JCTC Journal of Chemical Theory and Computation

Non-Covalent Interactions with Dual-Basis Methods: Pairings for Augmented Basis Sets

Ryan P. Steele,[‡] Robert A. DiStasio, Jr., and Martin Head-Gordon*

Department of Chemistry, University of California, Berkeley, California 94720

Received February 1, 2009

Abstract: Basis set pairings for dual-basis calculations are presented for the aug-cc-pVXZ (X = D, T, Q) series of basis sets. Fidelity with single-basis results is assessed at the second-order Møller−Plesset perturbation theory (MP2) level within the resolution-of-the-identity (RI) approximation, using the S22 set of noncovalent interactions and a series of electron affinities from the G3 set. Root-mean-squared errors for the S22 set are 0.019 kcal mol^{-1} or lower, with a maximum deviation of 0.44%, and errors in nuclear structures are 0.09% or lower. Cost savings of 60−93% (RI-MP2 energies) and 50−88% (RI-MP2 gradients) are demonstrated. Spin-component-scaled MP2 [SCS(MI)-MP2] scaling parameters are provided for the aug-cc-pVXZ series, and dual-basis results are shown to be consistent without reoptimization of the single-basis parameters. Explicit handling of linear dependence in the basis set projection scheme is also provided. These dual-basis pairings will be helpful for accelerating accurate Hartree−Fock, density functional theory (DFT), MP2 and scaled MP2, and so-called doubly hybrid DFT calculations of intermolecular interactions (and other systems), where augmented basis sets are physically important.

1. Introduction

Noncovalent interactions, encompassing $\pi-\pi$ stacking and dispersion forces, hydrogen-bonding, multipole−multipole interactions, and combinations thereof, have proven to be extremely difficult properties to quantitatively predict. Electrostatics alone cannot account for the existence of most of these phenomena, and thus, accurate electron correlation must be included to describe them properly. Further complicating matters, most wavefunction-based methods that properly account for electron correlation are extremely sensitive to the underlying atomic orbital (AO) basis.[1] These basis sets typically require functions with high angular momentum to account for polarization within and between molecules, as well as diffuse exponents to properly account for the long-range nature of the interactions. This combination of accurate electron correlation and large basis sets pushes the capability frontiers of modern electronic structure theory and has limited application to small prototype systems.

Density functional theory (DFT) within the Kohn−Sham formalism[2] has become the de facto method of choice for many theoretical applications. By combining parametrized electron correlation with a cost roughly equivalent to Hartree−Fock (HF), DFT typically performs quite well for thermochemistry[3−7] and molecular structures.[5−8] However, DFT functionals employing the local density approximation (LDA)[9] do not account for the inherently nonlocal effect of dispersion.[10] Even gradient-corrected (GGA)[11−16] functionals are corrections to the local density and do not treat dispersion forces properly. Empirical dispersion corrections (DFT-D)[17−20] have recently gained favor and perform remarkably well for interaction energies, sometimes approaching highly accurate coupled-cluster results. Such terms are purely empirical, however, and have their limitations, such as the neglect of the response of the electron density to the C_6 term. First-principles nonlocal dispersion functionals are also under development and show promise.[21−24]

Another successful method for interaction energies is symmetry-adapted perturbation theory.[25] By constructing the interaction energy between two (or more) distinct subunits directly, expensive computations on the entire system are

* To whom correspondence should be addressed. E-mail: mhg@bastille.cchem.berkeley.edu.

‡ Current address: Department of Chemistry, Yale University, 225 Prospect St., New Haven, CT 06405.

10.1021/ct900058p CCC: $40.75
Published on Web 04/30/2009

unnecessary. Results are typically accurate, and a recent study by Szalewicz[26] demonstrated that, when combined with DFT for the monomers' energies and orbitals, SAPT-DFT[27−32] could effectively and efficiently describe the entire potential energy surface of the benzene dimer, a well-studied prototype system[33−39] known to require extremely accurate electron correlation. These methods also allow for the decomposition of interaction energies into contributions for electrostatics, dispersion, etc., a key tool in determining the factors controlling an interaction. These methods are also not without their shortcomings, however. Applicability is somewhat limited; assessment of conformational energy differences[40] in polypeptides, proteins, or any system with *intra*molecular bonding would not be feasible in the absence of distinct subsytems. Furthermore, a description of chemistry (i.e., broken or changing bonds) is, as yet, undefined in these models.

Thus, an *ab initio* description of electron correlation with the supermolecule approach is still often necessary. The simplest treatment of correlation is second-order Møller−Plesset perturbation theory (MP2).[41] Coupled with the resolution-of-the-identity (RI) approximation,[42−45] MP2 provides an accurate estimate (∼90%) of correlation energies at a cost lower than the underlying self-consistent field (SCF) for many systems of interest. Unfortunately, the same practical shortcoming of all correlated methods also applies to MP2 theory: correlation energies are slowly convergent with respect to the underlying AO basis set.[1] Finally, it should be mentioned that scaling the spin components of the MP2 energy[46−53] has been demonstrated to give improved accuracy, in a statistical sense, both for covalent bond/reaction energies[46,48−51] and intermolecular interactions.[47,52,53] Unfortunately, significantly different scalings are required for these two different classes of interactions. The applicability of these methods in the current context will be discussed in section 3.1.4.

An interesting, practical amalgam of ab initio and DFT calculations has also recently emerged. So-called "doubly hybrid" density functionals[54−60] attempt to correct the poor performance of local functionals for nonlocal properties by including an orbital-dependent, MP2-type term, ideally without overcounting electron correlation or sacrificing accuracy in local properties. Several successful attempts at these qualities have appeared. The most important property of these methods in the present context, however, is that these methods, owing to the wavefunction-like correction, are once again strongly basis set dependent. They are, therefore, also prime candidates to benefit from the basis set pairings presented here.

The dual-basis (DB) method[61−67] has proven to be an accurate alternative to large basis set calculations and provides computational savings of 90% in this regime. In short, the DB scheme involves a standard, iterative SCF calculation in a small subset of the larger, target basis. A perturbative correction (amounting to an approximate Roothaan step) is applied to capture basis set relaxation effects. This scheme defines both DB-HF and DB-DFT. In addition, subsequent correlation corrections could also be added; doing so within the second-order perturbation theory framework defines the DB-MP2 method that is the main focus of this work.

The DB-RI-MP2 method can efficiently produce correlated-electron calculations approaching the AO basis set limit. We have previously presented basis set pairings for 6-31G*,[67] 6-311++G(3df,3pd), cc-pVTZ, and cc-pVQZ.[62] A practical need remains to provide pairings for the aug-cc-pVXZ (X = D,T,Q) series of Dunning-style basis sets,[68−70] commonly used for these noncovalent interactions as well as anionic systems. These basis sets are well-suited to cases where long-range interactions are present and where significant "in−out" flexibility is required in the basis. Frequently, augmented double-ζ results are of comparable quality with the much more expensive nonaugmented quadruple-ζ results for these systems; saturating the diffuse space often leads to faster convergence than higher angular momentum. Furthermore, the aug-cc-pV(D,T,Q)Z series is well-suited to extrapolation schemes,[71,72] with which results may be obtained at the equivalent of one higher angular momentum in the basis set, at essentially no cost. Dual-basis subsets for these bases are constructed and tested herein on several such systems. The main conclusion of this work is that subsets exist that faithfully reproduce target basis quantities, with cost savings equal to or greater than those previously seen for more compact basis sets.

Worth mentioning is a complementary set of correlated ab initio theories. The explicitly correlated methods, typically termed MP2-R12 (or MP2-F12) methods[73] in the context of MP2 theory, exploit the fact that correlation energies converge more slowly than SCF energies with respect to basis set size. These methods are also, in a sense, dual-basis methods; the correlation energies are computed in an additional basis of product functions, which provide basis-set-limit correlation energies. Recent efficient implementations with density fitting,[74,75] local approximations,[76,77] and the RI approximation[78−80] to the costly three- and four-electron integrals are working toward making these methods practical for large systems, and recent examinations of computational cost are very encouraging. Our experience indicates that double- or triple-ζ basis (where R12 energies are often calculated) SCF energies are still not fully converged with respect to basis set, indicated by the fact that the dual-basis correction is nonzero. Therefore, the R12 idea is an exciting correlation counterpart of the dual-basis idea, and future combination of the two methods would be even more efficient than either one individually.

We also note that in Wolinski and Pulay's demonstration of DB-MP2,[63] two truncations of the large aug-cc-pV5Z basis were presented. While this work focuses on the more pragmatic aug-cc-pV(D,T,Q)Z series, it is a testament to DB methods that aug-cc-pV5Z calculations are attainable. After discussing the design of our chosen basis set pairings, this larger basis set's truncation will be discussed in context.

2. Design of Basis Set Pairings

The dual-basis SCF (DB-SCF) method consists of a first-order approximation to basis set relaxation effects and is described in detail in refs 61 and 64 with available extensions

to RI-MP2[62] and its analytic gradient.[65] In short, an iterative self-consistent field (SCF) calculation is performed to convergence in a subset of the larger, target AO basis set, symbolically represented as ⟨target⟩ ← ⟨small⟩, as in 6-311G* ← 6-311G. A single Fock matrix F is constructed in the large basis and is subsequently diagonalized. The resulting molecular orbital (MO) coefficients are then used to form the DB energy correction to the SCF energy, as well as the subsequent correlation calculation:

$$E_{\mathrm{MP2}}^{\mathrm{dual}} = (E_{\mathrm{SCF}}^{\mathrm{small}} + \Delta E_{\mathrm{SCF}}^{\mathrm{target}}) + \Delta E_{\mathrm{MP2}}^{\mathrm{target}} \quad (1)$$

where

$$\Delta E_{\mathrm{SCF}}^{\mathrm{target}} = Tr(\Delta P \cdot F[P]) \quad (2)$$

Here, $\Delta P = P' - P$ is the relaxation of the density matrix upon diagonalization of the single large-basis Fock matrix, $F[P]$.

Savings are most significant when the smaller basis set is a strict subset of the target basis because of integral screening during the large Fock build, as well as during several steps in the analytical gradient. Forming subsets with a sufficiently small basis set ratio (small:target) that still preserve the high accuracy for which these basis sets are intended is the central design challenge.

The aug-cc-pVXZ series is not comprised of sequential supersets; therefore, manual construction of viable subsets is necessary, as was done for cc-pVTZ and cc-pVQZ.[62] Much like the target basis sets themselves, construction of the subsets is inherently empirical but may be guided by judicious chemical insight. Using these same tenets, we have constructed the strongest feasible truncations to the aug-cc-pVXZ series, with the lone restriction that DB results faithfully reproduce full-basis results. Several truncations were considered and tested for each target basis, and results are presented in section 3.

Accurate calculation of properties of noncovalent systems or anionic systems is typically the goal when utilizing these very diffuse basis sets because differential basis set effects (and differential electron correlation, which is strongly basis set dependent) are often crucial. Accordingly, test sets of these systems have been used to guide our empirical construction of subsets. Specifically, we have used the following two test sets for benchmarks:

• The S22 set[81] of Jurecka et al. consists of 22 noncovalent dimer systems, for which either accurate experimental or coupled-cluster binding energies are known. The set consists of 8 dispersion-dominated complexes, 7 hydrogen-bonding complexes, and 7 complexes containing a mixture of these two interactions. This set allows for the accurate parametrization of subsets for both heavy and hydrogen atoms.

• A 25-molecule/atom subset of the G3 set[82−84] is used to calculate adiabatic electron affinities (EA). In both cases, the resultant energy differences are highly sensitive to basis set quality and provide stringent tests of our truncations. The focus of this work is methodology for noncovalent systems. As such, the results for the S22 set have guided our choice of truncation schemes; results for the EA set using these pairings are provided for completeness.

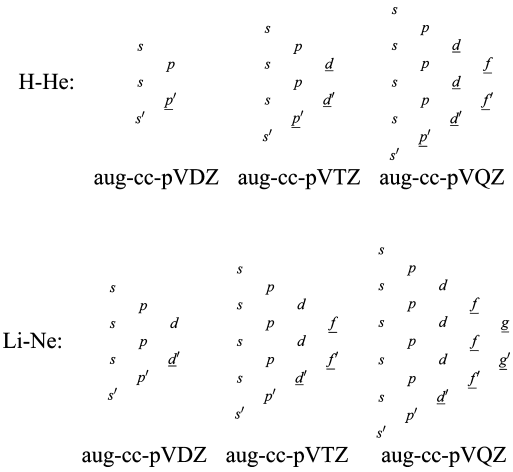

Figure 1. Structure of the aug-cc-pV(D,T,Q)Z series for first- and second-row atoms. The most compact functions are listed at the top of each set, and primed functions depict aug (diffuse) functions. Underlined functions are those eliminated in the truncated dual-aug-cc-pV(D,T,Q)Z series.

The energies of all systems were computed at the RI-MP2 and DB-RI-MP2 levels of theory, using fixed geometries. In all cases, SCF calculations were converged to a DIIS[85,86] error of at most 10^{-8} a.u., using integral thresholds of 10^{-12} a.u. For the RI calculations, the corresponding auxiliary basis sets of Weigend[87] were used. Linear dependence was handled as described in Appendix A, using a drop tolerance of 10^{-6}. The frozen core approximation was used in all correlated (RI-MP2) calculations.

The balanced structure of the aug-cc-pV*X*Z basis sets is shown in Figure 1. Lessons learned from previous DB truncations[62,67] help guide the choice of subsets. In particular, some polarization functions are required in the small basis for quantitative agreement with target basis properties.[67] The new design challenge for the augmented basis sets is, thus, choosing which diffuse functions may be discarded, as well as the proper balance between polarization and diffuse functions. All viable candidates for truncations of aug-cc-pV(D,T)Z are given in the Supporting Information (SI) and categorized according to the truncation level on heavy and hydrogen atoms.

The qualitative standards we sought to satisfy are as follows:

(1) The error caused by using the DB method should be sufficiently less than the difference between basis sets (i.e., between aug-cc-pVDZ and aug-cc-pVTZ).
(2) The accuracy should be balanced across anions, dispersion complexes, and hydrogen-bonded complexes.
(3) The cost savings resulting from the truncation should be on par with, or greater than, previous DB pairings (∼90−95% for RI-MP2 energies).

Because of the cost and number of possible permutations, the results for aug-cc-pVDZ and aug-cc-pVTZ were used to guide seven possible truncations for the aug-cc-pVQZ basis, also provided in the Supporting Information.

Table 1. DB-RI-MP2 Basis Pairing Results for aug-cc-pV(D,T,Q)Z on the S22 Set[81] [a]

	basis	relative to	CH ratio[b]	non-CP[c] rms[d]/ kcal mol^{-1}	CP rms/ kcal mol^{-1}
dual	aug-cc-pVDZ ← dual-aug-cc-pVDZ	aug-cc-pVDZ	0.750	0.085	0.043
	aug-cc-pVTZ ← dual-aug-cc-pVTZ	aug-cc-pVTZ	0.536	0.034	0.019
	aug-cc-pVQZ ← dual-aug-cc-pVQZ	aug-cc-pVQZ	0.357	0.020	0.015
single	cc-pVDZ	aug-cc-pVDZ		2.629	2.151
	cc-pVTZ	aug-cc-pVTZ		0.740	0.821
	aug-cc-pVDZ	aug-cc-pVTZ		1.462	0.730
	aug-cc-pVTZ	aug-cc-pVQZ		0.756	0.269

	basis	relative to	C_{OS}/C_{SS}[e]	non-CP rms/ kcal mol^{-1}	CP rms/ kcal mol^{-1}
dual	aug-cc-pVDZ ← dual-aug-cc-pVDZ	CCSD(T)/CBS		3.76	1.20
	aug-cc-pVTZ ← dual-aug-cc-pVTZ			2.29	1.04
	aug-cc-pVQZ ← dual-aug-cc-pVQZ			1.57	1.10
	aug-cc-pVDZ ← dual-aug-cc-pVDZ		0.00/1.83		0.26
	aug-cc-pVTZ ← dual-aug-cc-pVTZ		0.22/1.52		0.26
	aug-cc-pVQZ ← dual-aug-cc-pVQZ		0.31/1.40		0.26
single	cc-pVDZ	CCSD(T)/CBS		1.88	2.49
	cc-pVTZ			1.72	0.91
	cc-pVQZ			1.46	0.90
	aug-cc-pVDZ			3.73	1.19
	aug-cc-pVTZ			2.29	1.06
	aug-cc-pVQZ			1.57	1.11
	aug-cc-pVDZ		0.00/1.83		0.26
	aug-cc-pVTZ		0.22/1.52		0.26
	aug-cc-pVQZ		0.31/1.40		0.26

[a] The results shown in the upper section are errors (kcal mol^{-1}) relative to calculations in the basis shown in the second column, and those shown in the lower section are errors (kcal mol^{-1}) relative to complete basis set (CBS) estimates of coupled-cluster singles and doubles with perturbative triples (CCSD(T)). [b] Basis set size ratio for a molecule containing an equal number of heavy and hydrogen atoms. [c] rms = Root mean squared deviation. [d] CP = Counterpoise-corrected. [e] SCS(MI) scaling coefficients for same-spin (SS) and opposite-spin (OS) components of the MP2 correlation energy, refit to aug-cc-pV(D,T,Q)Z data using the method of ref 52.

Table 2. DB-RI-MP2 Basis Pairing Results for aug-cc-pV(D,T,Q)Z on a 25-Molecule Electron Affinities Subset of the G3 Set[82−84a]

	basis	relative to	rms[b]/kcal mol^{-1}
dual	aug-cc-pVDZ ← dual-aug-cc-pVDZ	aug-cc-pVDZ	0.109
	aug-cc-pVTZ ← dual-aug-cc-pVTZ	aug-cc-pVTZ	0.143
	aug-cc-pVQZ ← dual-aug-cc-pVQZ	aug-cc-pVQZ	0.261
	aug-cc-pVDZ ← dual-aug-cc-pVDZ	expt	6.256
	aug-cc-pVTZ ← dual-aug-cc-pVTZ	expt	5.397
	aug-cc-pVQZ ← dual-aug-cc-pVQZ	expt	5.549
single	cc-pVDZ	aug-cc-pVDZ	29.65
	aug-cc-pVDZ	aug-cc-pVQZ	3.907
	aug-cc-pVTZ	aug-cc-pVQZ	1.162
	aug-cc-pVDZ	expt	6.238
	aug-cc-pVTZ	expt	5.343
	aug-cc-pVQZ	expt	5.501

[a] Results are errors (kcal mol^{-1}) relative to the method shown in the second column. [b] rms = Root mean squared deviation.

3. Results

Optimal basis set pairing schemes are depicted in Figure 1. Summarized results for the S22 and EA sets are shown in Tables 1 and 2 for these pairings, along with pertinent basis set size information. A full statistical analysis of these results and the raw data are included in the Supporting Information for all tested pairings. The results for each basis set (D,T,Q) are here briefly discussed in turn.

3.1. Energies. *3.1.1. aug-cc-pVDZ.* The aug-cc-pVDZ basis[68] has a relatively limited number of feasible truncation schemes, but unlike 6-31G*,[88−90] three sets of diffuse functions are available for truncation. As the aug-cc-pVXZ series is often considered overly diffuse,[91] these diffuse functions are viable candidates for elimination. The removal of the diffuse d' functions (and p' functions on hydrogen) serves as a useful first attempt. Results for this pairing are quite good, as root-mean-squared (rms) errors are only 0.085 kcal mol^{-1} for the S22 set and 0.109 kcal mol^{-1} for electron affinities, relative to aug-cc-pVDZ results. The same errors for a single-basis cc-pVDZ calculation are 2.629 and 29.654 kcal mol^{-1}. Alternative elimination of the polarization functions, while retaining the diffuse functions, leads to significantly worse DB results (see Supporting Information). Given that the cost of this second pairing is actually greater

than the cost of the first, because of the inevitably larger number of significant shell pairs in the small basis, this pairing was eliminated. Once again, we find that retention of at least some polarization is necessary for accurate replication of target-basis properties. For completeness, a more drastic truncation was constructed, in which the two sets of diffuse functions of highest angular momentum were eliminated. This pairing performs the worst of the three, despite retention of polarization functions in the small basis. The lone set of diffuse s' functions on heavy atoms is not sufficient to account for these diffuse properties; in this small basis, even the aug functions serve the role of polarization functions. These trends appear to be constant for heavy and hydrogen atoms. No mixing of these pairing combinations performs better. Thus, we suggest the first pairing for use with aug-cc-pVDZ calculations, shown in Figure 1.

3.1.2. aug-cc-pVTZ. The scope for computational savings without loss of accuracy is greater at the aug-cc-pVTZ level and will be the main focus of our discussion. Using the previous 4s3p2d truncation of cc-pVTZ as a guide, the most logical extension was the additional elimination of the diffuse f' functions. Errors are small for this pairing, only 0.015 kcal mol^{-1} error for the S22 set. Three further truncations were attempted by removing one set of d functions, all of which lead to greater savings. Results for all of these pairings demonstrate rms errors below 0.08 kcal mol^{-1} for the S22 interaction energies. While removal of the central d performs nearly as well as the more costly f-only truncation, a more cost-effective option is available. Removal of the diffuse d' function performs just as well on the S22 set and nearly four times better for electron affinities. Thus, while augmentation of the basis set is crucial for electron affinities, accurately capturing polarization effects is also a necessary requirement. This choice is also entirely consistent with the previous nonaugmented cc-pVTZ truncation; for the augmented set presented here, an additional diffuse function has simply been removed.

A similar trend is observed for hydrogen, even in hydrogen-bonded complexes. Thus, the 4s2p truncation is the truncation of choice for aug-cc-pVTZ for hydrogen, in which the diffuse p' function was eliminated. In fact, in all of the optimal sets chosen here, diffuse functions beyond s' were found to be unnecessary for hydrogen.

For this pairing, errors are sufficiently small that dual-basis calculations may serve as a viable replacement for full-basis calculations of intermolecular interactions. CP-corrected rms errors in the S22 set are 0.019 kcal mol^{-1} (max 0.042 kcal mol^{-1}), an average absolute percent error of only 0.25% (max 1.17%). They are also significantly below the errors resulting from the use of aug-cc-pVDZ, one of the main requirements of our truncations. For example, the same errors from using single-basis aug-cc-pVDZ calculations are 0.730 kcal mol^{-1} (max 1.563 kcal mol^{-1}) and 10.3% (max 19.6%).

The results for adiabatic EAs are also reasonable. The rms error (relative to aug-cc-pVTZ) is 0.143 kcal mol^{-1}. The errors relative to experimental EAs for dual- and single-basis calculations are 5.40 and 5.34 kcal mol^{-1}, respectively. Even aug-cc-pVQZ remains 5.50 kcal mol^{-1} from experimental values, indicating that the greater influence in these errors is the MP2 model itself, rather than basis set effects. In this context and given the performance for S22, this pairing is quite adequate.

3.1.3. aug-cc-pVQZ. As is the standard story with dual-basis calculations, both savings and accuracy increase as the target basis becomes more complete. The same is seen for aug-cc-pVQZ here. Three main pairings were considered: a conservative truncation of the two highest angular momentum levels ($2f/f'$ and g/g'), a subsequent removal of diffuse functions of the next lower angular momentum (d'), and a final pairing in which one polarization and one diffuse function were retained in the d level of heavy atoms and the p level of hydrogen atoms. In the latter pairing, the "middle" polarization function was retained. Other permutations of this latter truncation could be performed, but given the expense of these calculations, this balanced truncation seems appropriate. Preliminary tests of more drastic truncations (not shown) proved to be significantly worse.

The least aggressive pairing expectedly performs best, with S22 rms errors of 0.013 kcal mol^{-1} (max 0.037). The estimated SCF savings (see section 3.3) of 90% is already significant, but further truncation is possible. Removal of the next set of diffuse functions produces errors of only 0.015 kcal mol^{-1}, with 93% savings. Thus, the sets of s' and p' diffuse functions are capable of capturing diffuse properties, a fact consistent with the diffuse function scheme in Pople-style basis sets.[92] Alternatively, the last tested pairing includes one set of d polarization functions (the middle of the three) and one set of diffuse d' functions. This set accounts for the aforementioned in−out flexibility and retains some diffuse character, with a CP-corrected rms error of only 0.016 kcal mol^{-1}. However, as was discussed for aug-cc-pVTZ above, a more cost-effective option is the second choice, in which the three polarization functions are retained and the d' diffuse function is eliminated. While this choice leaves three, as opposed to two, d functions in the smaller basis set, the diffuse d' function contributes more significantly to the overall cost (a fact not accounted for in the simple basis set scaling ratios). This choice is also more consistent with the previously published cc-pVQZ truncation. For hydrogen atoms, only s and p functions were required, and only the diffuse s' must be retained.

3.1.4. Spin-Component-Scaled MP2. To correct for deficiencies in MP2 theory, such as overestimating dispersion and poor treatment of atomization energies, Grimme[46] introduced two empirical parameters to the MP2 model, which separately scale the opposite-spin (singlet, OS) and same-spin (triplet, SS) components of the correlation energy. These methods have been termed spin component scaled (SCS) methods. While several flavors of SCS have since been introduced,[48−50] DiStasio[52] recently noted that the optimal scaling of OS and SS components for noncovalent systems is opposite the optimal relative scaling for thermochemical properties. A basis set-dependent set of scaling parameters was suggested for molecular interactions (MI) and termed SCS(MI); the resulting methodology was recently shown to perform the best among scaled MP2 methods for the uracil dimer.[93] Thus, a globally optimized set of empirical parameters appears not to exist, most likely because the empirical

scaling is accounting for different decay properties of the components of the correlation energy. For a well-defined class of systems, however, such as the S22 set considered here, this scaling is reasonable.

In Table 1, SCS(MI) results for single- and dual-basis interaction energies are shown. The original set of scaling parameters was optimized for the (nonaugmented) cc-pV(D,T,Q)Z series of basis sets. We have thus performed the same fitting of coefficients for the single-basis aug-cc-pV(D,T,Q)Z series here; scaling coefficients for OS and SS are also shown in Table 1. The optimal scaling parameters are markedly basis set-dependent (ranging from complete omission of OS components for aug-cc-pVDZ to $c_{OS} = 0.31$ for aug-cc-pVQZ), in contrast to Grimme's original assertion that SCS parameters are independent of basis set for thermochemical properties. The errors, relative to complete-basis-set (CBS) estimates of CCSD(T) binding energies, however, are remarkably basis set-independent when scaled appropriately. In fact, unscaled interaction energy errors also appear to be roughly independent of basis set when CP-corrected, as long as diffuse functions are included in the basis. Nonaugmented cc-pVDZ, for example, performs notably worse.[47,52]

Overall, SCS(MI) results are good for all basis sets considered. The across-the-board rms error of 0.26 kcal mol^{-1}, relative to CCSD(T), is a worthy improvement over the 1.1−1.2 kcal mol^{-1} error seen for unscaled energies. Importantly, dual-basis results are consistent with single-basis results, using the same scaling parameters optimized for single-basis energies.

3.2. Structures. Design of the pairings established above utilized fixed-geometry binding energies as the metric of interest. Structural optimization, however, is an important component for these noncovalent complexes; the recently developed DB-RI-MP2 analytical gradient[65] provides an efficient means for this additional comparison. For practical reasons, an 11-molecule subset of the S22 set was tested, which contains the systems with, at most, one aromatic ring. Again using single-basis MP2/aug-cc-pVQZ results as the benchmark of interest, the (mass-unweighted) center-to-center distances of the 11 molecules were tabulated with the aug-cc-pVXZ series and their DB counterparts. Such a metric is not unique but most directly describes the intermonomer aspect of the structural optimization. Nuclear geometries were optimized to a maximum gradient component of 3×10^{-5} a.u. and either a displacement of 12×10^{-5} a.u. or an energy change of 1×10^{-6} a.u., an order of magnitude tighter than the Q-Chem[94] default tolerances, since the potential energy surfaces involved are relatively flat.

Summarized results are presented as rms errors in Figure 2. The first noteworthy trend is the sizable basis set dependence of the structures. While intramolecular geometries are often less basis set dependent, the intermonomer spacing in these noncovalent complexes relies almost wholly on differential electron correlation and is a stringent test of the methodology presented here. The single-basis cc-pVDZ results, for example, are in error by more than 0.1 Å. Augmentation of the basis set reduces this error by better than a factor of 2. In fact, aug-cc-pVDZ results are, on average, more than a third better than (nonaugmented) cc-pVTZ structures. Even cc-pVQZ produces structures that are still 0.026 Å from structures of its augmented counterpart, and again, the augmented basis set of one lower angular momentum (aug-cc-pVTZ) outperforms the nonaugmented basis.

The DB-RI-MP2 results, obtained significantly faster, as discussed in the following section, are generally consistent with the single-basis results that they were designed to mimic. The rms errors for single- and dual-basis aug-cc-pVDZ are 0.041 and 0.042 Å, respectively. While the overall errors do not exactly reproduce aug-cc-pVQZ results, this augmented double-ζ pairing is the most likely basis set to be utilized for structural optimizations. The excellent fidelity of the DB results with target-basis quantities indicates that this pairing is a viable means for geometry optimizations. Results for DB aug-cc-pVTZ and aug-cc-pVQZ each show further significant improvements in accuracy, although the DB errors relative to the target basis are larger than those for aug-cc-pVDZ. While this observation is distinct from previous results for covalently bound systems,[62,64] it may also be expected from the progressively more aggressive truncation schemes presented herein. Note that the DB error is roughly consistent, however, with the previously published cc-pVQZ pairing. Furthermore, while the DB errors seen here are larger than those demonstrated for covalent systems, the intermonomer spacings (ranging from 1.6 to 3.7 Å) are also significantly larger than typical intramolecular bonds. The largest DB error (aug-cc-pVQZ), for example, is still within an average unsigned error of 0.085% of target-basis separations and is significantly less than the intrinsic error resulting from using aug-cc-pVTZ (0.464%).

The interaction energies of these optimized complexes are presented in the lower panel of Figure 2. Again, strong basis set dependence is demonstrated. The effect of basis set augmentation is less pronounced than in the structures but is still significant. Consistency between single- and dual-basis results is demonstrated, with differential errors in the augmented pairings of 0.03 kcal mol^{-1} or less.

3.3. Timings. Representative nuclear force timings are presented in Figure 3. All timings were performed on a single 2 GHz Opteron processor, using the same calculation parameters described above. The timings in the figure are broken down into constituent contributions.

For a molecule with an equal number of heavy and hydrogen atoms, the basis set truncation ratio for the aug-cc-pVDZ pairing is 0.750, leading to an estimated cost savings of 60% in the SCF calculation (see ref 62 for cost estimation formulas; 12 SCF cycles have been assumed for estimation purposes). For single-point energies, this savings is significant. For analytic SCF (Hartree−Fock or density functional theory) gradients, this truncation simply is not drastic enough to produce more than modest savings (28%). Using the recently developed DB-RI-MP2 analytical gradient,[65] however, we expect to see somewhat more significant savings. While the savings are system size-dependent, a simple timings comparison for the

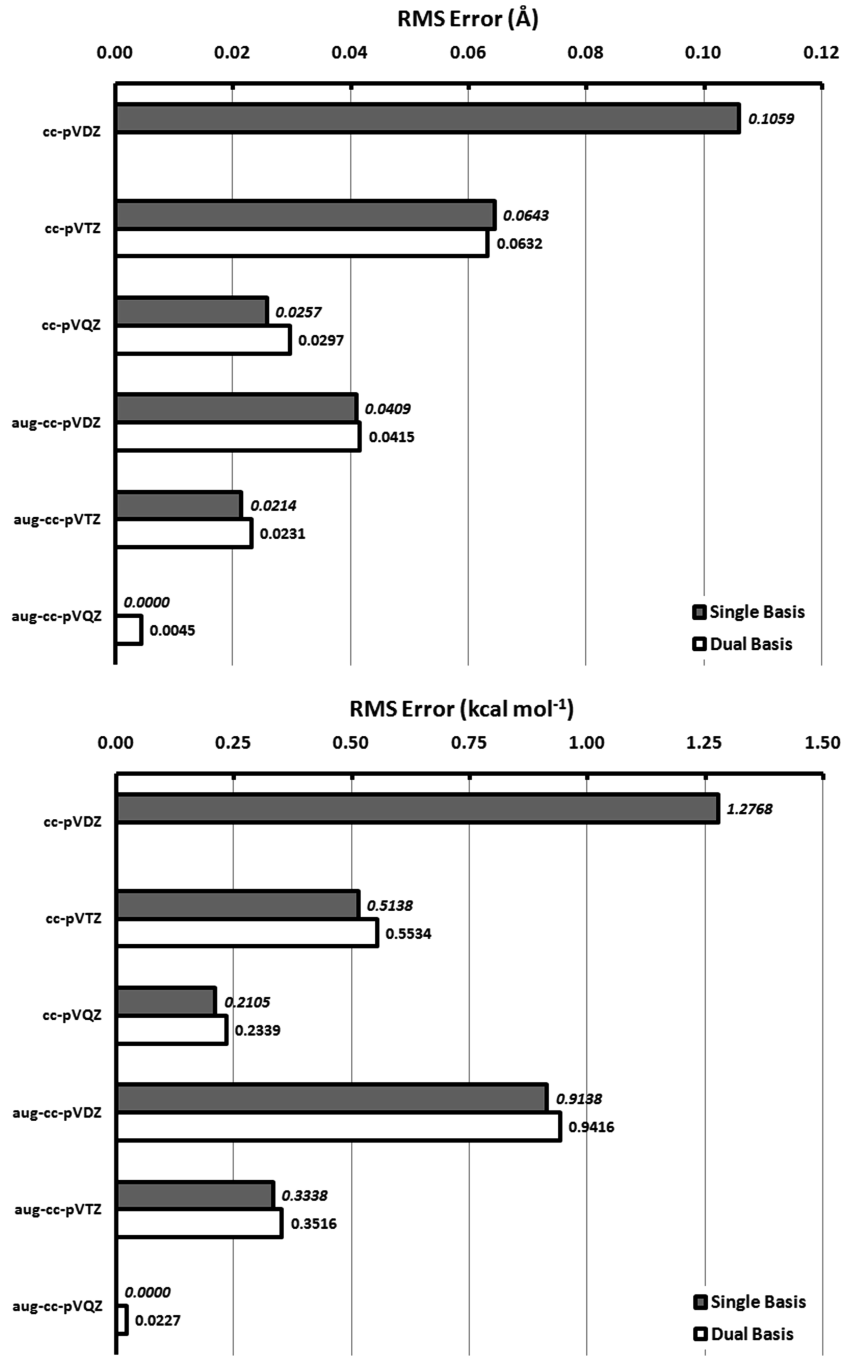

Figure 2. Statistical summary of DB-RI-MP2 performance on an 11-molecule subset of the S22 set. Shown are rms errors in optimized center-to-center distances (top), as well as errors in non-counterpoise-corrected interaction energies (bottom) on these optimized structures. Both panels utilize MP2/aug-cc-pVQZ results as reference numbers and present single- and dual-basis errors.

benzene dimer shows that total nuclear force savings are indeed 50%. Interaction energy errors using aug-cc-pVDZ are typically significant (though much improved over nondiffuse basis sets, such as cc-pVDZ or 6-31G*), relative to the basis set limit, but structures are often accurate and may be the only currently viable option for biologically relevant molecules.

Unlike the truncation for aug-cc-pVDZ, the estimated timings for aug-cc-pVTZ SCF nuclear gradients show significant promise. At the DB-SCF level, the basis truncation ratio of 0.536 leads to a 65% savings in a total nuclear force calculation. Thus, dual-basis aug-cc-pVTZ geometry optimizations may be performed three times faster than their single-basis counterparts. Further truncations may be feasible for geometry optimizations, but here we choose to use energies as the benchmark of interest. Additionally, savings in the DB-SCF gradient are somewhat tempered by the need to solve a response (z-vector)[95] equation. For the DB-RI-MP2 gradient, the response equation is already present in single-basis calculations, and the dual-basis savings should be more significant. With benzene dimer, for example, DB-RI-MP2 nuclear

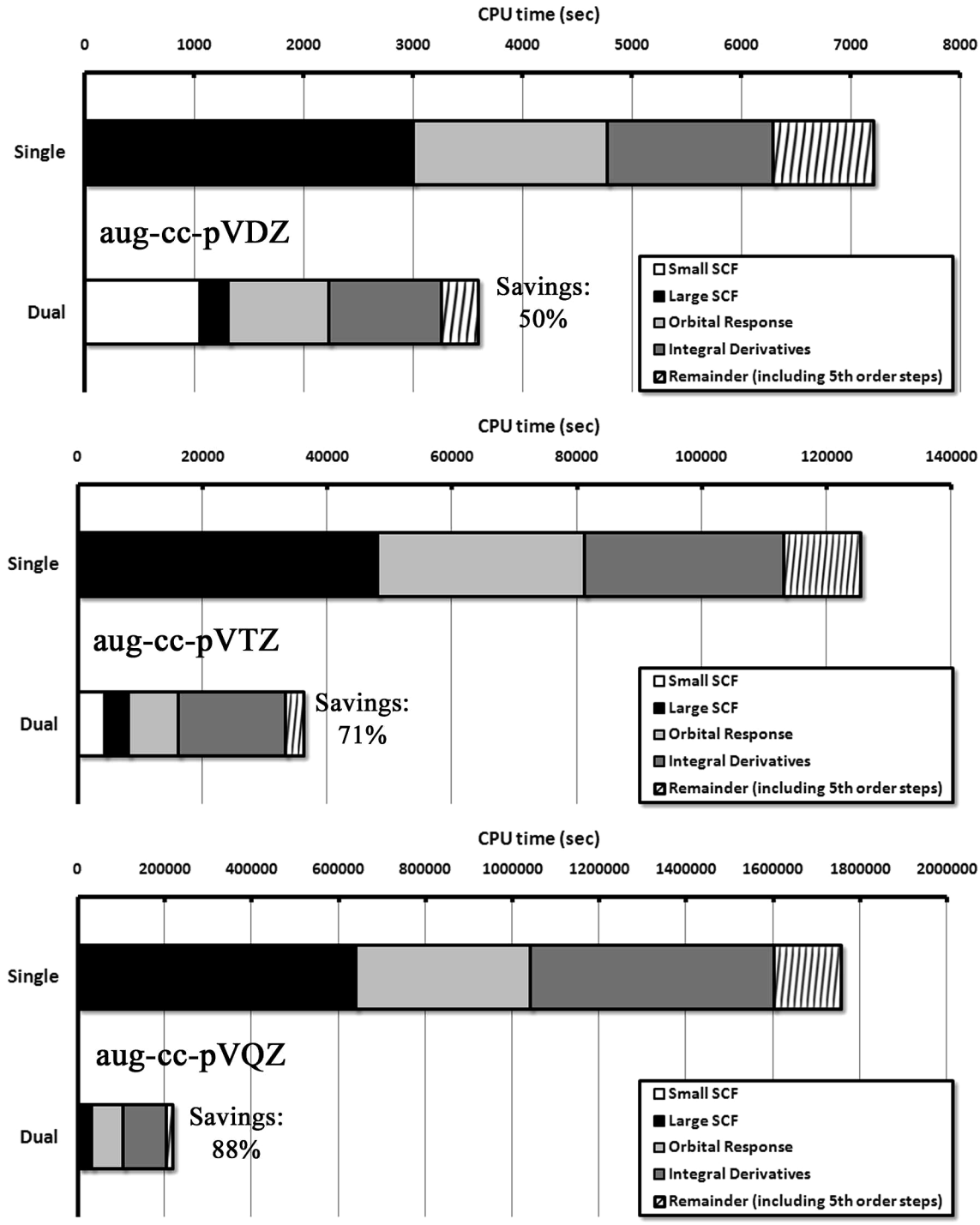

Figure 3. Nuclear force calculation timings for the benzene dimer. Shown are constituent contributions to the nuclear force for single- and dual-basis aug-cc-pV(D,T,Q)Z. Note the vastly different scales.

force timings show a 71% savings. In fact, the total DB-RI-MP2 nuclear force calculation was 25% faster than the SCF alone in the comparative single-basis calculation!

For aug-cc-pVQZ, an estimated 93% reduction in SCF cost is anticipated. Further extension to DB-RI-MP2 nuclear gradients, still admittedly prohibitive for many systems, exhibit noteworthy savings. For the benzene dimer, an 88% reduction in nuclear force timings is observed, reducing a *single* nuclear force calculation from nearly two weeks to two days.

4. Discussion

The results of this brief analysis indicate that DB pairings for augmented Dunning-style basis sets are quite viable. Accuracy closely approaches the results of target basis quantities, at savings of 60−93%. Further analysis indicates that basis set superposition error,[96] an unfortunate artifact of finite-basis calculations, particularly in noncovalent interactions, is slightly larger in DB calculations, but counterpoise-corrected results are just as accurate (see Supporting Information). Performance on noncovalent com-

plexes is superior to performance for electron affinities (in terms of absolute error), and this latter area does expose one limitation of DB methods, within this choice of basis set pairings. However, since anionic and open-shell complexes are, admittedly, a difficult class of systems for MP2 anyway,[97] these errors are tolerable.

Given the above results, the aug-cc-pV5Z truncation of Wolinski and Pulay[63] appears reasonable. The original 7s6p5d4f3g2h basis is truncated to 7s6p4d3f in their best pairing. Based upon our findings for aug-cc-pVQZ, this pairing most likely could be truncated further, either by elimination of all but the central *f* function or by elimination of the *f* functions altogether. Without empirical testing on our part, however, we recommend their original truncation for this extremely large basis set regime.

The SCS(MI) results are encouraging, in that the spin-component scaling is transferrable to DB pairings. A critical assessment of the basis set dependence in the S22 results, however, must include the observations that (1) The CP-corrected *aug* single-basis errors are essentially basis set independent, and (2) the dual- and single-basis results are independent of basis set after spin-component scaling. Both raise the question of why to use these larger basis sets at all. Several justifications exist. First, non-CP results are significantly more basis set dependent, as was shown in Figure 2. The CP correction is only viable for well-defined monomeric units. Intramolecular interactions in biological systems would necessarily suffer the much larger basis set dependence seen in the non-CP column of Table 1. Second, the better performance of medium-sized basis sets is solely caused by error cancellation, where the overestimate of dispersion interactions inherent to MP2 is canceled by basis set incompleteness effects. This error cancellation will not necessarily hold across the entire potential energy surface, where dispersion contributions are variable. And finally, noncovalent systems are often coupled with chemical systems, for example, enzymatic reactions, where larger basis sets more strongly influence thermochemistry. A balanced description of both properties is essential in this class of systems, and large basis sets are still required.

5. Conclusions

An accurate description of noncovalent interactions requires the use of correlated-electron methodologies and large basis sets. This pair of requirements, however, comes with a steep computational cost. The above work has shown that dual-basis MP2 is a viable alternative to single-basis calculations for noncovalent interactions. Optimal basis set pairings for the aug-cc-pV(D,T,Q)Z series were constructed, and, despite relatively aggressive basis truncations, errors for the S22 set are on the order of hundredths of a kcal mol^{-1}, a necessary requirement for a class of systems displaying small energy differences. Because of the ability to truncate the overly complete diffuse space of the augmented Dunning basis sets, computational time savings are significant for all three basis sets considered, and savings grow with the size of the basis. Since correlated-electron calculations converge slowly with respect to basis set size, these results allow for the accurate calculation of nearly complete basis set properties at significantly reduced cost.

While our assessment of both accuracy and computational savings are performed for dual-basis RI-MP2 calculations, we expect that these augmented dual basis sets will be useful in a variety of other electronic structure methods, as well. For instance, if calculations of intermolecular interactions are performed using a density-functional method[17−24,54,55,98,99] or post-Hartree−Fock method[100−103] that includes dispersion corrections, then large augmented basis sets are also important. These DB methods will provide speedups that will be greater than those reported here. Likewise, our conclusions regarding computational savings in MP2 calculations directly transfer to other methods which make second-order perturbative corrections to DFT energies, such as the increasingly popular "double-hybrid" functionals of Grimme.[54,54−59] Calculations using double-hybrid methods converge similarly with basis set size to MP2 itself. Finally, a natural synergism exists between DB methods, which correct SCF energies, and the emerging R12/F12 methods,[73] which provide a dual, two-particle basis for describing electron correlation effects.

Acknowledgment. Funding for this work has been provided by the National Science Foundation under Grant No. CHE-0535710, with additional support for code development from a subcontract from Q-Chem, Inc. for an NIH Small Business Innovation Research (SBIR) grant. M.H.-G. is a part owner of Q-Chem, Inc.

Supporting Information Available: Tested basis set pairings, detailed S22 results, BSSE analysis, detailed electron affinities data, basis set files in Q-Chem format. This information is available free of charge via the Internet at http://pubs.acs.org/.

Appendix A: Linear Dependence

In a dual-basis calculation, a projection of the occupied molecular orbital coefficients is required. The projection from one atomic orbital basis set to its corresponding superset is straightforward, but the presence of linear dependence in one or both basis sets requires attention. This Appendix describes these linear dependence possibilities and their associated solutions. The following methods have been implemented in Q-Chem.[94]

For reference, the following notational conventions are used:

$p, q, r, \ldots$ = all molecular orbitals (MOs)
$i, j, k, \ldots$ = occupied MOs
$a, b, c, \ldots$ = virtual MOs
$u, v, \lambda, \ldots$ = atomic orbitals (AOs)

Unadorned indices represent small-basis quantities. Indices bearing a tilde ($\tilde{\mu}$) represent large-basis-only quantities ("new functions"), while barred indices ($\bar{\mu}$) represent the full large-basis space; that is, $\{\mu\}\oplus\{\tilde{\mu}\} = \{\bar{\mu}\}$. The nonorthogonal AO space may be transformed to an orthogonalized AO (OAO) space by

$$|\underline{\mu}\rangle = \sum_{\nu} |\nu\rangle X_{\nu\underline{\mu}}$$

where underlined indices signify OAO functions. The nonunitary transformation matrix X may be chosen as any matrix that leaves the basis orthogonal; however, two forms are most commonly used.[104] Symmetric orthogonalization employs

$$X = S^{-1/2}$$

where the AO overlap matrix is defined as $S_{\mu\nu} = \langle\mu|\nu\rangle$. Canonical orthogonalization utilizes

$$X = Us^{-1/2}$$

where the unitary U matrix diagonalizes S, giving $USU^T = s$.

If the chosen basis set is linearly dependent, symmetric orthogonalization is not possible, since the S matrix becomes singular and its inverse ill-defined. Canonical orthogonalization is typically employed in this case, where the columns of U are scaled by $s^{-1/2}$. If any eigenvalue is below a chosen numerical threshold, the column is instead eliminated, thus reducing the number of OAOs by the number of linear dependencies and removing the null space.

An MO $|p\rangle$ may, therefore, be expanded in the small-basis AO space as

$$|p\rangle = \sum_{\mu\nu}^{\mathrm{AO}} \sum_{\underline{\lambda}}^{\mathrm{OAO}} |\mu\rangle X_{\mu\underline{\lambda}} X_{\underline{\lambda}\nu} \langle\nu|p\rangle \tag{3}$$

$$= \sum_{\mu} C_{\mu p} |\mu\rangle \tag{4}$$

Starting with this same MO, the large-basis analogue may be constructed via projection

$$|\bar{p}\rangle = \sum_{\bar{\lambda}\bar{\sigma}} \sum_{\mu\nu} \sum_{\underline{\gamma}} \sum_{\underline{\delta}} |\bar{\lambda}\rangle X'_{\bar{\lambda}\,\underline{\delta}} X'_{\underline{\delta}\,\bar{\sigma}} \langle\bar{\sigma}|\mu\rangle X_{\mu\underline{\gamma}} X_{\underline{\gamma}\nu} \langle\nu|p\rangle \tag{5}$$

$$= \sum_{\bar{\lambda}\bar{\sigma}} \sum_{\mu} \sum_{\underline{\delta}} |\bar{\lambda}\rangle X_{\bar{\lambda}\,\underline{\delta}} X_{\underline{\delta}\bar{\sigma}} S_{\bar{\sigma}\mu} C_{\mu p} \tag{6}$$

where $S_{\bar{\sigma}\mu}$ is the (rectangular) overlap matrix in the mixed large-small space. (The primed X's are simply a reminder that different orthogonalization schemes could potentially be used for the two basis sets.) An occupied MO coefficient in the large space is thus constructed as

$$C_{\bar{\lambda}i} = \sum_{\bar{\sigma}} \sum_{\mu} \sum_{\underline{\delta}} X_{\bar{\lambda}\,\underline{\delta}} X_{\underline{\delta}\,\bar{\sigma}} S_{\bar{\sigma}\mu} C_{\mu i} \tag{7}$$

In the absence of linear dependence, several simplifications occur. Equation 3 becomes

$$|p\rangle = \sum_{\mu\nu}^{AO} |\mu\rangle S^{-1}_{\mu\nu} \langle\nu|p\rangle \tag{8}$$

and a projected large-basis occupied MO coefficient reduces to

$$C_{\bar{\lambda}i} = \sum_{\bar{\sigma}} \sum_{\mu} S^{-1}_{\bar{\lambda}\bar{\sigma}} S_{\bar{\sigma}\mu} C_{\mu i} \tag{9}$$

This latter scheme properly handles a projection between any two linearly independent basis sets (activated by the BASISPROJTYPE=OVPROJECTION keywords in Q-Chem). Importantly, when $\{\mu\} \subset \{\bar{\mu}\}$, the only case considered for our basis set truncations (for computational efficiency reasons), the rectangular projection matrix

$$T_{\bar{\lambda}\mu} = \sum_{\bar{\sigma}} S^{-1}_{\bar{\lambda}\bar{\sigma}} S_{\bar{\sigma}\mu} \tag{10}$$

simply becomes a rectangular delta function matrix

$$T_{\bar{\lambda}\mu} = \delta_{\bar{\lambda}\mu} \tag{11}$$

The large-space occupied coefficients are, therefore, identical to the small-space occupied coefficients, with additional null elements corresponding to new basis functions. The same is true for the large-basis density matrix, defined as

$$P_{\bar{\mu}\bar{\nu}} = \sum_{i} C_{\bar{\mu}i} C_{i\bar{\nu}}$$

The small-basis elements of P are identical to elements in its small-basis analogue, p, with all other elements equal to zero.

When two or more basis functions cause the basis to become (numerically) linearly dependent, however, this scheme requires revision. Standard routines to eliminate linear dependence, such as canonical orthogonalization (or equivalently, a "square" singular value decomposition),[105] take linear combinations of the offending basis functions to produce an orthonormal basis of reduced dimension. Several issues make this choice tenuous for a dual-basis projection.

First, while canonical orthogonalization in eq 7 is qualitatively and numerically correct, it, by definition, mixes the contribution of each AO to the large-basis MO coefficients. Thus, the zero-structure in the resultant coefficients vanishes, leaving the integral screening in the large basis ineffective. The addition of an exactly linearly dependent function in the large basis, for example, would produce an occupied MO coefficient with two identical elements, both equal to half of the corresponding element in the small basis.

The second (related) issue involves the large-basis density matrix. When the aforementioned mixing occurs, the small-basis elements of P also change. Thus, the reference energy before projection, $E[p]$, is no longer equal to the energy after projection, $E[P]$, and ambiguity arises in the definition of the dual-basis energy.[62]

This X-dependent ambiguity also arises in a standard SCF calculation and can even leave the SCF energy effectively nonvariational. Consider, for example, a calculation of the energy of a hydrogen atom, for which a single s function's exponent has been variationally optimized. If a second s function causes linear dependence, the resultant "mixed" s function will necessarily cause the energy to rise. While this latter property is rare, the fact remains that the SCF energy is dependent on X for linearly dependent basis sets. The same is true for a dual-basis calculation, with the added possibility that the *same* calculation may employ different X.

We thus desire a projection matrix T that accomplishes the following:

(1) The resultant large-basis occupied MO coefficients are orthonormal.
(2) The resultant density matrix remains idempotent.
(3) The corresponding large-basis density matrix retains its designed "zero structure".

(4) The small-basis energy is unchanged by the projection ($E[p] = E[P]$) and is unambiguously defined.

The ansatz for T that we have chosen is simply the delta function projector shown in eq 11, applied whether or not linear dependencies exist. The large-basis occupied MO coefficients are a strict superset of the small-basis coefficients and are used to construct the large basis density matrix P. Canonical orthogonalization is subsequently used for later stages of the dual-basis SCF, such as the transformation to an OAO basis prior to diagonalization of the large-basis Fock matrix, $F[P]$.

Such a choice satisfies the four criteria above, by definition. Additionally, the special cases unique to dual-basis calculations are properly handled. For linear dependence among functions residing in the small basis only, the delta function projector preserves the mixing already used to account for these dependencies in the small-basis SCF, while producing properly orthonormal occupied MO coefficients. The reference energy in a dual-basis calculation is thus unambiguously defined. For linear dependence among large-basis functions, canonical orthogonalization performs adequately ($X^2S = 1$) but only for *exact* linear dependence. The delta function projector produces null large-basis elements of the occupied MO coefficients for this case, even for the more likely event of purely numerical linear dependence. Finally, for the more complicated case of a new, large-basis function becoming linearly dependent with a small-basis function, we have some choice in the projection matrix. Items 1 and 2 above impose two standard constraints on our projection, whereas items 3 and 4 impose constraints specific to dual-basis calculations, uniquely satisfied by the delta function projector. This choice of projection also avoids unnecessary complications in the analytic gradient of the dual-basis energy.

While the reference energy is uniquely defined by T, the dual-basis energy correction is still dependent on the choice of X. For the H-atom "thought experiment" above, the dual-basis correction will be nonzero (and dependent on the X used in the orthogonalized Roothaan equations, for example). However, this situation is identical to a standard SCF calculation. Differences among orthogonalization routines will be minimal, and the widespread use of canonical orthogonalization should further minimize these small discrepancies.

This delta function projector is easily implemented. Since the AO ordering is not {small + large} but instead is atom-ordered, a loop over elements of $S_{\mu\bar{\nu}}$ is required to distinguish between small and large basis functions. A value of 1.0 (to within the job-specified precision) in a column of this rectangular overlap matrix designates the column-index basis function as a small-basis function. The remaining functions are large-basis functions and correspond to null columns in the projection matrix. The same projection scheme is used for projections within the dual-basis analytic gradient.

One final caveat remains. The delta function projector is the method of choice for subset constructions, as defined in this paper. For 6-31G* ← 6-4G calculations,[67] in which the small basis is a subset by *primitives* only, the original projection scheme must be utilized. In practice, this fact is not a limitation, as linear dependencies most likely will never be problematic for these small basis sets. A case-dependent switch in the projection code allows for either type to be handled.

References

(1) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *J. Chem. Phys.* **1997**, *106*, 9639.

(2) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133–A1138.

(3) Becke, A. D. *J. Chem. Phys.* **1992**, *96*, 2155–2160.

(4) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(5) Andzelm, J.; Wimmer, E. *J. Chem. Phys.* **1992**, *96*, 1280–1303.

(6) Johnson, B. G.; Gill, P. M. W.; Pople, J. A. *J. Chem. Phys.* **1993**, *98*, 5612–5626.

(7) Becke, A. D. *J. Chem. Phys.* **1986**, *84*, 4524–4529.

(8) Delley, B. *J. Chem. Phys.* **1991**, *94*, 7245–7250.

(9) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–B871.

(10) Kristyàn, S.; Pulay, P. *Chem. Phys. Lett.* **1994**, *229*, 175–180.

(11) Langreth, D. C.; Perdew, J. P. *Phys. Rev. B* **1980**, *21*, 5469–5493.

(12) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1986**, *33*, 8800–8802.

(13) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822–8824.

(14) Perdew, J. P. *Phys. Rev. B* **1988**, *34*, 7406.

(15) Langreth, D. C.; Mehl, M. J. *Phys. Rev. B* **1983**, *28*, 1809–1834.

(16) Langreth, D. C.; Mehl, M. J. *Phys. Rev. B* **1983**, *29*, 2310.

(17) Wu, Q.; Yang, W. *J. Chem. Phys.* **2002**, *116*, 515–524.

(18) Jureka, P.; erný, J.; Hobza, P.; Salahub, D. R. *J. Comput. Chem.* **2006**, *28*, 555–569.

(19) Antony, J.; Grimme, S. *Phys. Chem. Chem. Phys.* **2006**, *8*, 5287–5293.

(20) Chai, J.-D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.

(21) Thonhauser, T.; Puzder, A.; Langreth, D. C. *J. Chem. Phys.* **2006**, *124*, 164106.

(22) Thonhauser, T.; Cooper, V. R. S; Li, A. P.; Hyldgaard, P.; Langreth, D. C. *Phys. Rev. Lett.* **2007**, *76*, 125112.

(23) Puzder, A.; Dion, M.; Langreth, D. C. *J. Chem. Phys.* **2006**, *124*, 164105.

(24) Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D. C.; Lundqvist, B. I. *Phys. Rev. Lett.* **2004**, *92*, 246401.

(25) Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887–1930.

(26) Podeszwa, R.; Bukowski, R.; Szalewicz, K. *J. Phys. Chem. A* **2006**, *110*, 10345–10354.

(27) Misquitta, A. J.; Szalewicz, K. *Chem. Phys. Lett.* **2002**, *357*, 301–306.

(28) Misquitta, A. J.; Jeziorski, B.; Szalewicz, K. *Phys. Rev. Lett.* **2003**, *91*, 033201–033204.

(29) Heβelmann, A.; Jansen, G. *Chem. Phys. Lett.* **2002**, *357*, 464–470.

(30) Heβelmann, A.; Jansen, G. *Chem. Phys. Lett.* **2002**, *362*, 319–325.

(31) Heβelmann, A.; Jansen, G. *Chem. Phys. Lett.* **2003**, *367*, 778–784.

(32) Williams, H. L.; Chabalowski, C. F. *J. Phys. Chem. A* **2001**, *105*, 646–659.

(33) Janowski, T.; Pulay, P. *Chem. Phys. Lett.* **2007**, *447*, 27–32.

(34) DiStasio, R. A.; von Helden, G.; Steele, R. P.; Head-Gordon, M. *Chem. Phys. Lett.* **2007**, *437*, 277–283.

(35) Hill, J. G.; Platts, J. A.; Werner, H. J. *Phys. Chem. Chem. Phys.* **2006**, *8*, 4072–4078.

(36) Park, Y. C.; Lee, J. S. *J. Phys. Chem. A* **2006**, *110*, 5091–5095.

(37) Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2004**, *108*, 10200–10207.

(38) Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. *J. Am. Chem. Soc.* **2002**, *124*, 10887–10893.

(39) Takatani, T.; Sherrill, C. D. *Phys. Chem. Chem. Phys.* **2007**, *9*, 6106–6114.

(40) DiStasio, R. A., Jr.; Steele, R. P.; Rhee, Y. M.; Shao, Y.; Head-Gordon, M. *J. Comput. Chem.* **2007**, *28*, 839.

(41) Møller, C.; Plesset, M. *Phys. Rev.* **1934**, *46*, 618.

(42) Eichkorn, K.; Treutler, O.; Ohm, H.; Heiser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *240*, 283.

(43) Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. *Chem. Phys. Lett.* **1998**, *294*, 143.

(44) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. *Chem. Phys. Lett.* **1993**, *208*, 359.

(45) Jung, Y.; Sodt, A.; Gill, P. M. W.; Head-Gordon, M. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6692.

(46) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095–9102.

(47) Antony, J.; Grimme, S. *J. Phys. Chem. A* **2007**, *111*, 4862–4868.

(48) Lochan, R. C.; Shao, Y.; Head-Gordon, M. *J. Chem. Theory Comput.* **2007**, *3*, 988–1003.

(49) Jung, Y.; Shao, Y.; Head-Gordon, M. *J. Comput. Chem.* **2007**, *28*, 1953–1964.

(50) Lochan, R. C.; Head-Gordon, M. *J. Chem. Phys.* **2007**, *126*, 164101.

(51) Rhee, Y. M.; Head-Gordon, M. *J. Phys. Chem. A* **2007**, *111*, 5314–5326.

(52) DiStasio, R. A., Jr.; Head-Gordon, M. *Mol. Phys.* **2007**, *105*, 1073–1083.

(53) Hill, J. G.; Platts, J. A. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2785–2791.

(54) Grimme, S. *J. Chem. Phys.* **2006**, *124*, 034108.

(55) Schwabe, T.; Grimme, S. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3397–3406.

(56) Schwabe, T.; Grimme, S. *Phys. Chem. Chem. Phys.* **2006**, *8*, 4398–4401.

(57) Grimme, S.; Mück-Lichtenfeld, C.; Würthwein, E.-U.; Ehlers, A. W.; Goumans, T. P. M.; Lammertsma, K. *J. Phys. Chem. A* **2006**, *110*, 2583–2586.

(58) Grimme, S.; Steinmetz, M.; Korth, M. *J. Org. Chem.* **2007**, *72*, 2118–2126.

(59) Neese, F.; Schwabe, T.; Grimme, S. *J. Chem. Phys.* **2007**, *126*, 124115.

(60) Benighaus, T.; DiStasio, R. A., Jr.; Lochan, R. C.; Chai, J. D.; Head-Gordon, M. *J. Phys. Chem. A* **2008**, *112*, 2702–2712.

(61) Liang, W.-Z.; Head-Gordon, M. *J. Phys. Chem. A* **2004**, *108*, 3206–3210.

(62) Steele, R. P.; DiStasio, R. A., Jr.; Shao, Y.; Kong, J.; Head-Gordon, M. *J. Chem. Phys.* **2006**, *125*, 074108.

(63) Wolinski, K.; Pulay, P. *J. Chem. Phys.* **2003**, *118*, 9497–9503.

(64) Steele, R. P.; Shao, Y.; DiStasio, R. A.; Head-Gordon, M. *J. Phys. Chem. A* **2006**, *110*, 13915–13922.

(65) DiStasio, R. A., Jr.; Steele, R. P.; Head-Gordon, M. *Mol. Phys.* **2007**, *105*, 2731–2742.

(66) Nakajima, T.; Hirao, K. *J. Chem. Phys.* **2006**, *124*, 184108.

(67) Steele, R. P.; Head-Gordon, M. *Mol. Phys.* **2007**, *105*, 2455–2473.

(68) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(69) Kendall, R. A.; Dunning, T. H., Jr. *Chem. Phys. Lett.* **1992**, *96*, 6796.

(70) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, *98*, 1358.

(71) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243.

(72) Feller, D. *J. Chem. Phys.* **1992**, *96*, 6104–6114.

(73) Armour, E. A. G.; Franz, J.; Tennyson, J. *Explicitly Correlated Wavefunctions*, 2nd ed.; Collaborative Computational Project on Molecular Quantum Dynamics, Daresbury Laboratory, Daresbury, Warrington, U.K., 2006.

(74) Manby, F. R. *J. Chem. Phys.* **2003**, *119*, 4607.

(75) May, A. J.; Manby, F. R. *J. Chem. Phys.* **2004**, *121*, 4479.

(76) Werner, H.-J.; Manby, F. R. *J. Chem. Phys.* **2006**, *124*, 054114.

(77) Manby, F. R.; Werner, H.-J.; Adler, T. B.; May, A. J. *J. Chem. Phys.* **2006**, *124*, 094103.

(78) Klopper, W.; Samson, C. C. M. *J. Chem. Phys.* **2002**, *116*, 6397.

(79) Kutzelnigg, W.; Klopper, W. *J. Chem. Phys.* **1991**, *94*, 1985.

(80) Valeev, E. F. *Chem. Phys. Lett.* **2004**, *395*, 190.

(81) Jureka, P.; Sponer, J.; Cerny, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985.

(82) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* **1991**, *94*, 7221–7230.

(83) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764–7776.

(84) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **2000**, *112*, 7374–7383.

(85) Pulay, P. *Chem. Phys. Lett.* **1980**, *73*, 393.

(86) Pulay, P. *J. Comput. Chem.* **1982**, *3*, 556.

(87) Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175–3183.

(88) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257–2261.

(89) Hariharan, P. C.; Pople, J. A. *Theor. Chem. Acc.* **1973**, *28*, 213.

(90) Francl, M. M.; Petro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. *J. Chem. Phys.* **1982**, *77*, 3654.

(91) Mintz, B.; Lennox, K. P.; Wilson, A. K. *J. Chem. Phys.* **2004**, *121*, 5629–5634.

(92) Clark, T.; Chandrasekhar, J.; Spitnagel, G. W.; v. R. Schleyer, P. *J. Comput. Chem.* **1983**, *4*, 294.

(93) Pitoák, M.; Riley, K. E.; Neogrády, P.; Hobza, P. *Chem. Phys. Chem.* **2008**, *9*, 1636–1644.

(94) Shao, Y.; et al. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172.

(95) Handy, N. C.; Schaefer, H. F., III *J. Chem. Phys.* **1984**, *81*, 5031.

(96) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.

(97) Byrd, E.; Sherrill, C.; Head-Gordon, M. *J. Phys. Chem. A* **2001**, *105*, 9736–9747.

(98) Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2005**, *123*, 154101.

(99) Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2007**, *127*, 124108.

(100) Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2005**, *122*, 154104.

(101) Johnson, E. R.; Becke, A. D. *J. Chem. Phys.* **2005**, *123*, 024101.

(102) Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2006**, *124*, 014104.

(103) Johnson, E. R.; Becke, A. D. *J. Chem. Phys.* **2006**, *124*, 174104.

(104) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications, Inc.: Mineola, NY, 1982; pp 142−145.

(105) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in C++: The Art of Scientific Computing*; Oxford University Press: Oxford, U.K., 1994; pp 62−68.

CT900058P