Property-optimized Gaussian basis sets for molecular response calculations

Dmitrij Rappoport^{a)} and Filipp Furche^{b)}

Department of Chemistry, University of California, Irvine, 1102 Natural Sciences II, Irvine, California 92697, USA

(Received 29 June 2010; accepted 9 August 2010; published online 4 October 2010)

With recent advances in electronic structure methods, first-principles calculations of electronic response properties, such as linear and nonlinear polarizabilities, have become possible for molecules with more than 100 atoms. Basis set incompleteness is typically the main source of error in such calculations since traditional diffuse augmented basis sets are too costly to use or suffer from near linear dependence. To address this problem, we construct the first comprehensive set of property-optimized augmented basis sets for elements H-Rn except lanthanides. The new basis sets build on the Karlsruhe segmented contracted basis sets of split-valence to quadruple-zeta valence quality and add a small number of moderately diffuse basis functions. The exponents are determined variationally by maximization of atomic Hartree–Fock polarizabilities using analytical derivative methods. The performance of the resulting basis sets is assessed using a set of 313 molecular static Hartree–Fock polarizabilities. The mean absolute basis set errors are 3.6%, 1.1%, and 0.3% for property-optimized basis sets of split-valence, triple-zeta, and quadruple-zeta valence quality, respectively. Density functional and second-order Møller-Plesset polarizabilities show similar basis set convergence. We demonstrate the efficiency of our basis sets by computing static polarizabilities of icosahedral fullerenes up to C720 using hybrid density functional theory. © 2010 American Institute of Physics. [doi:10.1063/1.3484283]

I. INTRODUCTION

The prediction of molecular response properties from first principles^{1–3} has become an integral part in the search for novel materials for applications such as molecular electronics, light harvesting, and nonlinear optics. Accurate calculations can identify new lead structures and predict substituent effects in a systematic and cost-efficient way. Time-dependent density functional theory (TDDFT) (Ref. 4) has emerged as a good compromise between accuracy and computational cost and is widely used for computing linear and nonlinear response properties of large molecules. In addition, coupled-cluster response methods^{5,6} are available for accurate calculations and as benchmarks for density functional methods.

Perhaps surprisingly, incompleteness of the one-electron basis set is a major cause of error in molecular response calculations, particularly for systems with more than 20–50 atoms. In the vast majority of molecular calculations, atomcentered basis functions (BFs) are used to represent molecular orbitals (MOs) and their response to external perturbations.⁷ It has long been established that the basis set requirements for an accurate description of the orbital response are fundamentally different from that of the unperturbed MOs.⁸ In particular, diffuse, i.e., low-exponent BFs are crucial for an adequate description of the orbital response. However, the addition of a significant number of diffuse BFs has an adverse effect on numerical stability of both ground-state and response calculations.

With existing diffuse augmentation schemes, the balance between improved response properties and numerical stability is delicate and often difficult to achieve. In the early 1980s, Schleyer and co-workers^{9,10} were the first to use energy minimization of anions to construct diffuse augmented basis sets on a larger scale. They were already aware of the computational drawbacks of diffuse augmentation and noted that their basis sets "may not be ideal"¹⁰ for properties such as excitation energies of neutral systems. Later, Dunning and co-workers constructed the augmented correlation-consistent basis sets by energy minimization of atomic or molecular anions^{11,12} or by extrapolation.¹³ Similarly, the augmented polarization-consistent basis sets of Jensen¹⁴⁻¹⁷ were obtained by extrapolation using the basis set convergence of molecular electron affinities as a guiding principle. These basis sets approach completeness in the diffuse region with increasing basis set size¹⁸ and are highly useful for smaller systems. However, they tend to include large numbers of diffuse BFs and rapidly become impractical for larger molecules. An alternative are the basis sets developed by Sadlej and co-workers, which are specially constructed for polarizability calculations.^{19,20} The Sadlei basis sets often vield excellent response properties with a comparatively small valence space, but their use for larger molecules can be affected by near linear dependence. Recently, Benkova, Sadlej, Oakes, and Bell published modified basis sets for first- and

^{a)}Present address: Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138. Electronic mail: rappoport@chemistry.harvard.edu.

^{b)}Electronic mail: filipp.furche@uci.edu.

second-row atoms containing a smaller number of diffuse functions. $^{21,22} \,$

The aim of the present work is to develop hierarchical basis sets that (i) reduce the basis set error in linear polarizabilities below typical errors of Hartree-Fock (HF) and density functional treatments, (ii) are applicable to systems with more than 100 atoms, and (iii) yield systematic convergence to the basis set limit. Already in 1976, Werner and Meyer pointed out²³ that the Hylleraas variational principle^{8,24} can be used to construct basis sets for molecular properties. This principle implies that optimal values of parameters, such as basis function exponents entering the first-order wave function, maximize the electronic polarizability. Here we use analytical derivative techniques based on Lagrangian polarizability theory²⁵ to implement Werner's and Meyer's strategy. We start from polarized Karlsruhe def2 basis sets,²⁶ describing the core and valence electrons. These segmented contracted basis sets developed by Ahlrichs and co-workers over the past two decades²⁷⁻²⁹ are well-known for their robustness and their excellent cost-to-performance ratio in large-scale HF and density functional theory (DFT) calculations. We augment the Karlsruhe def2 basis sets with a small number of diffuse BFs determined by maximization of isotropic static HF polarizabilities of atoms. The variational principle guarantees that the chosen augmentation is optimal and leads to compact basis sets with few diffuse functions. We report on the first comprehensive set of property-optimized diffuse augmented basis sets across the Periodic Table, which sheds light on the general requirements for an efficient representation of molecular response in Gaussian basis sets and may also be used for constructing other diffuse basis set hierarchies.

This paper is structured as follows. We show how to use the polarizability variational principle for basis set optimization in Sec. II. Section III summarizes the implementation of analytical basis set gradients of the HF polarizability. The basis set convergence as a function of the number of primitives and their *l* quantum number is investigated in detail for fluorine atom in Sec. IV. This leads to the design of augmented basis sets of split-valence, triple-zeta, and quadruplezeta valence quality for elements H–Rn (except lanthanides) presented in Sec. V. The quality of the property-optimized basis sets is assessed using atomic polarizabilities and a large molecular test set in Sec. VI, followed by recommendations. Some applications using density functional and second-order Møller–Plesset (MP2) theory are presented in Sec. VII.

II. VARIATIONAL OPTIMIZATION OF BASIS SETS FOR RESPONSE PROPERTIES

Most quantum chemistry codes employ contracted Cartesian Gaussian-type orbitals (CGTOs) as BFs,

$$\chi_{\mu}(\mathbf{r}|\mathbf{R},\mathbf{l}) = (x-X)^{l_{x}}(y-Y)^{l_{y}}(z-Z)^{l_{z}}\sum_{d}c_{\mu d}e^{-\zeta_{\mu d}(\mathbf{r}-\mathbf{R})^{2}},$$
(1)

where $\mathbf{R} = (X, Y, Z)$ denotes the center of the BF and $\mathbf{l} = (l_x, l_y, l_z)$ is the *l*-quantum number vector. In analogy to hydrogenic orbitals, CGTOs with $l = l_x + l_y + l_z = 0, 1, 2, 3, ...$

are denoted by s, p, d, f, \ldots orbitals, respectively. CGTOs are usually grouped in shells containing all BFs of a given angular-momentum quantum number l. Optimization of the basis set parameters $\zeta_{\mu d}$ and $c_{\mu d}$ requires a variational principle for some target functional, which ensures convergence toward the complete basis set limit. The deviation of the target functional from the basis set limit is a quality measure for the basis set at hand, and quasi-Newton optimization algorithms³⁰ can be used to minimize the basis set error. Basis set optimizations for ground-state wave functions are based on the variational principle for the ground-state energy and are often performed in the framework of HF approximation (see, e.g., Refs. 26-29). Although most molecular properties can be obtained from the response of the total energy to a suitably chosen perturbation, basis sets are usually optimized in the absence of such external perturbations. Thus, energy-optimized basis sets perform poorly for many response properties.

There are two systematic solutions to this dilemma. First, explicitly perturbation dependent basis sets can be used. The most successful example of this strategy are gauge-including atomic orbital basis sets,³¹ which explicitly depend on a static magnetic field. For electric perturbations, a balanced description of first- and higher-order response properties is difficult to achieve,³² and electric-field-variant basis sets^{33–35} are rarely used.

The second strategy is to augment energy-optimized basis sets by a few additional functions to improve the description of response properties. This method is conceptually simple and efficient if the additional functions can be determined by a variational procedure guaranteeing quadratic convergence of the desired properties. Such a procedure starts from Hylleraas' variational principle.^{8,24} The functional

$$G[\psi^{(1)}] = \frac{1}{2} \langle \psi^{(1)} | \hat{H}^{(0)} - E^{(0)} | \psi^{(1)} \rangle - \langle \psi^{(1)} | \hat{H}^{(1)} | \psi^{(0)} \rangle$$
(2)

is minimized by the exact first-order wave function $\psi^{(1)}$. $\hat{H}^{(0)}$ and $\hat{H}^{(1)}$ denote the zero- and first-order Hamiltonians, and $E^{(0)}$ and $\psi^{(0)}$ denote the zero-order energy and wave function, respectively; the variation of *G* is restricted to $\psi^{(1)}$ orthogonal to $\psi^{(0)}$. For an electric dipole perturbation,

$$\hat{H}^{(1)} = -\hat{\boldsymbol{\mu}} \cdot \mathbf{E},\tag{3}$$

where **E** is a uniform static electric field, the minimum of *G* equals the negative polarizability. Thus, basis sets for electric response properties may be optimized by maximizing the molecular polarizability. As a result of the variational principle, the basis set error in the polarizability is quadratic in the basis set error in $\psi^{(1)}$. Thus, it should be possible to achieve high accuracy in electric response properties by adding a few polarization functions if their exponents are fully optimized. This method is analogous to the optimization of polarization functions by maximization of the MP2 energy.

G is not variational with respect to changes in the zeroorder wave function. Thus, a rigorous lower bound for *G* exists only if the basis set for $\psi^{(0)}$ is entirely fixed, which may be achieved, e.g., by dual basis set methods. In practice, keeping the valence basis set fixed and optimizing only the diffuse augmentation was found to be stable in all cases. However, for certain metal atoms, small valence basis sets can lead to spuriously large polarizabilities, as discussed in detail in Sec. VI A.

A crucial advantage of our approach is that basis sets for molecular property calculations are determined directly by optimization of response properties rather than by energy optimization of negative ions, which has been a common practice so far. The ground-state wave function of negative ions has a much slower exponential decay than the groundstate wave function of the corresponding neutral atom. Thus, basis set augmentations optimized using negative ions typically contain very diffuse functions. (Additional complications arise for atoms whose negative ions are unbound.) For molecules with more than a few atoms, such augmentations become impractical, causing near linear dependence and inefficiency of integral prescreening. This is undesirable, especially because highly diffuse basis functions may not even be necessary for accurate response property calculations. The first-order wave function $\psi^{(1)}$ generally has the same exponential decay as the zero-order wave function $\psi^{(0)}$. Thus, adding more and more diffuse functions will not improve $\psi^{(1)}$ significantly once the exponents drop below a certain threshold determined by $\psi^{(0)}$.

We illustrate this argument for the 1s ground state of hydrogen atom with zero-order wave function

$$\psi^{(0)}(\mathbf{r}) = \frac{1}{\sqrt{\pi}} e^{-r}.$$
(4)

It is well-known that for a dipole perturbation by a uniform electric field $\mathbf{E} = E\mathbf{e}_z$, the first-order wave function is

$$\psi^{(1)}(\mathbf{r}) = -zE\left(1+\frac{r}{2}\right)\psi^{(0)}(\mathbf{r}).$$
(5)

Obviously, $\psi^{(1)}$ has the same exponential decay as $\psi^{(0)}$, whereas augmentation optimized for negative ions or excited states will exhibit a spuriously slow exponential decay. Equation (5) also shows that basis sets for first-order electric dipole perturbed wave functions must span products of the occupied MOs and the monomials x, y, z, which is achieved by polarization functions with angular momentum ± 1 (+1 for *s* functions). A qualitative description of the radial factor (1+r/2) by nodeless Gaussians requires at least *two* primitive polarization functions with different exponents.

III. ANALYTICAL BASIS SET GRADIENTS OF ELECTRONIC POLARIZABILITIES

Analytical basis set derivatives are a crucial prerequisite for efficient basis set optimization. Analytical polarizability gradients are conveniently derived and implemented using polarizability Lagrangian theory.²⁵ First, a polarizability Lagrangian L is constructed, which is stable with respect to *all* parameters, including ground-state MO coefficients. L equals the negative polarizability at its minimum. After L has been minimized, analytical polarizability gradients are obtained from the Hellmann–Feynman theorem. The static HF polarizability basis set gradients considered here are a special case of the analytical polarizability derivative methods developed in Ref. 25, to which the reader is referred for details. The



FIG. 1. Effect of incremental augmentation of SVP, TZVPP, and QZVPP basis sets on the isotropic static HF polarizability of fluorine atom.

only difference of the present work from Ref. 25 is that in the present work derivatives are taken with respect to basis set exponents rather than nuclear coordinates. The basis set gradients of polarizabilities were implemented in the EGRAD module^{36,37} of the TURBOMOLE program package.^{38,39}

IV. CASE STUDY: FLUORINE ATOM

We begin by investigating the effect of diffuse basis functions on the static polarizability of fluorine atom in the spin-unrestricted HF (UHF) approximation. Sets of primitive diffuse functions were added to Karlsruhe def2 ground-state basis sets and fully optimized by maximizing the atomic polarizability. Thus, the maximum polarizability achievable by a given diffuse augmentation is obtained. This procedure allows us to assess the relative contributions of individual BF shells to polarizabilities and to select the BFs yielding the largest increase in polarizability. In addition, we can map out the range of Gaussian exponents important for first-order response.

The effect of a successive augmentation of the def2 basis sets for the fluorine atom is shown in Fig. 1. The splitvalence plus polarization (def2-SVP) basis set underestimates the basis set limit for the static polarizability by as much as 65%. The basis set limit is 3.30 a.u. in the UHF approximation, while the restricted open-shell HF method yields 3.28 a.u.⁴⁰ The largest improvement in polarizability is obtained by adding a BF shell of d symmetry (1d augmentation). The resulting basis set, denoted by SVP+1d, has 23% basis set error. Further 1s augmentation and reoptimization of all additional BF exponents (yielding the SVP +1s1d basis set) reduces the basis set error to 14%. These results are not surprising since the polarization of the fluorine atom is dominated by the deformation of the highest-lying occupied p orbitals (l=1), whose first-order electric dipole response has s and d symmetry. The case of the fluorine atom also illustrates the need to balance diffuse polarization and diffuse valence BFs. The largest improvement relative to the SVP+1s1d basis set for fluorine is achieved by an additional shell of diffuse p functions. After reoptimization of BF exponents (resulting in the SVP+1s1p1d basis set) the basis set error is reduced to less than 3%. The additional diffuse p BF



FIG. 2. Incremental contributions of d basis functions to the isotropic static HF polarizability of fluorine atom with SV, TZV, and QZV basis sets.

shell polarizes the 2s valence shell and improves the outer region of the occupied p orbitals, but contributes little to the ground-state energy.

A similar picture emerges for triple- and quadruple-zeta basis sets of fluorine (see Fig. 1). The unaugmented triplezeta valence basis set with two sets of polarization functions (def2-TZVPP) and the corresponding quadruple-zeta basis set (def2-QZVPP) are far from the basis set limit for the static polarizability, having 45% and 33% basis set error, respectively. This sharply contrasts with the relative error of def2-QZVPP for the fluorine HF ground-state energy, which is approximately 3 ppm.²⁶ The largest improvements of the atomic polarizability are again obtained by 1d augmentation, with TZVPP+1d and QZVPP+1d basis sets reducing the basis set errors to 9% and 8%, respectively. Augmented TZVPP+1s1p1d and QZVPP+1s1p1d basis sets are already close to the basis set limit, with errors of 0.8% and 0.5%, respectively. Thus, 1s1p1d augmentation is adequate irrespective of the orbital basis set size.

The effect of multiple augmentation with diffuse polarization functions is illustrated in Fig. 2. Starting with the unpolarized split-valence (SV), triple-zeta valence (TZV), and quadruple-zeta valence (QZV) basis sets for fluorine,²⁶ we successively added primitive d functions. All d BF exponents were reoptimized in each step. As is seen from Fig. 2, the polarizability increment $\Delta \alpha_n$ of the *n*-fold augmented basis set relative to the n-1 augmentation decreases exponentially with n. Therefore, each additional d function yields an improvement about one order of magnitude smaller than the previous BF. The observed strong decrease of the polarizability increment suggests that a large number of diffuse BFs are unnecessary if optimized exponents are used. The polarizability increment is virtually independent of the initial basis set here since d orbitals have no significant occupation in the UHF ground state of fluorine atom.

V. BASIS SET DESIGN

Augmented Gaussian basis sets were constructed for the elements H-Rn (except lanthanides) by diffuse augmentation of the Karlsruhe def2-SVP, def2-TZVPP, and def2-QZVPP basis sets.²⁶ Stuttgart effective core potentials (ECPs) were used^{41,42} for the elements Rb–Rn. All basis set optimizations

TABLE I. Augmentation for def2 basis sets.

Groups (block)		Augmentation	
	SVP	TZVP/TZVPP	QZVP/QZVPP
1-2(s)	1 <i>p</i>	$1p^{a,b}$	1 <i>p</i> ^c
3–12 (<i>d</i>)	$1p^{d}$	$1p^{e,f}$	$1p^{g}$
13–15 (<i>p</i>)	1s1d	1s1d	1s1d
16–18 (<i>p</i>)	1s1p1d	1s1p1d	1s1p1d

^aAugmentation only for TZVPP basis sets of H, He, Li, Be, Na, and Mg. ^b2p polarization set from Ref. 43 was used for K instead of 1p. ^cAugmentation only for QZVPP basis sets of H and He.

^d1p was inserted in the SVP polarization basis sets of Y–Cd, La, and Hf–Hg.

^eAugmentation only for TZVPP basis sets of Sc-Zn, La, and Hf-Hg.

1p was inserted in the TZVPP polarization basis sets of Au and Hg.

^gAugmentation only for QZVPP basis sets of Hf-Re.

were performed for isotropic static UHF polarizabilities of the atoms in their ground states. The starting def2 basis sets were successively augmented with uncontracted diffuse BFs, with a reoptimization of all diffuse basis set exponents performed in each augmentation step. The target accuracy for the augmented SVP basis sets, denoted by SVPD in the following, was 6%-8% basis set error for atomic polarizabilities. Augmented TZVPP basis sets (TZVPPD) were required to have no more than 2% basis set error, and augmented QZVPP basis sets (QZVPPD) were required to be within 1% of the atomic basis set limits.

The resulting augmentation is shown in Table I. The augmentation patterns for the individual elements are determined by the angular momentum of the highest occupied orbital. Augmentation with one set of diffuse p functions (1p) is used for s- and d-block elements. 1s1d augmentation is employed for early *p*-block elements (groups 13-15), while late p-block elements (groups 16-18) require additional diffuse valence p functions, corresponding to a 1s1p1daugmentation. As in the case of fluorine atom, there is no need to increase the diffuse augmentation with the size of the initial basis set to achieve the target accuracy. For s- and d-block metals, we even observe the opposite trend, i.e., fewer additional functions are needed because the def2-TZVPP and def2-QZVPP basis sets for these elements already include fairly diffuse valence and polarization BFs. As a result, the def2-QZVPP basis sets of 2s, 3s, and 4d elements as well as the def2-TZVPP and def2-QZVPP basis sets of most 4s-6s and 5d elements meet our accuracy criteria without further augmentation. Exceptions are def2-TZVPP basis sets of K, Au, and Hg and def2-QZVPP basis sets of Hf–Re, which require an additional p function in the polarization (rather than diffuse) subset. def2-SVP basis sets of 4dand 5d elements were also augmented using an additional ppolarization function per basis set. The exponents of the additional polarization functions were obtained by geometric interpolation of the neighboring BF exponents. These additional polarization functions have little effect on atomic ground-state energies but appear to be essential for a correct description of first-order response. The exponents of diffuse basis functions follow the general trend of the lowest exponents of unaugmented basis sets and differ from the latter by factors of 0.2-0.6. The ratio depends on the position in the

TABLE II. Basis set errors of isotropic static HF polarizabilities of neutral atoms H–Rn (except lanthanides) with def2 basis sets and augmented def2 basis sets in %. Mean errors, mean unsigned (mean uns.) errors, and maximum negative (max. neg.) and positive (max. pos.) deviations relative to the extended def2-QZVPP basis sets. See text and Table I for the basis set definitions.

	SVP	SVPD	TZVPP	TZVPD	TZVPPD	QZVPP	QZVPD	QZVPPD
Mean	-33.1	-2.4	-20.2	-2.2	-1.4	-8.1	-0.3	-0.2
Mean uns.	33.1	3.8	20.2	2.3	1.5	8.1	0.5	0.4
Max. neg.	-88.3	-10.8	-57.8	-15.8	-6.1	-38.4	-2.2	-1.8
	(H)	(Al)	(H)	(Ti)	(Ga)	(H)	(Pt)	(Nb)
Max. pos.	3.5	10.6		1.0	1.0	0.0	1.3	1.3
	(La)	(Li)		(Na)	(Na)	(Cs)	(Ne)	(Ne)

Periodic Table and the size of the underlying orbital basis set. The diffuse augmentation developed for def2-TZVPP and def2-QZVPP basis sets may also be combined with the smaller def2-TZVP and def2-QZVP basis sets, ²⁶ yielding the TZVPD and QZVPD basis sets, respectively. Further details of the basis set construction and the final optimized basis sets are available as supplementary information.⁴⁴ The new basis sets are also available for free download from http://www.turbomole.com and from EMSL basis set exchange.⁴⁵

VI. ASSESSMENT

A. Atomic polarizabilities

Table II summarizes statistical error measures for augmented and unaugmented Karlsruhe def2 basis sets. (A complete list of results for atomic polarizabilities is provided as supplementary information.) The relative basis set errors of isotropic static polarizabilities of the atoms H-Rn (except lanthanides) in the UHF approximation are shown. The reference atomic polarizability values, which estimate the HF basis set limits, were computed using extended def2-QZVPP basis sets constructed by downward extrapolation from def2-QZVPP basis sets²⁶ $(1s_1p_1d_1f$ extrapolation for the elements H–Be and 1s1p1d1f1g extrapolation for all others). These basis sets yield results within 1% of the published basis set limits for atomic HF polarizabilities.⁴⁰ Mean errors, mean unsigned errors, as well as the largest positive and negative deviations are given for each basis set. Positive basis set errors are mostly due to the effect of diffuse augmentation on ground-state wave functions and are discussed below. The addition of optimized diffuse basis functions strongly improves atomic polarizabilities, reducing both mean unsigned errors and maximal errors. For instance, a mean unsigned error of 33.1% with def2-SVP basis sets is reduced to 3.8% for SVPD, which perform better than the unaugmented def2-QZVPP basis sets (8.1% mean unsigned error). Augmented triple-zeta basis sets TZVPPD have 1.5% mean unsigned error for the atoms, while QZVPPD results are within 1% of the basis set limit (0.4% mean unsigned error relative to the reference basis sets). TZVPD and QZVPD basis sets perform well for atomic polarizabilities, with mean unsigned errors of 2.3% and 0.5%, respectively. The largest negative deviations are found for the 3p elements Al–P and their 4p congeners. The addition of a diffuse 1p basis set would not only reduce the basis set errors further but would also affect occupied ground-state orbitals and was thus not implemented.

Due to the maximum property of the polarizability (see Sec. II), one might expect that all deviations from the basis set limits would be negative. While this is true for most atoms, some elements (in particular, metals) show substantial positive deviations with small basis sets, exceeding the basis set limit by as much as 10%, e.g., for Li with the SVPD basis set. This deviation from the variational property of static polarizabilities can have two causes. Besides the incompleteness of reference basis sets, which can account for up to 1% positive error, the large positive deviations seen in metal atoms result from the ground-state bias introduced to some extent by diffuse augmentation, leading to a partial breakdown of the polarizability variational principle.

The balance between the valence, polarization, and diffuse subsets is crucial for accurate molecular properties. This balance is particularly important for atoms with relatively low-lying diffuse excited states such as alkali metals. For example, the polarizability of Li atom is underestimated by 38.6% using the def2-SVP basis set. However, upon 1p augmentation (SVPD basis set), the atomic polarizability is overestimated by 10.6%. Using triple-zeta basis sets strongly reduces the ground-state bias: the TZVPPD polarizability of Li atom is 1.7% smaller than the basis set limit, in accord with the variational principle. We conclude that the groundstate bias and the resulting relative overestimation of polarizabilities may lead to fortuitous error cancellation if small basis sets with extensive augmentation are used, but it also produces numerical instability in larger molecular applications and a strongly system-dependent tendency to overpolarize. Our strategy to minimize this imbalance is to start from fully optimized valence basis sets and add the smallest possible number of diffuse functions necessary to obtain accurate polarizabilities.

B. Molecular tests of augmented basis sets

We used a set of 313 small molecules including the elements H–At (except lanthanides and noble gases) to assess the quality of the optimized augmented basis sets across a broad range of bonding situations and oxidation states. The test set employed here is based on the compilation of Weigend and Ahlrichs,²⁶ with a few substitutions.⁴⁶ Mean errors, mean unsigned errors, and maximal negative and positive deviations are presented in Table III for SVPD, TZ-VPD, TZVPPD, QZVPD, and QZVPPD basis sets. The results for the def2-SVP, def2-TZVPP, and def2-QZVPP basis

TABLE III. Basis set errors of isotropic static HF polarizabilities of 313 small molecules with def2 basis sets and augmented def2 basis sets in %. Mean errors, mean unsigned (mean uns.) errors, and maximum negative (max. neg.) and positive (max. pos.) deviations relative to the extended def2-QZVPP basis sets. See text and Table I for the basis set definitions.

	SVP	SVPD	TZVPP	TZVPD	TZVPPD	QZVPP	QZVPD	QZVPPD
			1 <i>s</i> , 2 <i>p</i> e	lements (40	molecules)			
Mean	-26.0	-3.1	-10.8	-0.6	-0.4	-5.8	-0.1	-0.1
Mean uns.	26.0	3.1	10.8	0.6	0.4	5.8	0.1	0.1
Max. neg.	-49.8	-5.9	-27.3	-2.2	-1.2	-17.4	-0.3	-0.3
U	(HF)	(N_2)	(HF)	(H_2)	(C_2H_2)	(HF)	(N_2H_2)	(N_2H_2)
Max. pos.		••••		0.2	0.4		0.6	0.6
				(F_2)	(H_2)		(HF)	(HF)
				· 2/	× 2/			. ,
			3p ele	ments (31 m	nolecules)			
Mean	-29.2	-2.5	-13.6	-0.3	-0.3	-5.7	-0.0	-0.0
Mean uns.	29.2	2.5	13.6	0.3	0.3	5.7	0.1	0.1
Max. neg.	-50.4	-9.3	-30.0	-1.9	-1.9	-14.4	-0.5	-0.5
	(HCl)	(P ₂)	(HCl)	(P_2)	(P_2)	(HCl)	(P_2)	(P_2)
Max. pos.		•••		0.3	0.3		0.4	0.4
		•••	•••	(Cl_2)	(Cl_2)	•••	(HCl)	(HCl)
			4n ele	ments (26 m	olecules)			
Mean	-24.7	-1.7	-11.6	-0.3	-0.3	-6.8	-0.1	0.0
Mean uns.	24.7	1.9	11.6	0.4	0.4	6.8	0.2	0.1
Max. neg.	-41.1	-3.5	-22.3	-1.2	-1.1	-18.2	-0.5	-0.1
U	(HBr)	(GaF ₂)	(HBr)	(AsH ₃)	(AsH ₃)	(HBr)	(GeF_4)	(GaF_2)
Max. pos.	•••	1.4		0.3	0.3		0.3	0.3
*		(GaO)		(SeH ₂)	(SeH ₂)		(SeH ₂)	(SeH ₂)
			5p ele	ments (22 m	nolecules)			
Mean	-25.8	-1.7	-12.6	-0.2	-0.2	-4.4	-0.0	-0.0
Mean uns.	25.8	1.7	12.6	0.2	0.2	4.4	0.1	0.1
Max. neg.	-41.5	-3.8	-22.3	-0.7	-0.7	-8.3	-0.2	-0.2
	(HI)	$(InCl_3)$	(HI)	(SbF)	(SbF)	(HI)	(SbF)	(SbF)
Max. pos.		•••		(0.1)	(0.10)		(0.1	(01.01-)
				(SDO_2)	(SDO_2)		$(SDCI_6)$	$(SDCI_6)$
			6p ele	ments (18 m	nolecules)			
Mean	-23.5	-2.0	-11.7	-0.2	-0.1	-4.1	-0.0	-0.0
Mean uns.	23.5	2.0	11.7	0.3	0.2	4.1	0.1	0.1
Max. neg.	-38.6	-3.7	-24.5	-0.9	-0.9	-8.2	-0.3	-0.3
	(BiF)	(TlCl ₃)	(BiF)	(BiF)	(BiF)	(BiF)	(BiF)	(BiF)
Max. pos.		0.0		0.4	0.4		0.1	0.1
		(At_2)		(PbF_3)	(PbF_3)		(At_2)	(At ₂)
			2 ((F Q				
Mean	-18.1	_3 3	-10.6	-1.3	-1.2	_38	-0.4	-0.3
Mean uns	18.3	3.0	10.0	1.3	1.2	3.0	0.4	0.5
Max neg	-40.9	-12.4	-24.8	-7.0	-54	-10.1	-21	-2.4
Max. neg.	(LiCl)	$(I \downarrow O)$	(LiF)	$(\mathbf{I};\mathbf{O})$	(N_2, Ω)	(N ₂ E)	(B2E)	(N_2, Ω)
Max nos	(LICI) 5.6	(L120)	(LII') 	(LI ₂ O)	17	(1111)	0.6	0.7
Max. pos.	(Na ₂ O)	(K₂P)		(K ₂ P)	(K ₂ P)		(MgF2)	(NaF)
	(11420)	(1131)		(1131)	(1131)		(11812)	(1 (112)
			3d ele	ments (50 m	nolecules)			
Mean	-29.8	-4.4	-10.4	-1.5	-1.1	-3.0	-0.6	-0.2
Mean uns.	29.8	5.0	10.4	1.6	1.2	3.0	0.7	0.3
Max. neg.	-53.7	-29.5	-21.9	-9.9	-9.1	-7.3	-6.5	-3.7
	(NiO)	(FeO)	(NiO)	(FeO)	(FeO)	(ZnF_2)	(FeO)	(TiO)
Max. pos.	•••	4.4	•••	2.1	1.1	•••	0.3	0.4
	•••	(CrO_3)		(ZnO)	(ZnO)		(MnO_4^-)	(MnO)

	SVP	SVPD	TZVPP	TZVPD	TZVPPD	QZVPP	QZVPD	QZVPPD
			4d elem	ments (35 m	olecules ^a)			
Mean	-13.8	-3.3	-7.6	-2.3	-2.0	-2.5	-0.9	-0.3
Mean uns.	16.8	4.7	7.6	3.0	2.2	2.5	0.9	0.3
Max. neg.	-30.8	-15.1	-19.2	-12.5	-10.1	-6.9	-5.8	-1.4
	(PdF)	(PdF)	(PdF)	(PdF)	(PdF)	(CdF_2)	(PdF)	(MoH)
Max. pos.	40.7	23.1	0.3	6.0	3.4		0.6	0.2
	(RhF_4)	(RhF_4)	(RhF_4)	(RhF_4)	(RhF_4)		(RhF_4)	(RhF_4)
			5d ele	ments (33 m	olecules)			
Mean	-14.9	-3.5	-8.3	-2.4	-2.1	-2.9	-1.1	-0.7
Mean uns.	15.9	5.0	8.3	3.2	2.5	2.9	1.5	1.1
Max. neg.	-28.8	-17.8	-25.2	-17.8	-14.1	-12.9	-11.4	-11.5
	(HfF_3)	(WO)	(ReH)	(TaF)	(WO)	(WO)	(WO)	(WO)
Max. pos.	16.5	8.1		6.2	6.0		5.9	5.9
	(ReO_2)	$(AuCl_3)$		$(OsOF_5)$	$(OsOF_5)$	•••	$(OsOF_5)$	$(OsOF_5)$
			Tota	al (313 mole	cules ^b)			
Mean	-22.7	-3.1	-10.6	-1.1	-1.0	-4.2	-0.4	-0.2
Mean uns.	23.2	3.6	10.6	1.4	1.1	4.2	0.5	0.3
Max. neg.	-53.7	-29.5	-30.0	-17.8	-14.1	-18.2	-11.4	-11.5
	(NiO)	(FeO)	(HCl)	(TaF)	(WO)	(HBr)	(WO)	(WO)
Max. pos.	40.7	23.1	0.3	6.2	6.0		5.9	5.9
	(RhF_4)	(RhF_4)	(RhF_4)	(OsOF ₅)	(OsOF ₅)		(OsOF ₅)	(OsOF ₅)

TABLE III. (Continued.)

^a34 molecules for SVP basis sets (ZrF excluded).

^b312 molecules for SVP basis sets (ZrF excluded).

sets are shown for comparison. Extended def2-QZVPP basis sets (see Sec. V for definition) were used as a reference. The geometric structures and electronic occupations of all molecules in the test set are provided as supplementary information.

The test set is divided into subsets corresponding to different blocks of the Periodic Table in order to analyze the consistency of the basis set construction. The 1s, 2p subset includes compounds consisting of the elements H and B-F. Unaugmented basis sets show rather slow convergence toward the basis set limit. Even with extensive def2-QZVPP basis sets, which are virtually at the basis set limit for HF ground-state energies,^{26,29} the mean unsigned error is 5.8% for polarizabilities, while the maximal deviation is -17.4%. On the other hand, SVPD basis sets are already of fair accuracy (3.1% mean unsigned error and -5.9% maximum error). TZVPD and TZVPPD results are more accurate and very similar to each other, with mean unsigned errors of less than 1%. QZVPD and QZVPPD results are essentially at the basis sets limit for polarizabilities, featuring less than 0.5% residual basis set error.

Results for compounds of 3p (Al–Cl) and 4p (Ga–Br) elements are similar to those for 2p elements. SVPD basis sets make for a reasonable accuracy with less than 3% mean unsigned error. Polarizabilities computed with TZVPD and TZVPPD basis sets are accurate to within 1% relative to the reference, while QZVPD and QZVPPD basis sets are again very close to (occasionally above) the reference value. The use of ECPs for core electrons in the heavier 5p (In–I) and

6p (Tl-At) elements has little impact on polarizabilities and does not alter the basis set convergence. SVPD basis sets are accurate to within 2%, while the augmented triple- and quadruple-zeta basis sets are within 0.5% of the reference value for compounds of 5p and 6p elements.

Alkali and alkali earth metals form strongly ionic compounds, which pose a challenge for any basis set optimization method based on neutral atoms. High electron densities on the electronegative counterions slow down the basis set convergence of polarizabilities, especially in fluorides and oxides, and lead to large basis set errors for unaugmented basis sets. In addition, positive deviations from the basis set limit are observed for small metal clusters such as Mg₄ and Li₈ and the phosphides Na₃P and K₃P. The positive deviations are largest for SVPD basis sets and become smaller for TZVPPD and QZVPPD basis sets, which indicate that they are caused by an imbalance between the valence/polarization and diffuse parts of the basis sets, similar to the parent metal atoms (see Sec. VIA). The mean unsigned errors for the s-block compounds are 3.9% with SVPD basis sets, 1.3% with TZVPD and TZVPPD basis sets, and 0.5% with QZ-VPD and QZVPPD basis sets, respectively.

Many transition metal compounds exhibit significant multireference character, i.e., multiple determinants contribute strongly to the ground-state wave function. Frequently, the multireference character manifests itself as an instability of the HF ground-state solution,^{47–49} leading in extreme cases even to negative static HF polarizabilities, e.g., -250.8 a.u. for ZrF with def2-SVP basis sets. The basis set conver-

gence is generally slower in these molecules compared to compounds with stable HF ground states, with larger deviations, both positive and negative, occurring for small basis sets. Typical examples are the low-valent metal oxides, e.g., FeO and TiO, and fluorides, e.g., PdF and HfF.

The basis set convergence of 3*d* transition metal compounds is affected both by stability issues and in part by ionic character. Polarizabilities are underestimated by 29.8% on average with def2-SVP basis sets; the maximum deviation is -53.7% for NiO. Diffuse 1*p* augmentation reduces the mean unsigned error to 5.0%, while the largest negative and positive deviations are -29.5% for FeO and 4.4% for CrO₃, respectively. Mean unsigned errors of augmented basis sets decrease to 1.2% and 0.3% for the TZVPPD and QZVPPD basis sets, respectively. However, maximum deviations can still be significant with augmented basis sets in the presence of HF ground-state instabilities. For example, the deviation is as large as -3.7% for TiO with the QZVPPD basis set.

Similarly, larger maximum deviations are seen for the 4dtransition metal compounds, particularly with small basis sets. def2-SVP basis sets show a maximum negative deviation of -30.8% for PdF and a maximum positive deviation of 40.7% for RhF₄. The mean unsigned error is 16.8%. Augmentation with a 1p set reduces the maximum errors to -15.1% and 23.1\%, respectively, while the mean unsigned error decreases to 4.7%. Both the mean error and the spread are reduced with TZVPPD and QZVPPD basis sets compared to their unaugmented counterparts. Mean unsigned errors are 2.2% and 0.3% with TZVPPD and QZVPPD basis sets, respectively. Maximum positive deviations are 3.4% and 0.2% with TZVPPD and QZVPPD basis sets, respectively. The basis set convergence of 5d element compounds shows complications similar to that with 4d metal compounds. While large deviations persist in problematic cases, such as WO and OsOF₅, even with large augmented QZVPD and QZVPPD basis sets, the mean unsigned errors show satisfactory convergence in the sequence of augmented basis sets. SVPD basis sets have a mean unsigned error of 5.0%, TZVPPD basis sets show 2.9% basis set error, while QZ-VPPD basis sets are off the reference by 1.1%.

C. Summary and recommendations

SVPD basis sets exhibit a mean unsigned error of 3.6% across the entire test set and are thus sufficiently accurate for polarizability estimates and for exploratory calculations. With mean unsigned errors of 1.4% and 1.1%, respectively, TZVPD and TZVPPD basis sets may be recommended for accurate routine calculations. They might also be preferable for exploratory calculations in problematic cases such as metal compounds. QZVPD and QZVPPD polarizabilities are essentially at the basis set limit, with 0.5% and 0.3% mean unsigned errors, respectively. Nevertheless, basis set convergence may be severely impacted by ground-state instability problems. In addition, the moderate augmentation developed in this work is likely to be insufficient for notoriously difficult cases such as (small) anions,^{50,51} which require multiple sets of diffuse BFs.



FIG. 3. Basis set convergence of isotropic static polarizabilities of HCN (upper panel) and SO₂ (lower panel) at experimental equilibrium geometries with HF, density functional, and MP2 methods. See text for computational details.

VII. APPLICATIONS

A. Basis set convergence of polarizabilities with density functional and second-order Møller–Plesset methods

The basis set convergence of static polarizabilities of HCN and SO₂ using HF, density functional, and MP2 theory is displayed in Fig. 3. Approximate exchange-correlation functionals include the local spin density approximation (LSDA) of Perdew and Wang,⁵² the gradient-corrected functional of Becke and Perdew (BP),^{53,54} and the hybrid functional of Perdew, Burke, and Ernzerhof (PBE0).⁵⁵ In addition, the effect of the resolution-of-the-identity approximation (RI-*J*) for the Coulomb operator^{56–60} is investigated for the BP functional (BP/RI). The RI-*J* approximation yields speedups of 10–100 for density functional calculations of ground-state energies^{58,60} and response properties^{61,62} with nonhybrid exchange-correlation functionals.

All polarizabilities were computed at the experimental structures from Ref. 63 for HCN and from Ref. 64 for SO₂, respectively. HF and density functional results were obtained by static linear-response calculations, while static MP2 polarizabilities were computed by finite-field methods. For comparison, the experimental static polarizabilities of HCN and SO₂ are 17.5 a.u. (Ref. 65) and 23.45 a.u.,⁶⁶ respectively. HF theory is known to underestimate molecular polarizabilities,^{67–69} while LSDA and gradient-corrected functionals tend to overestimate polarizabilities.^{68,70} Hybrid functionals yield quite accurate polarizabilities and are considered a good compromise between accuracy and computational cost.^{71,72} The MP2 method significantly improves

TABLE IV. Static isotropic polarizability α_{iso} (in a.u.) of fullerene C₆₀.

Method	Basis set	$lpha_{ m iso}$
PBE0	SVP	505.8
PBE0	SVPD	551.1
PBE0	TZVPPD	555.5
PBE0	QZVPPD	555.9
PBE0	aug-cc-pVDZ	544.2
PBE0	Sadlej	556.0
HF	QZVPPD	543.6
HF	6-31++G	507.0 ^a
CC2	aug-cc-pVDZ	623.7 ^b
CCSD	ZPolX 554.7 ^c	
Expt.		516 ± 54 , ^d 533 ± 27^{e}

^aReference 73.

^bReference 74.

^cReference 75.

^dMolecular beam deflection (Ref. 76).

^eMass spectrometry (Ref. 77).

upon HF for static polarizabilities and compares well to higher-level correlation methods.^{67,69} However, all computed polarizabilities exhibit slow convergence toward basis set limit with unaugmented basis sets. In fact, as seen from Fig. 3, the basis set errors outweigh the differences between all methods for def2-SVP basis sets and are larger than or comparable to method errors if def2-TZVPP basis sets are used. Even with def2-QZVPP, the polarizabilities are not converged with respect to the basis set, as indicated by the slope of polarizability curves in Fig. 3, in agreement with the HF results of Sec. VI B.

Our property-optimized augmented basis sets improve basis set convergence significantly for all methods considered here; note the different scales for augmented basis sets in Fig. 3. The basis set errors are clearly smaller than the method errors if TZVPPD or QZVPPD basis sets are used. The residual error of TZVPPD basis sets compared to QZ-VPPD is less than 1% for HF and density functional methods and for MP2. Thus, the performance of the augmented basis sets developed here appears to be transferable to density functional and correlated wave function methods. More extensive tests involving the 1s, 2p, and 3p subsets of the molecular test set from Sec. VI B show that SVPD basis sets are within 4% of the basis sets limit for the BP and PBE0 functionals, while TVZPPD and QZVPPD basis sets are within 1% and 0.5% of the reference values, respectively. (See supplementary information for the complete results.) The RI-J errors are small even though unaugmented auxiliary basis sets⁵⁹ were used. The effect of the RI-J approximation is $\sim 0.1\%$ and thus well below the atomic orbital basis set error. Full statistical information on the molecular test calculations with BP and PBE0 functionals is available as supplementary information.

B. Polarizabilities of icosahedral fullerenes

The static polarizability of fullerene C_{60} is of fundamental and technological importance and has repeatedly triggered theoretical studies,^{73,74} most recently a massively parallel coupled-cluster singles doubles calculation on 1024

TABLE V. Lowest Gaussian exponents ζ_{\min} of *s*, *p*, and *d* basis functions and overlap condition numbers κ of fullerene C₆₀ with various diffuse augmented basis sets.

Basis set	$\zeta_{\min}(s)$	$\zeta_{\min}(p)$	$\zeta_{\min}(d)$	к
SVPD	0.067	0.153	0.117	4.1×10^{7}
TZVPPD	0.048	0.101	0.091	4.8×10^{9}
QZVPPD	0.045	0.076	0.077	$6.8 imes 10^{11}$
aug-cc-pVDZ ^a	0.047	0.040	0.151	3.8×10^{9}
aug-cc-pVTZ ^a	0.044	0.036	0.100	4.8×10^{10}
aug-cc-pVQZ ^a	0.041	0.032	0.077	5.4×10^{11}
aug-pc-1 ^b	0.045	0.034	0.101	4.9×10^{10}
aug-pc-2 ^b	0.037	0.026	0.079	2.1×10^{12}
aug-pc-3 ^b	0.032	0.019	0.057	3.3×10^{15}
Sadlej ^c	0.048	0.039	0.039	5.7×10^{9}

^aReference 11.

^bReference 17.

^cReferences 19 and 20.

CPUs.⁷⁵ Table IV displays the performance of the basis sets developed here along with previous theoretical and experimental results. The structures for the present calculations were optimized at the BP/SVP level; quadrature grids were of size m4.

The SVPD, TZVPPD, and QZVPPD basis results show rapid, monotonous basis set convergence of the isotropic polarizability. Taking the QZVPPD result as a reference, the SVPD basis set error is already below 1%. This rapid convergence is in part due to the size of C₆₀, making basis set convergence less critical than in smaller systems. The augcc-pVDZ basis set performs slightly worse than SVPD even though it contains an extra *p* shell and considerably more diffuse exponents. The aug-cc-pVDZ, Sadlej, and QZVPPD basis sets required tighter integral neglect thresholds to converge the ground-state energy to 10^{-7} hartree. Thus, while the Sadlej result is appealingly close to the basis set limit, the use of this basis set comes at a high cost and may not be practical for larger or less symmetric systems.

Although the agreement of the converged PBE0 result of 555.9 a.u. with the CCSD result of 554.4 a.u. is somewhat fortuitous, Table IV is strong evidence for the excellent price to performance ratio of PBE0/SVPD. In contrast, the PBE0/SVP result is off by 50.1 a.u., approximately five times the difference between the converged HF and the CCSD results. This once more underlines the importance of basis set error versus method error in polarizability calculations, especially in larger systems, where augmented basis sets were previously prohibitive.

The numerical difficulties, which are frequently encountered with diffuse augmented basis sets, can be traced back to the condition of the overlap matrix. The presence of very low basis function exponents in augmented basis sets increases the condition number by several orders of magnitude, resulting in slow convergence of iterative procedures and a significant accumulation of round-off errors in finiteprecision arithmetic. In Table V, we illustrate the evolution of the lowest Gaussian exponents and overlap condition numbers of fullerene C_{60} with increasing basis set size. As expected, condition numbers increase with growing basis set size for all basis set hierarchies. However, the SVPD, TZ-

TABLE VI. Static isotropic polarizability of icosahedral fullerenes (in a.u.).

	PPP ^a	ADFT ^b	PBE/5 $s3p1d^{c}$	PBE0/SVPD
C ₆₀	526.3	533.1	553.4	551.1
C ₂₄₀	2017.8	2 024.5	1990.8	2 017.8
C540	6222.0	7 794.3	8050.8	8 395.4
C ₇₂₀	9494.9	12 470.9	N/A	12 342.9

^aPariser–Parr–Pople theory (Ref. 78).

^bSum-over-states analytic DFT with screening correction (Ref. 79).

^cFinite-field PBE calculations (Ref. 79).

VPPD, and QZVPPD basis sets show condition numbers one to two orders of magnitude smaller than the corresponding aug-cc-pVXZ basis sets¹¹ and about three orders of magnitude smaller than the aug-pc-*n* basis set hierarchy¹⁷ for C_{60} . The main difference stems from the fact that energy minimizations performed on anions yield diffuse p functions for carbon, which have relatively small contributions to polarizabilities and are not included in property-optimized augmented basis sets (see Sec. V and Table I). The exponents of s functions are somewhat higher, while exponents of d functions are slightly lower for property-optimized basis sets compared to the aug-cc-pVXZ hierarchy. These results show that while higher overlap condition numbers due to diffuse augmentation cannot generally be avoided, full optimization of diffuse basis functions appears to result in the smallest increase in overlap numbers for a given level of accuracy.

Table VI compares PBE0/SVPD isotropic polarizabilities of icosahedral fullerenes with 60, 240, 540, and 720 carbon atoms to previous results. In view of its performance for C₆₀, the PBE0/SVPD method should be accurate to within few percent for higher fullerenes and will be used as a reference. For C_{60} and C_{240} , the PBE/5s3p1d finite-field results by Zope and co-workers⁷⁹ agree well with PBE0/ SVPD; there is a significant $\sim 4\%$ deviation for C₅₄₀, however. The sum-over-states analytic DFT results with semiempirical screening correction by the same authors agree well with PBE0/SVPD for C240 and C720 but are significantly too small for C_{60} and $C_{540}.$ The semiempirical Pariser–Parr–Pople method 78 is very accurate for C_{240} but underestimates the PBE0/SVPD polarizability of C720 by 23%. Clearly, the basis sets developed here make it possible to compute polarizabilities of molecules in the 100-1000 atoms regime at a level of theory that was too costly in the past.

The isotropic polarizability per atom (see Fig. 4) exhibits a minimum for C_{240} for all methods considered here. Remarkably, this trend does not correlate with the energy difference between the highest occupied and the lowest unoccupied MO, which is the largest for C_{60} , but is in line with strain arguments,⁸⁰ suggesting that C_{240} is the most inert fullerene. The PBE0/SVPD curve differs from the others by exhibiting a negative curvature for larger fullerene sizes, possibly indicating a saturation of the polarizability per atom for large icosahedral fullerenes.

VIII. CONCLUSIONS

Bigger is not necessarily better for diffuse basis set augmentation. The balance between improved response proper-



FIG. 4. Polarizability per atom (in a.u.) of icosahedral fullerenes as obtained from Pariser–Parr–Pople theory (PPP, Ref. 78), sum-over-states analytic DFT with screening correction (ADFT, Ref. 79), and PBE0 using the SVPD basis set.

ties and stability of the ground state can be delicate, especially for smaller valence basis sets. Excessive diffuse augmentation leads to spurious overpolarization, illconditioned overlap matrices, and a breakdown of integral prescreening and other low-order scaling methods. Thus, it has been very challenging to compute converged polarizabilities and other response properties for molecules with more than ~50 atoms in the past.

The basis sets developed here effectively address these problems for most density functional and low-cost correlated wave function treatments. The central ideas underlying our basis sets are as follows: (i) basis sets for molecular response calculations must be optimal for response properties rather than ground-state energies of atomic anions and (ii) the smallest possible number of diffuse functions should be employed. This strategy is implemented by variational maximization of atomic HF polarizabilities using analytical derivative methods. The resulting property-optimized basis sets are smaller and less diffuse than augmented Dunning or Sadlej basis sets, yet their accuracy for molecular polarizabilities is similar. The numerical stability and computational efficiency of the present basis sets makes it possible to compute near basis set limit polarizabilities for systems in the 100-1000 atoms regime. This eliminates a major source of error in computed polarizabilities of large systems since method errors of hybrid density functionals are typically a few percent, while basis set error using conventional nonaugmented basis sets can easily exceed 10%.

The basis sets developed here were optimized for dipole polarizabilities. Preliminary results indicate that they are also viable for transition dipole moments, Raman intensities, optical rotations, and up to third-order nonlinear-response properties. Further work along these lines is in progress.

ACKNOWLEDGMENTS

The authors thank Reinhart Ahlrichs and Florian Weigend for helpful suggestions and discussions. Support by the NSF Center for Chemical Innovation "Chemistry at the Space-Time Limit" (CaSTL), Grant Nos. CHE-0533162 and CHE-0802913, is acknowledged.

- ¹*Computational Molecular Spectroscopy*, edited by P. R. Bunker and P. Jensen (Wiley, Chichester, 2000).
- ²S. Mukamel, *Principles of Nonlinear Optical Spectroscopy* (Oxford University Press, New York, 1995).
- ³F. Neese, Coord. Chem. Rev. 253, 526 (2009).
- ⁴ Time-Dependent Density Functional Theory, Lecture Notes in Physics Vol. 706, edited by M. A. L. Marques, C. A. Ullrich, F. Nogueira, A. Rubio, K. Burke, and E. K. U. Gross (Springer, Berlin, 2006).
- ⁵O. Christiansen, P. Jørgensen, and C. Hättig, Int. J. Quantum Chem. 68, 1 (1998).
- ⁶I. Shavitt and R. J. Bartlett, *Many-Body Methods in Chemistry and Physics* (Cambridge University Press, Cambridge, 2009).
- ⁷E. R. Davidson and D. Feller, Chem. Rev. (Washington, D.C.) **86**, 681 (1986).
- ⁸J. O. Hirschfelder, W. Byers Brown, and S. T. Epstein, Adv. Quantum Chem. 1, 255 (1964).
- ⁹J. Chandrasekhar, J. G. Andrade, and P. von Ragué Schleyer, J. Am. Chem. Soc. **103**, 5609 (1981).
- ¹⁰T. Clark, J. Chandrasekhar, G. W. Spitznagel, and P. von Ragué Schleyer, J. Comput. Chem. 4, 294 (1983).
- ¹¹ R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).
- ¹²D. E. Woon and T. H. Dunning, Jr., J. Chem. Phys. **98**, 1358 (1993).
- ¹³D. E. Woon and T. H. Dunning, Jr., J. Chem. Phys. 100, 2975 (1994).
- ¹⁴F. Jensen, J. Chem. Phys. **115**, 9113 (2001).
- ¹⁵F. Jensen, J. Chem. Phys. **116**, 3502 (2002).
- ¹⁶F. Jensen, J. Chem. Phys. **116**, 7372 (2002).
- ¹⁷ F. Jensen, J. Chem. Phys. **117**, 9234 (2002).
- ¹⁸ P. Manninen and J. Vaara, J. Comput. Chem. 27, 434 (2006).
- ¹⁹A. J. Sadlej, Collect. Czech. Chem. Commun. **53**, 1995 (1988).
- ²⁰A. J. Sadlej, Theor. Chim. Acta **79**, 123 (1991).
- ²¹Z. Benkova, A. J. Sadlej, R. E. Oakes, and S. E. J. Bell, J. Comput. Chem. 26, 145 (2005).
- ²²Z. Benkova, A. J. Sadlej, R. E. Oakes, and S. E. J. Bell, Theor. Chem. Acc. **113**, 238 (2005).
- ²³H.-J. Werner and W. Meyer, Mol. Phys. **31**, 855 (1976).
- ²⁴E. A. Hylleraas, Z. Phys. **65**, 209 (1930).
- ²⁵D. Rappoport and F. Furche, J. Chem. Phys. **126**, 201104 (2007).
- ²⁶F. Weigend and R. Ahlrichs, Phys. Chem. Chem. Phys. 7, 3297 (2005).
- ²⁷ A. Schäfer, H. Horn, and R. Ahlrichs, J. Chem. Phys. **97**, 2571 (1992).
- ²⁸ A. Schäfer, C. Huber, and R. Ahlrichs, J. Chem. Phys. **100**, 5829 (1994).
- ²⁹ F. Weigend, F. Furche, and R. Ahlrichs, J. Chem. Phys. **119**, 12753 (2003).
- ³⁰K. Faegri, Jr. and J. Almlöf, J. Comput. Chem. 7, 396 (1986).
- ³¹F. London, J. Phys. Radium 8, 397 (1937).
- ³²C. L. Darling and H. B. Schlegel, J. Phys. Chem. **98**, 5855 (1994).
- ³³A. J. Sadlej, Chem. Phys. Lett. **47**, 50 (1977).
- ³⁴J. A. Hudis and R. Ditchfield, Chem. Phys. Lett. 77, 202 (1981).
- ³⁵T. Helgaker and J. Almlöf, J. Chem. Phys. **89**, 4889 (1988).
- ³⁶ F. Furche and R. Ahlrichs, J. Chem. Phys. **117**, 7433 (2002).
- ³⁷ F. Furche and R. Ahlrichs, J. Chem. Phys. **121**, 12772 (2004).
- ³⁸R. Ahlrichs, M. Bär, M. Häser, H. Horn, and C. Kölmel, Chem. Phys. Lett. **162**, 165 (1989).
- ³⁹ TURBOMOLE, version 6.2, see http://www.turbomole.com.
- ⁴⁰ J. Stiehler, and J. Hinze, J. Phys. B 28, 4055 (1995).
- ⁴¹M. Dolg, H. Stoll, and H. Preuss, J. Chem. Phys. 90, 1730 (1989).
- ⁴²D. Andrae, U. Häußermann, M. Dolg, H. Stoll, and H. Preuß, Theor. Chim. Acta 77, 123 (1990).
- ⁴³A. J. H. Wachters, J. Chem. Phys. **52**, 1033 (1970).
- ⁴⁴ See supplementary material at http://dx.doi.org/10.1063/1.3484283 for orbital occupations and HF energies of atoms and molecules in the test set, static isotropic HF, DFT, and MP2 polarizabilities of the test set, basis set definitions of augmented basis sets SVPD, TZVPD, TZVPPD, QZVPD, and QZVPPD.

- ⁴⁵EMSL basis set exchange, see http://bse.pnl.gov.
- ⁴⁶Compared to the test set of Weigend and Ahlrichs (Ref. 26), XeF₂, XeF₄, XeOF₄, MnO₃F, and HfF were removed. P₄, At₂, Bi₄, PoO, PoO₂, Cs₂, and ZnO were added.
- ⁴⁷D. J. Thouless, *The Quantum Mechanics of Many-Body Systems*, 2nd ed. (Academic Press, New York, 1972).
- ⁴⁸ J. Čížek and J. Paldus, J. Chem. Phys. **47**, 3976 (1967).
- ⁴⁹R. Seeger and J. A. Pople, J. Chem. Phys. 66, 3045 (1977).
- ⁵⁰ M. G. Papadopoulos, J. Waite, and A. D. Buckingham, J. Chem. Phys. 102, 371 (1995).
- ⁵¹O. Ventura, M. Kieninger, and I. Cernusak, J. Mol. Struct. **436–437**, 489 (1997).
- ⁵²J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).
- ⁵³A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
- ⁵⁴J. P. Perdew, Phys. Rev. B **33**, 8822 (1986).
- ⁵⁵ J. P. Perdew, M. Ernzerhof, and K. Burke, J. Chem. Phys. **105**, 9982 (1996).
- ⁵⁶E. Baerends, D. E. Ellis, and P. Ros, Chem. Phys. **2**, 41 (1973).
- ⁵⁷ B. I. Dunlap, J. W. D. Connolly, and J. R. Sabin, J. Chem. Phys. **71**, 3396 (1979).
- ⁵⁸K. Eichkorn, O. Treutler, H. Öhm, M. Häser, and R. Ahlrichs, Chem. Phys. Lett. **242**, 652 (1995).
- ⁵⁹ F. Weigend, Phys. Chem. Chem. Phys. **8**, 1057 (2006).
- ⁶⁰ F. Weigend, M. Kattannek, and R. Ahlrichs, J. Chem. Phys. **130**, 164106 (2009).
- ⁶¹ R. Bauernschmitt, M. Häser, O. Treutler, and R. Ahlrichs, Chem. Phys. Lett. 264, 573 (1997).
- ⁶² F. Furche and D. Rappoport, in *Computational Photochemistry*, edited by M. Olivucci (Elsevier, Amsterdam, 2005), pp. 93–128.
- ⁶³ F. Pawłowski, P. Jørgensen, J. Olsen, F. Hegelund, T. Helgaker, J. Gauss, K. L. Bak, and J. F. Stanton, J. Chem. Phys. **116**, 6482 (2002).
- ⁶⁴G. Graner, E. Hirota, T. Iijima, K. Kuchitsu, D. A. Ramsay, J. Vogt, and N. Vogt, in *Springer Materials—The Landolt–Börnstein Database*, edited by K. Kuchitsu (Springer, Berlin, 1998), Vol. 25A.
- ⁶⁵ A. L. McClellan, *Tables of Experimental Dipole Moments* (Freeman, San Francisco, 1963).
- ⁶⁶ R. Feng, G. Cooper, G. R. Burron, and C. E. Brion, Chem. Phys. **240**, 371 (1999).
- ⁶⁷M. A. Spackman, J. Phys. Chem. **93**, 7594 (1989).
- ⁶⁸C. Van Caillie and R. D. Amos, Chem. Phys. Lett. **291**, 71 (1998).
- ⁶⁹O. Christiansen, C. Hättig, and J. Gauss, J. Chem. Phys. **109**, 4745 (1998).
- ⁷⁰ S. J. A. van Gisbergen, P. R. T. Schipper, O. V. Gritsenko, E. J. Baerends, J. G. Snijders, B. Champagne, and B. Kirtman, Phys. Rev. Lett. **83**, 694 (1999).
- ⁷¹C. Van Caillie and R. D. Amos, Chem. Phys. Lett. **328**, 446 (2000).
- ⁷²C. Adamo, M. Cossi, G. Scalmani, and V. Barone, Chem. Phys. Lett. 307, 265 (1999).
- ⁷³ K. Ruud, D. Jonsson, and P. R. Taylor, J. Chem. Phys. **114**, 4331 (2001).
- ⁷⁴ T. B. Pedersen, A. M. J. Sánchez de Merás, and H. Koch, J. Chem. Phys. 120, 8887 (2004).
- ⁷⁵ K. Kowalski, J. Hammond, W. de Jong, and A. J. Sadlej, J. Chem. Phys. 129, 226101 (2008).
- ⁷⁶ R. Antoine, P. Dugourd, D. Rayane, E. Benichou, and J. Broyer, J. Chem. Phys. **110**, 9771 (1999).
- ⁷⁷ A. Ballard, K. Bonin, and J. Louderback, J. Chem. Phys. **113**, 5732 (2000).
- ⁷⁸A. Ruiz, J. Bretón, and J. M. G. Llorente, J. Chem. Phys. **114**, 1272 (2001).
- ⁷⁹ R. R. Zope, T. Baruah, M. R. Pederson, and B. I. Dunlap, Phys. Rev. B 77, 115452 (2008).
- ⁸⁰ R. C. Haddon, G. E. Scuseria, and R. E. Smalley, Chem. Phys. Lett. 272, 38 (1997).