B. Here, the protons taken from the C atoms are largely unshielded by the linking σ bond, giving rise to repulsive electrostatic contributions from the $\sigma \cdots \sigma$ multipoles. This contribution is chemically unexpected, but is correct given the choice of particle assignment to fragments. We note that such unexpected multipole-multipole electrostatic interactions can also occur in other systems if the linking σ bonds are given to C (see below for such examples as dipole-dipole interactions in methyl-methyl interactions). At present, we can offer no solution to this issue, but note that this anomaly is certainly smaller than the charge-charge or radical contamination inherent in many other choices of molecular partition and is consistent with our definition of zeroth-order wavefunction.

A more torturous test is the noncovalent interaction of the methyl groups in propane, depicted for various C–C–C bond angles in Figure 3. Note that for this system, cut-and-cap analysis is not even qualitatively correct; the capping protons on the corresponding methyl dimer would be less than 1 Å apart and the resultant steric clash would overwhelm the analysis. Here, ISAPT is able to produce a smooth and largely

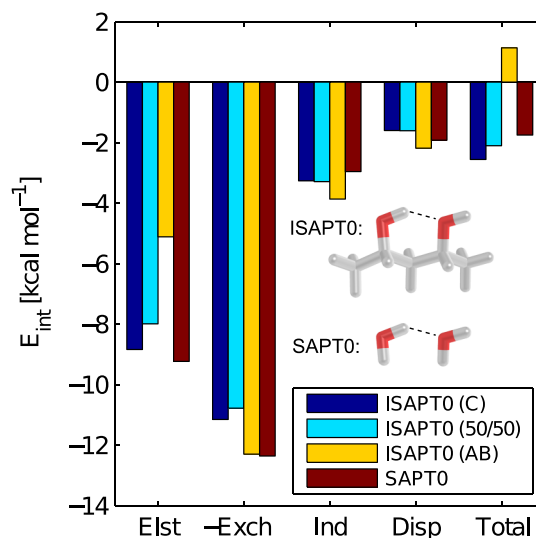


FIG. 2. Comparison of ISAPT0/jun-cc-pVDZ for the hydroxyl groups of 2,4-pentanediol to standard intermolecular SAPT0/jun-cc-pVDZ for the corresponding water dimer. Note that $-E_{\text{exch}}$ is plotted for convenience. ISAPT partitions for the linking σ bonds assigned to C, shared evenly between C and A or B, or assigned to A or B are depicted.

sensible picture, by defining each of the methyl units A and B as the three C–H σ bonds plus the C $1s$, embedded in the field of the CH_2 linking unit. At small angles, exchange dominates the interaction, though significant dispersion, induction, and charge penetration (electrostatic) contributions are also apparent. The exchange term exhibits a minimum at around 132° , beyond which exchange increases slightly, consistent with a mild overlap developing between the two previously uninvolved C–H σ bonds (the outermost two C–H σ bonds in the figure). At large angles, the interaction is dominated by repulsive electrostatic interactions, which are consistent with the anti-aligned dipoles present in the zeroth-order wavefunctions for the methyl groups. These dipoles arise from the sum

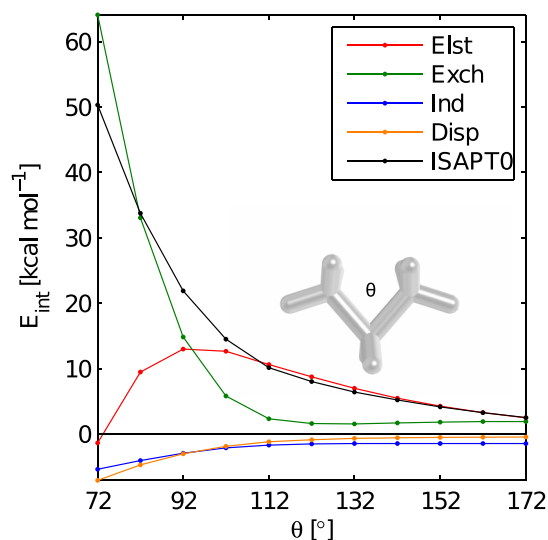


FIG. 3. ISAPT0/jun-cc-pVDZ analysis for the interaction between the two terminal methyl groups in propane vs. the C–C–C bond angle θ . Links are assigned to C.

of the three C–H bond dipoles of each methyl group, which are not compensated by the link bond (which is assigned to C). Several alternative choices of link bond assignment are not appropriate in this case: if the link bonds are shared or wholly assigned to A or B, the overlap between the two linking σ bonds causes enormous charge penetration contributions to the electrostatics and induction terms. Thus, although ISAPT reasonably reflects the dependence of the energy components on the geometry, the absolute value of the 1,3-methyl-methyl interaction (central to the protobranching debate^{15–19}) remains to be accurately determined once these artifacts are eliminated.

A key example of the utility of ISAPT is hexaphenyl-ethane (5H), and its *t*-Bu-substituted derivative (5T), depicted in Figure 4. These compounds have recently been of substantial interest in the experimental literature because the latter has been synthesized and found to have one of the longest C–C σ bonds known (on the order of 1.67 Å), but the former has not been experimentally observed. This is perhaps counterintuitive, as one would expect substantially more steric hindrance in the *t*-Bu-substituted variant. Schreiner²⁰ and later Schreiner and Grimme²¹ have hypothesized that 5T does experience increased steric clash between the two groups on opposite sides of the long σ bond, but this is countered by an even greater dispersion stabilization afforded by extra *t*-Bu \cdots *t*-Bu and *t*-Bu \cdots Ph interactions. To this point, however, no direct quantification of the interaction energy components has been possible for this class of systems, principally because they are intractable with a cut-and-cap style analysis. ISAPT can easily handle these systems, by defining the link unit C as the long σ bond and assigning the remainder to A and B. Note that the theoretical [TPSS-D3(BJ)/TZV2P] central bond length of 5T is 1.661 Å, whereas the one in 5H is 1.713 Å (significantly longer). To determine the substituent effects without contamination from geometric relaxation, we perform a difference ISAPT0/jun-cc-pVDZ analysis at geometries for 5H and 5T both constrained to the 5T equilibrium bond distance (blue in Figure 4).²² Repeating this same analysis at the 5H equilibrium bond distance reveals nearly quantitatively the same picture (red in Figure 4). Below, we quote the results for the 5T equilibrium bond distance.

The increase in ISAPT exchange repulsion for 5T vs. 5H is ~ 24 kcal mol^{–1}, while the increase in dispersion attraction is ~ 34 kcal mol^{–1}: dispersion interactions compensate for increased steric repulsion by ~ 10 kcal mol^{–1}. However, ISAPT reveals another, less intuitive physical contribution: electrostatic repulsion *decreases* upon substitution by ~ 10 kcal mol^{–1}, possibly from charge penetration or CH- π contacts. The induction contributes a rather modest ~ 1 kcal mol^{–1}, favoring 5T. Overall, the *t*-butyl substitution enhances the intramolecular interaction energy by ~ 22 kcal mol^{–1}, with roughly equal contributions from exchange + dispersion (hypothesized by Schreiner and Grimme) and electrostatics (a new finding).

Note that the insights into the substituent effect above are only possible by performing a carefully designed ISAPT experiment in which the effects of the anti-aligned dipoles on monomers A and B cancel between 5H and 5T—this mandates the use of a common central bond distance above. If one considers the ISAPT interaction energies at the respective

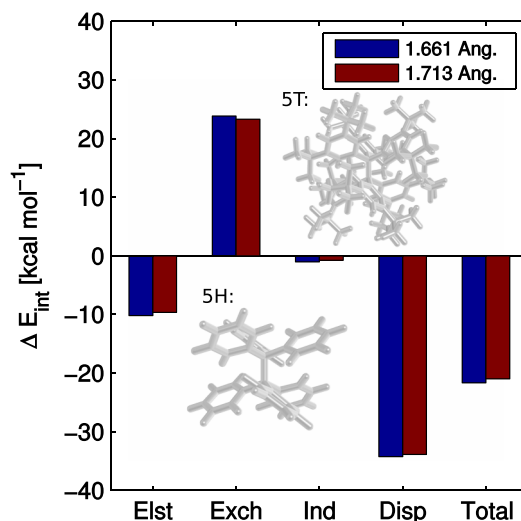


FIG. 4. ISAPT0/jun-cc-pVDZ difference analysis obtained by subtracting the hexaphenyl-ethane (5H) results from the all-*meta-t*-Bu-hexaphenyl-ethane (5T) results, at the fixed central C–C bond length of 5T (1.661 Å) or 5H (1.713 Å).

equilibrium bond distances of 5H and 5T, the closer contact between the anti-aligned dipoles in 5T leads to unfavorable dipole-dipole contributions to 5T, muddling the analysis to the point that the favorability of 5T is reduced to only ~ 10 kcal mol^{–1}. This exemplifies the problem of chemically unexpected dipoles in ISAPT, but demonstrates that useful results may still be obtained with a judiciously chosen computational experiment.

It is apparent that ISAPT allows for quantification of non-covalent interactions in systems for which cut-and-cap analysis is unsuitable. The method is easy to implement within an existing two-body SAPT code and exhibits the same tractability limit. Future developments will involve the application of F-SAPT techniques to the methodology (e.g., to elucidate the fraction of CH- π vs. π - π interactions in 5T) and efforts to mitigate the chemically unexpected multipole-multipole interactions near the linking σ bonds.

R.M.P. is a DOE Computational Science Graduate Fellow (Grant No. DE-FG02-97ER25308). J.F.G. acknowledges the Swiss NSF fellowship (Grant No. P2ELP2_155351) for the financial support. C.C. acknowledges the Swiss NSF Grant No. 200021_156001. This work was supported by the National Science Foundation (Grant No. CHE-1300497).

¹K. Kitaura and K. Morokuma, *Int. J. Quantum Chem.* **10**, 325 (1976).

²Y. Mo, J. Gao, and S. D. Peyerimhoff, *J. Chem. Phys.* **112**, 5530 (2000).

³R. J. Azar and M. Head-Gordon, *J. Chem. Phys.* **136**, 024103 (2012).

⁴B. Jeziorski *et al.*, in *Methods and Techniques in Computational Chemistry: METECC94*, Medium-Size Systems Vol. B, edited by E. Clementi (STEF, Cagliari, 1993), p. 79.

⁵B. Jeziorski, R. Moszynski, and K. Szalewicz, *Chem. Rev.* **94**, 1887 (1994).

⁶J. F. Gonthier and C. Corminbœuf, *J. Chem. Phys.* **140**, 154107 (2014).

⁷I. Mayer, *Int. J. Quantum Chem.* **23**, 341 (1983).

⁸R. M. Parrish and C. D. Sherrill, *J. Chem. Phys.* **141**, 044115 (2014).

⁹R. M. Parrish, T. M. Parker, and C. D. Sherrill, *J. Chem. Theory Comput.* **10**, 4417 (2014).

¹⁰F. R. Manby, M. Stella, J. D. Goodpaster, and T. F. Miller, *J. Chem. Theory Comput.* **8**, 2564 (2012).

- ¹¹J. M. Turney *et al.*, *WIREs Comput. Mol. Sci.* **2**, 556 (2012).
- ¹²E. G. Hohenstein, R. M. Parrish, C. D. Sherrill, J. M. Turney, and H. F. Schaefer, *J. Chem. Phys.* **135**, 174107 (2011).
- ¹³T. M. Parker, L. A. Burns, R. M. Parrish, A. G. Ryno, and C. D. Sherrill, *J. Chem. Phys.* **140**, 094106 (2014).
- ¹⁴G. Knizia, *J. Chem. Theory Comput.* **9**, 4834 (2013).
- ¹⁵S. Gronert, *J. Org. Chem.* **71**, 1209 (2006).
- ¹⁶S. Gronert, *J. Org. Chem.* **71**, 9560 (2006).
- ¹⁷S. Gronert, *Chem. - Eur. J.* **15**, 5372 (2009).
- ¹⁸M. Wodrich *et al.*, *Chem. - Eur. J.* **13**, 7731 (2007).
- ¹⁹W. C. McKee and P. v. R. Schleyer, *J. Am. Chem. Soc.* **135**, 13008 (2013).
- ²⁰P. R. Schreiner *et al.*, *Nature* **477**, 308 (2011).
- ²¹S. Grimme and P. R. Schreiner, *Angew. Chem., Int. Ed. Engl.* **50**, 12639 (2011).
- ²²See supplementary material at <http://dx.doi.org/10.1063/1.4927575> for geometries, additional figures, and discussion of IBO localization solutions in 5H/5T.