

Gaussian Basis Sets for Molecular Applications

J. Grant Hill

The choice of basis set in quantum chemical calculations can have a huge impact on the quality of the results, especially for correlated *ab initio* methods. This article provides an overview of the development of Gaussian basis sets for molecular calculations, with a focus on four popular families of modern atom-centered, energy-optimized bases: atomic natural orbital, correlation consistent, polarization consistent, and def2. The

terminology used for describing basis sets is briefly covered, along with an overview of the auxiliary basis sets used in a number of integral approximation techniques and an outlook on possible future directions of basis set design. © 2012 Wiley Periodicals, Inc.

DOI: 10.1002/qua.24355

Introduction

Ab initio and density functional theory (DFT) calculations typically require two main approximations, the “method” (the expansion of the many-electron wavefunction or the choice of exchange-correlation functional) and the “basis set” (the expansion of the one-electron orbitals). The choices made in these basic considerations for solving the electronic Schrödinger equation will determine the overall accuracy of the results and the associated computational cost required to obtain them. Although knowledge of basis set development is not a prerequisite for performing quantum chemical calculations, especially as many basis sets are included in widely available program packages or basis set archives,^[1–3] the aim of this review is to aid the theoretical/computational chemist in making informed basis set choices for molecular calculations, focusing on insights into the design, development, and optimization of the more recent developments in the field. It is the belief of the author that this will allow a feeling for the basis set incompleteness error in a given calculation to be developed, or “the right results for the right reason” to be obtained.

Detailed descriptions of the functions found within basis sets and their classifications can be found in a number of texts (e.g., Ref. [4]), hence the following description is kept purposefully brief to serve as a reminder of terminology. This review is concerned only with Gaussian basis sets,^[5,6] where the basis set is comprised of basis functions taking the form:

$$\chi(\zeta, n, m, l; r, \theta, \phi) = N Y_{l,m}(\theta, \phi) r^{2n-2-l} e^{-\zeta r^2} \quad (1)$$

or an analogous representation using Cartesian, rather than polar, coordinates. For the Gaussian-type orbital (GTO) shown in Eq. (1), N is a normalization constant, ζ is known as the exponent, and $Y_{l,m}$ are spherical harmonic functions. The indices n , m , and l determine the type of orbital (*s*, *p*, *d*, etc.). Basis functions of this explicit form are also known as Gaussian primitives (the Gaussian prefix shall be assumed herein) and, for reasons of balance and efficiency, a linear combination of a number of primitives is often undertaken to produce contracted functions:

$$\varphi = \sum_i n_i \chi_i \quad (2)$$

where n_i is known as a contraction coefficient.^[7] The methods and philosophies of optimizing the exponents and contraction coefficients, and their subsequent collection into a single defined whole (the basis “set”) will be the focus of this article.

How the exponents of a basis set are optimized has a large dependence on its intended usage. For example, if an investigation is focused solely on the electric properties of molecules, then using a basis set that has been optimized specifically for that purpose is likely to produce very good results.^[8] However, attempting to address all basis sets specialized for every different type of physicochemical property would represent a huge undertaking, hence this review will be mainly concerned with atom-centered, “energy-optimized” basis sets, where the exponents are optimized to minimize an electronic energy. Such basis sets are considered to be more general-purpose, yet are often augmented with additional functions when considering specific problems (see later). Although all exponents may be optimized using common minimization algorithms such as Broyden-Fletcher-Goldfarb-Shanno (BFGS) and simplex,^[9] an alternative option is to optimize a single exponent, then produce additional exponents using mathematical relationships such as “well-tempered,” “even-tempered,” and “Legendre polynomials.” Such methods are incredibly useful for large basis sets where optimization of every individual primitive is difficult, and have been discussed by Petersson et al.^[10] Even-tempered schemes also find use in extending basis sets with additional diffuse functions for the description of, for example, molecular anions. Such extensions often take the form:

$$\zeta_3 = \frac{\zeta_2}{\zeta_1/\zeta_2}, \quad (3)$$

J. G. Hill
School of Chemistry, University of Glasgow, Glasgow G12 8QQ, United Kingdom
E-mail: grant.hill@glasgow.ac.uk

© 2012 Wiley Periodicals, Inc.

J. Grant Hill obtained his PhD from the University of York (2006) for a modern valence bond study of aromaticity in pericyclic reactions. He then undertook a postdoc at Cardiff University using local electron correlation methods to study the noncovalent interactions occurring in models of biological systems. A second period of postdoctoral study saw him move to Washington State University where he worked on high-accuracy thermochemistry for the post-*d* main group elements. In 2010, he began an independent research program at the University of Glasgow and has recently been awarded a prestigious Royal Society of Edinburgh/Scottish Government Personal Research Fellowship. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]



where ζ_3 is the new exponent, ζ_2 is the next most diffuse exponent, and so forth.

Quality and contraction pattern

Working within the assumed framework of atom-centered, energy-optimized GTOs, basis sets can be further classified based on the number and type of functions included. The simplest basis sets are termed “minimal basis sets,” as they possess only the functions necessary to contain the electrons within an atom. For example, a minimal basis set for a neutral Ne atom would contain two *s* functions and a single three-component *p* function. In practice, minimal basis sets are rarely used due to the fact that they allow for little or no electron correlation.

One method of improving on a minimal basis set is to simply double the number of functions, leading to four *s* functions and two (three-component) *p* functions in our Ne example. In a similar fashion, one could triple, quadruple, and so forth, the number of functions, and such basis sets are known as double zeta (DZ), triple zeta (TZ), quadruple zeta (QZ), respectively. Most commonly used basis sets use a slightly modified arrangement where only the basis functions corresponding to the valence electrons of a given element are increased in zeta, leading to “split-valence” basis sets. Such basis sets may still be referred to as, for example, DZ, but the abbreviations VDZ and DZV (valence double zeta and double zeta valence, respectively) are also in use and provide additional clarity. Beyond increasing the zeta of a given basis set, its quality may also be improved by adding additional higher angular momentum functions beyond those required to contain the neutral atom electrons, for example, adding *d*-type functions (or higher) to a basis set for Ne. Functions of this type are termed “polarization functions” or “correlation functions” and interested readers are referred to Ref. [4] for a more detailed description of the physical interpretation of the effects of adding such functions to a basis set. Suffice it to say that they are incredibly important for investigations using electron correlation.

Although the contraction of basis functions according to Eq. (2) is a simple linear combination, the design decisions undertaken when developing new basis sets has led to the emergence of two distinct methods of selecting which GTOs are contracted and the pattern for doing so. The first of these is the “segmented contraction,” where each primitive contributes to only one contraction. This can lead to the situation where several primitives are cloned to appear in a number of differ-

ent contractions. The second method is known as “general contraction,”^[11] and in this case all primitives of the same angular momentum are present in every contracted function of the same angular momentum. A given primitive will then have a different contraction coefficient in each contracted function. For either type of contraction, it is also common to include a number of primitives in their uncontracted form (essentially a single primitive with a contraction coefficient of 1). When listing a basis set composition, both types of contraction are expressed in the style (12s6p3d2f1g) → [5s4p3d2f1g], where the parenthesis denotes the primitives and the square brackets the contracted functions. It is important to note that this provides almost zero information about the contraction pattern, and to uncover the details one must usually refer to some list of exponents and contraction coefficients or the original publication. Although most modern quantum chemistry codes support both segmented and generally contracted basis sets, the algorithms used are usually significantly more efficient for one contraction scheme than the other. A user should use such information when deciding on which basis set or software package to use for a given investigation.

With the terminology of basis sets in place, the development of several families of modern basis sets can now be discussed. For earlier developments and further history of Gaussian basis sets, there are a number of well-written reviews, such as Ref. [12].

Atomic Natural Orbital Basis Sets

Almost every modern basis set draws at least some inspiration from the atomic natural orbital (ANO) basis sets of Almlöf and Taylor.^[13–15] One of the major goals of this seminal advance was to reduce the contraction error (defined as the energy difference between calculations with an uncontracted primitive basis and those using a contracted version) in post-Hartree-Fock (HF) calculations. The approach taken was based on the primitive *sp* basis sets of van Duijneveldt,^[16] which were subsequently generally contracted in a way that produced HF contraction errors to within “chemical-accuracy” of 1 kcal mol^{−1}. Configuration interaction with singles and doubles (CISD) calculations were carried out using the uncontracted primitives to produce ANOs, with the expansion coefficients of these NOs then used as the contraction coefficients within the general contraction of the basis. Even-tempered polarization functions were added to the *sp* set and contracted in the same fashion, producing a family of basis sets for the first and second row atoms (and hydrogen). To address the transition

Table 1. Compositions of selected contracted Gaussian basis sets

Basis set	Elements		
	H	B—Ne	Al—Ar
ANO-L	[4s3p2d]	[6s5p3d2f]	[7s6p4d3f]
ANO-S	[4s3p]	[4s3p2d]	[5s4p3d]
ANO-RCC	[6s4p3d1f]	[8s7p4d3f2g]	[9s9p5d3f2g]
cc-pVDZ	[2s1p]	[3s2p1d]	[4s3p1d]
cc-pVTZ	[3s2p1d]	[4s3p2d1f]	[5s4p2d1f]
cc-pVQZ	[4s3p2d1f]	[5s4p3d2f1g]	[6s5p3d2f1g]
cc-pV5Z	[5s4p3d2f1g]	[6s5p4d3f2g1h]	[7s6p4d3f2g1h]
cc-pV(T + d)Z	[3s2p1d]	[4s3p2d1f]	[5s4p3d1f]
aug-cc-pVTZ	[4s3p2d]	[5s4p3d2f]	[6s5p3d2f]
cc-pwCVTZ	[3s2p1d]	[6s5p3d1f]	[7s6p4d2f]
cc-pVTZ-F12	[4s3p1d]	[6s6p3d2f]	[7s7p4d2f]
cc-pCVTZ-F12	[4s3p1d]	[7s7p4d2f]	[8s8p5d3f]
pc-0	[2s]	[3s2p]	[4s3p]
pc-1	[2s1p]	[3s2p1d]	[4s3p1d]
pc-2	[3s2p1d]	[4s3p2d1f]	[5s4p2d1f]
pc-3	[5s4p2d1f]	[6s5p4d2f1g]	[6s5p4d2f1g]
pc-4	[7s6p3d2f1g]	[8s7p6d3f2g1h]	[7s6p6d3f2g1h]
aug-pc-2	[4s3p2d]	[5s4p3d2f]	[6s5p3d2f]
def2-SVP	[2s1p]	[3s2p1d]	[4s3p1d]
def2-TZVPP	[3s2p1d]	[5s3p2d1f]	[5s5p3d1f]
def2-QZVPP	[4s3p2d1f]	[7s4p3d2f1g]	[9s6p4d2f1g]
def2-TZVPPD	[3s3p1d]	[6s3p3d1f]	[6s5p4d1f]
6-31G	[2s]	[3s2p]	[4s3p]

Where bases have multiple contractions available the set with the greatest number of contracted functions has been selected.

metal elements, the above procedure was modified such that density matrices were obtained for several electronic states, which were then averaged to produce the natural orbitals used in the contraction.^[17,18]

The averaging of density matrices was also used in an alternative formulation of ANO basis sets (ANO-L) described by Roos and coworkers,^[19] in which the matrix and resulting orbitals are averaged over anions and cations, atoms in electric fields and several electronic states for each atom. This produced a balanced description of hydrogen, helium, and the first row elements for properties including ionization potentials, electron affinities (EAs), and polarizabilities. The same methodology was later used to produce revised ANO basis sets for the second row elements,^[20] and for the first row transition metal elements.^[21] The composition of the largest of the resulting bases (along with compositions of many of the other sets detailed in this review) are displayed in Table 1. Smaller ANO-type basis sets (ANO-S) were also developed by Roos and coworkers^[22] for the elements H—Kr, with the intention that they could be used for larger molecules.

A third iteration of ANO basis sets (ANO-RCC) have also been published, with the design goal of ensuring basis sets of the same quality for the whole periodic table of elements. The initial publication in this direction targeted the alkali and alkaline earth metals,^[23] introduced scalar relativistic effects via the use of a Douglas-Kroll-Hess (DKH) Hamiltonian,^[24,25] and substituted the CISD method with complete-active-space self-consistent field (CASSCF) and complete-active-space second-order perturbation theory (CASPT2).^[26,27] The same design principles have since been used in the development of ANO-RCC sets for the main group elements (groups 13–18),^[28] the

transition metal elements [including spin-orbit (SO) effects],^[29] and the actinides and lanthanides (again including SO effects).^[30,31] The ANO-S, ANO-L, and ANO-RCC basis sets are the standard basis sets within the MOLCAS software.^[32]

Recently, Neese and Valeev have presented a series of ANO basis sets that take many cues from the correlation consistent basis sets (see below). The initial publication covers the elements H—Ar,^[33] with the multireference averaged coupled-pair functional method used in the optimization.^[34] A number of variations in polarization functions and contractions were benchmarked in terms of HF energy, correlation energy, and total energy, before comparison with several basis sets from other families. Although the initial results with these ano-pVnZ ($n = D, T, Q$) bases appear promising, at the time of writing, it is unclear how well the basis sets perform for atomic/molecular properties and for relative energies.

Correlation Consistent Basis Sets

The large number of primitive functions in the original ANO basis sets, particularly with high angular momentum, and the associated computational cost of evaluating them posed a significant barrier to adoption. An alternative approach, originally developed for the first row atoms (B—Ne and H) by Dunning,^[35] has led to the correlation consistent (cc) family of basis sets. The sets are energy-optimized and use general contraction schemes for the functions describing the occupied orbitals (the HF wavefunction). The key advance of the cc basis sets came from inspecting the incremental change in correlation energy (at the CISD level) when additional uncontracted higher angular momentum primitives were added, with the striking results depicted in Figure 1. It can easily be seen that the polarization functions can be separated into discrete groups based on the magnitude of correlation energy they recover. For example, adding a single *d* function has by far the largest effect on the energy, but adding a second *d* function has approximately the same effect as adding a single *f*

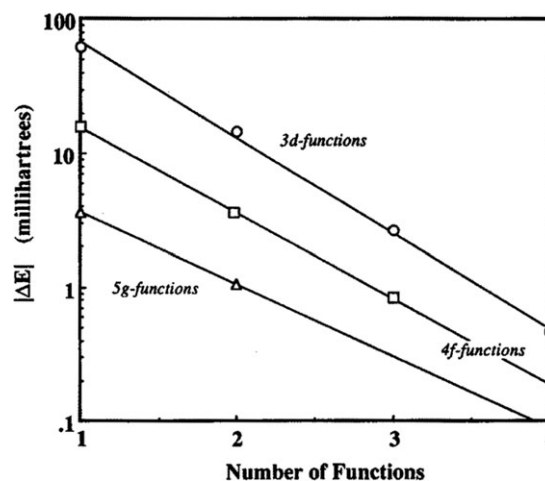


Figure 1. Incremental CISD correlation energy contributions for the oxygen atom, as used in the design of the correlation consistent basis sets. Reprinted from Ref. [35] with permission. Copyright 1989, American Institute of Physics.

function, forming a second grouping. Likewise, a third grouping is formed by the third *d* function, second *f* function and first *g* function. The naming of the basis set family now becomes somewhat obvious, each function within a grouping adds a (approximately) consistent amount of correlation energy. Combining these polarization groupings with split-valence DZ, TZ, and QZ quality groups of *s* and *p* functions (plus a number of *s* and *p* polarization functions), naturally led to basis sets that are known as cc-pVDZ, cc-pVTZ, and cc-pVQZ (correlation consistent polarized valence *n* zeta).

One of the primary reasons for the cc basis set family's lasting popularity is due to a series of empirical observations that as the cardinal number (*n* in cc-pV*n*Z) of the basis set is increased, energies and various properties converge smoothly toward the complete basis set (CBS) limit. This provides an established route for the systematic improvement of *ab initio* calculations toward the exact solution of the time-independent nonrelativistic electronic Schrödinger equation. Just as one may improve a coupled cluster operator from single and double excitations to single, double, and triple excitations,^[36] the basis set can be improved from, for example, cc-pVTZ to cc-pVQZ. This systematic improvement with respect to cc basis sets paved the way for a large number of extrapolation formulas that can be used to estimate the CBS limit. A review of extrapolation procedures is beyond the scope of the current work, and interested readers are referred to Ref. [37] and references therein.

The same principles used in the original cc investigation have since been used to extend the prescription up to cc-pV5Z or cc-pV6Z for H–Ne (and even beyond, such as cc-pV10Z for Ne),^[38–41] and to produce analogous basis sets for the second row elements (Na–Ar),^[41–43] and third row main group elements (Ga–Kr).^[44] A number of years after the initial publications, it was observed that adding a single additional “tight” (large exponent) *d*-function and reoptimizing the *d* primitives produced significantly better bond lengths and dissociation energies for molecules involving the second row elements. The resulting basis sets are denoted cc-pV(*n* + *d*)Z and should be used in place of the original sets for those atoms.^[41,45]

Extensions to the cc basis sets

In addition to the systematic improvement toward the CBS limit, cc basis sets have proven popular due to the large number of “extensions” available that adapt the original scheme to heavier elements and augment the sets with primitives for describing properties of interest. Many of these extensions were reviewed by Peterson in 2007,^[46] hence they will only be mentioned briefly here, along with descriptions of more recent developments.

Diffuse functions. An accurate description of EAs and noncovalent interactions requires “diffuse” (small exponent) functions within the basis set to satisfy the long-range part of the wavefunction. The cc philosophy to address such situations is to augment a “parent” set with an additional diffuse function for each angular momentum symmetry occupied within that basis.^[47] These functions are typically optimized for atomic anions and the resulting set prefixed aug-, for example, aug-cc-pVDZ. When cc basis sets were developed for heavier ele-

ments, it became commonplace that diffuse augmenting functions were published alongside their parent and, in situations where even more diffuse functions are required, doubly (d-aug-cc-pV*n*Z) and triply (t-aug-cc-pV*n*Z) augmented basis sets are available.^[38] Additional augmentation can be produced in an *ad hoc* fashion using the even-tempered scheme of Eq. (3).

Core-valence correlation. By default, most *ab initio* programs use a frozen-core approximation for correlated methods and all of the parent cc basis sets were designed with this in mind. For high-accuracy studies, it is frequently important to correlate at least an additional shell of core electrons, which requires a number of tight functions optimized for this purpose. There are two established approaches for optimizing such exponents; in the cc-pCV*n*Z basis sets, the exponents are optimized based on the difference between valence-only and core and valence correlation energies,^[48–51] whereas in the cc-pwCV*n*Z bases, the core-valence contribution to the correlation energy is strongly weighted over the core-core (hence the *w* in the basis set abbreviation).^[49,50] Typically, the cc-pwCV*n*Z sets produce better results for small zeta, with the results from both approaches converging rapidly for larger basis sets.

Scalar relativistic effects. For the lighter elements of the periodic table scalar relativistic effects are typically small, but, in very-high accuracy calculations where, for example, wavenumber accuracy is the goal, they should still be included. Perhaps the easiest way of doing so is to use a DKH Hamiltonian, and a comprehensive study has indicated that simply using the standard cc basis sets in such calculations can lead to large errors.^[52] The same study went on to show that recontracting the basis sets (keeping all exponents the same) in atomic DKH calculations was sufficient to produce good results for the elements H–He, B–Ne, Al–Ar, and Ga–Kr,^[52] with a later study developing similar sets (denoted cc-pV*n*Z-DK) for Li, Be, Na, and Mg.^[41] Alternatives to DKH calculations include fully-relativistic 4-component Dirac–Hartree–Fock calculations (fully-relativistic herein) or those including relativistic pseudopotentials [PPs, also known as effective core potentials (ECPs)], which shall be discussed in the context of the heavier elements below.

Heavier elements. All the cc basis sets described above share a common methodology in their design and optimization, yet challenges posed by heavier elements such as greater relativistic effects, a proliferation of common oxidation states, and a simple increase in the total number of electrons have necessitated some changes to the approach. The use of PPs remedies two of these problems as they accurately include scalar relativistic effects in a simple fashion, and they replace a number of core electrons, reducing the size of the basis set. For further details on the design and constructions of PPs, the reader is referred to a recent review,^[53] but for present purposes it is important to note two different sizes of PP: large-core (replacing all but the valence electrons), and small-core (an extra core-valence shell of electrons remain). In the latter, the primary goal is an accurate description of relativistic and

correlation effects, rather than decreasing the overall computational cost.

Some of the first cc basis sets matched to PPs were produced for the elements Ga–Kr and In–Xe, at TZ and QZ quality.^[54] These sets are based on the large-core Stuttgart–Dresden–Bonn PPs,^[55] and are denoted SDB-cc-pVnZ. Subsequently, cc basis sets from DZ to 5Z quality have been developed that are matched to the small-core Stuttgart–Köln PPs.^[56–58] These basis sets span the whole post-*d* main group elements, are denoted cc-pVnZ-PP, and have diffuse augmented and core-valence variants available.^[59] As had been suggested previously,^[54,60] calibration of the core-valence effect confirmed that correlation of the $(n - 1)d$ electrons is very important for systems containing the group 13–15 elements, with the effects becoming larger further down the groups. Correlating the $(n - 1)s$ and *p* electrons is much less important, but still advisable for high-accuracy studies.^[59] A revised PP and accompanying cc bases have been published for I, and due to the increase in accuracy should be used in preference to the originals.^[61] In terms of basis sets for fully relativistic calculations on post-*d* main group elements, cc sets of DZ–QZ quality are available,^[62–65] along with additional functions for core correlation.^[66] Dyal has also recently published fully relativistic sets for the 7*p* elements, managing to overcome the problem of very large SO coupling in the process.^[67]

High-accuracy calculations involving transition metal elements are demanding on both method and basis set due to the likelihood of a large number of low lying electronic states, and care must be taken to ensure a basis set provides a balanced description of all these states (as seen in the ANO basis sets where an averaged density matrix was used). The first cc basis sets for transition metals were published for Ti and Fe,^[68,69] where the approach chosen was to contract the primitive functions developed by Partridge,^[70] before optimizing polarization functions. All-electron basis sets for all of the 3*d* elements were later produced by Balabanov and Peterson,^[71] with care taken to ensure the sets were flexible enough to describe the 4*p* orbitals and to average the primitive *d* functions over (up to) three electronic states. Although only TZ–5Z sets were described in the original publication, DZ sets developed using the same strategy are available,^[72] and included in packages such as MOLPRO.^[73,74] The same publication also detailed cc sets for use in DKH calculations, diffuse augmented bases, and core-valence sets. The latter differ slightly from the cc-pwCVnZ bases for lighter elements as a number of the valence-only exponents were reoptimized in addition to augmenting with a number of extra functions. This change of approach was necessitated by significant overlap between valence and core-valence functions in initial test optimizations.

Relativistic effects can no longer be neglected once the 4*d* and 5*d* transition metal elements are reached, hence cc basis sets have been developed that are matched to the small-core Stuttgart–Köln relativistic PPs for those elements and Cu and Zn.^[75–78] As the reader may have already surmised, aug- and cc-pwCVnZ variants were introduced simultaneously with their parent sets (DZ–5Z), but these investigations also optimized all-electron DKH basis sets at the TZ level for the purposes of

quantifying the error introduced by using the PP approximation. These errors were shown to be negligible, except when core-valence calculations were carried out on electronic excited states. Basis sets for fully relativistic calculations (DZ–QZ) are also available for the 4*d* and 5*d* elements.^[79,80]

The heavier alkali and alkaline earth elements are one of the areas of the periodic table that are not covered particularly well by the cc family of basis sets. For calcium, cc-pVnZ and cc-pCVnZ (*n* = D–5) sets have been published,^[81] and it has been established that correlation of the 3*sp* outer core electrons is important for a number of calcium containing molecules.^[81,82] However, fully relativistic sets are available for 4*s*, 5*s*, 6*s*, and 7*s* elements,^[83] along with the lanthanides and actinides.^[84,85]

Basis sets for explicit correlation. Over the course of the last decade, there has been a great deal of excitement in the theoretical chemistry community surrounding explicitly correlated wavefunctions (for recent reviews, see Refs. [86–89]). By incorporating functions depending explicitly on the interelectronic distance into the wavefunction, the convergence of energies and properties toward the basis set limit is greatly accelerated. In many cases, explicitly correlated methods with TZ quality basis sets produce results comparable to 5Z sets in conventional methods.^[90] Although using “standard” diffuse augmented correlation consistent basis sets (aug-cc-pVnZ) in explicitly correlated calculations produces good results, these methods also offer an attractive proposition from a basis set development point of view. In conventional post-HF calculations, the basis set must be able to describe the HF wavefunction, the correlation hole (associated with the Coulomb cusp), and the longer-range component of the wavefunction. The additional correlation factor in explicitly correlated methods is a two-electron function and thus is able to accurately describe the cusp region, allowing the one-particle basis set to focus on the HF and longer-range components.

The approach of Peterson and coworkers to producing basis sets for use with explicitly correlated methods mirrors the original cc work of Dunning as correlation consistent groupings were determined by plotting the incremental contribution of individual functions to the total correlation energy at the explicitly correlated second-order Møller–Plesset perturbation theory level (MP2-F12,^[91,92] where F12 denotes a nonlinear correlation factor).^[93] A comparison of the incremental correlation energy contributions at the MP2 and MP2-F12 levels made it obvious that higher angular momentum functions are much less important in explicitly correlated calculations, meaning only DZ–QZ sets were developed. At the time of the original investigation, open-shell explicitly correlated codes were not commonly available, hence the exponents for most elements were determined by molecular calculations that covered a number of bonding environments. The choice of HF and *s* and *p* polarization functions was based on the HF functions from the cc-pV(*n* + 1)Z basis sets along with diffuse augmenting functions from the corresponding aug-cc-pV(*n* + 1)Z set. At first glance, this leads to a surprising result, the composition of the new cc-pVTZ-F12 basis set for Ne is [6*s*6*p*3*d*2*f*],

while the aug-cc-pVTZ composition is [5s4p3d2f]. However, the reasoning behind this minor increase in basis set size is very sound. Contrary to the situation in conventional post-HF calculations, for F12 methods the basis set error in the total energy is often dominated by the error in the HF energy. The increase in the number of *s* and *p* functions thus reduces the HF error with little increase in computational cost.

A comparison of the MP2-F12 correlation energies for Ne, Ar, N₂, and P₂ demonstrated that the new cc-pVnZ-F12 basis sets produce results roughly equivalent to those with an aug-cc-pV(*n* + 1)Z basis (it is important to note that an F12 method is used in both cases), and benchmark calculations on a larger set of relative total energies indicate similar performance.^[93] Systematic convergence toward the CBS limit is again a feature of these basis sets and an extrapolation procedure has been reported for use in high-accuracy studies of, for example, thermochemistry and *ab initio* spectroscopy.^[94] The F12 basis sets were originally developed for the elements H, He, B–Ne, and Al–Ar and have since been extended to cover Li, Be, Na, and Mg.^[95] Although the cc-pVnZ-F12 basis sets outperform the equivalent aug-cc-pVnZ sets in most cases, the latter do exhibit a more rapid convergence for the interaction energies of molecules bound by noncovalent interactions.^[96] This is due to the presence of additional diffuse higher angular momentum functions in the aug-sets; however, for the same reasons, the cc-pVnZ-F12 sets suffer from significantly less basis set superposition error and may be preferred when counterpoise correction calculations are not carried out.^[96,97]

Correlation consistent basis sets for accurately describing core-core and core-valence correlation in an F12 framework have also been published for the elements Li–Ar.^[51,95] Two alternatives were proposed and denoted cc-pCVnZ-F12 and aug-cc-pC_{F12}VnZ, based on adding tight functions to the cc-pVnZ-F12 and aug-cc-pVnZ basis sets, respectively. The number of additional functions added to construct the cc-pCVnZ-F12 sets is very small, especially at the QZ level, where only [1s1p1d] are added (for the first row elements), compared to [3s3p2d1f] in aug-cc-pCVnZ. A subsequent benchmarking of these new sets for the calculation of equilibrium geometries, dissociation energies, and harmonic frequencies of diatomic molecules at the coupled cluster level indicates that when using explicitly correlated methods the core-core and core-valence effects are effectively converged at the QZ level and that the cc-pCVnZ-F12 sets outperform alternative bases. For more pragmatic studies, the core-valence effect is reasonably well converged at the cc-pCVTZ-F12 level, with cc-pCVDZ-F12 producing good results for diatomic molecules containing second row elements.

Polarization Consistent Basis Sets

Most of the above review of the ANO and cc basis sets has been considered from a correlated *ab initio* perspective, where the scaling of computational effort with respect to basis set size is steep and convergence of the total energy is typically slow (inverse polynomial with respect to the highest angular momentum functions).^[98–101] However, the DFT (and HF)

methods converge significantly faster (exponential with respect to basis set size),^[102,103] and hence the basis set requirements are obviously different. It also bears mentioning that DFT correlates all electrons, but most post-HF methods and the basis sets designed for them use a frozen-core approximation. Although the rapid convergence with respect to basis set has led to a proliferation of DFT calculations carried out with the older Pople style basis sets, for example, 6-31G and extensions,^[104] it is still desirable to be able to systematically approach the CBS limit. Not only does this enable one to produce presumably more accurate energetics, but it also allows for an inspection of the intrinsic errors of a specific exchange-correlation functional. Basis sets that systematically approach the DFT (Kohn–Sham) and HF limit have been developed by Jensen and coworkers, and are denoted polarization consistent (pc).^[105,106]

Determining basis sets for molecular calculations using the HF method leads to a problem as the atomic HF energy does not depend on polarization functions. Jensen's solution to this was to minimize the HF energy for symmetric homonuclear molecules containing the elements H, C, N, O, and F, with atomic *s* and *p* functions (*s* only for H), thus ensuring that only the polarization functions may be biased toward the molecules in the optimization set. In a similar fashion to the cc basis sets, plotting the energy contribution from individual polarization functions for the N₂ molecule, on top of the (26s17p) atomic functions of Koga et al.,^[107] produces the pattern displayed in Figure 2. A comparison of Figures 1 and 2 shows that the composition of the first two groupings of (non *s* and *p*) polarization functions is the same for both the cc and pc schemes, yet the subsequent pc groupings contain slightly more *d* angular momentum symmetry functions than the cc groupings. For example, summing the first three pc groupings results in 4d2f1g functions, but 3d2f1g functions in the cc groupings. Figure 2 shows energy contributions for N₂ with internuclear separations of both 2.068 and 2.680 a.u., demonstrating that a change in geometry produces only a minor effect on the energies and does not alter the groupings. Analogous plots for C₂, N₄, O₂, O₃, and F₂ indicate little change in

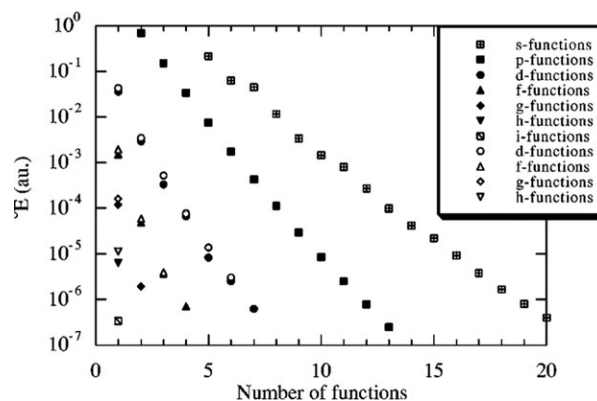


Figure 2. Incremental HF energy contributions for N₂ (filled symbols for a bond distance of 2.068 a.u., open symbols for 2.680 a.u.), as used in the design of the polarization consistent basis sets. Reprinted from Ref. [45] with permission. Copyright 2001, American Institute of Physics.

the optimal groupings for each molecule, allowing for a series of consistent groupings. The abbreviation *pc-n* is then used to designate the basis sets, where *n* indicates the level of polarization. For example, *pc-0* has no polarization functions, *pc-1* has one polarization function (1*d*), and *pc-4* has polarization functions of four angular momentum symmetry higher than that needed for the atom (6*d3f2g1h*). Although the polarization functions are left uncontracted, a number of the *s* and *p* functions are generally contracted using the atomic HF orbital coefficients as the contraction coefficients. Subsequently, the exponents and contraction coefficients were recalculated using the BLYP functional,^[108,109] resulting in slightly more diffuse *s* and *p* functions, and slightly tighter exponents for the polarization functions.^[110] The composition of all sets was fixed to those of the earlier HF based work.

Applications where diffuse functions are important, such as EAs, produced some interesting problems in the development of diffuse augmented *pc* basis sets (denoted *aug-pc-n*).^[111] Much of this is related to the self-interaction error in DFT (see Refs. [112] and [113], and references therein), which is particularly problematic for anionic systems. In the process of developing these new sets, Jensen examined how the BLYP EA of a number of systems changed as basis set size is increased and concluded that the good agreement between DZ and TZ basis set results and experiment is somewhat fortuitous. Pragmatically, such levels of theory can be thought of as “parameterized models” for the calculation of EAs. Analogous to the *aug-cc* sets, a single additional diffuse function was added to each angular momentum symmetry occupied in a given basis set, with the *s* and *p* exponents determined by a scaled even-tempered extension [a modified version of Eq. (3)], while the diffuse polarization exponents were produced by scaling the resulting diffuse *p* exponents. In some cases, the use of partially augmented basis sets can be advocated, such as in the calculation of EAs or for the largest *pc-4*-based sets. Although the DFT and HF methods preclude the need for core-valence basis sets, the *pc-n* family has been augmented with tight exponents, primarily for use in the calculation of spin-spin coupling constants.^[114] The resulting sets are denoted *pcJ-n*, and have been further improved by subsequent recontraction.^[115]

Following the same principles used in the initial investigations, *pc-n* and *aug-pc-n* basis sets have been developed for the elements Si–Cl.^[116] The number of polarization functions remains the same as for C–F, with small changes in the number of *s* and *p* functions. Just as for the lighter elements,^[105,110,117] it has been demonstrated that the *pc-n* family of basis sets for the second row elements can produce accurate atomization energies, equilibrium geometries, harmonic vibrational frequencies, and can easily be extrapolated to the CBS limit.^[116] Extensions of the *pc-n* family to the remaining elements in the first two rows of the periodic table (plus He),^[118] encountered some small problems for the alkali and alkaline earth elements due to the energetic groupings of the polarization functions being different to those of the *p*-block elements. Although a number of other basis set families ensure that the *s*-block elements of a given row have the same number of functions as the *p*-block, the driving force

behind the *pc-n* sets having a different composition is based on the basis sets producing a balanced description of energetic properties across an entire row. The *pc-n* basis sets have also recently been extended to include the 4*s* (K, Ca) and 4*p* (Ga–Kr) elements.^[119]

The benchmarking and calibration of the *pc-n* and *aug-pc-n* basis sets has shown that, for the HF and DFT methods, this family of basis sets typically outperforms other sets of a roughly equivalent size for atomization energies, dipole moments, ionization potentials, relative molecular energies, and equilibrium geometries. The *pc-3* and *pc-4* basis sets are somewhat larger than basis sets commonly used in DFT investigations, but *pc-2* produces results that are reasonably well converged and the errors associated with *pc-1* can be of the same order as the intrinsic error of functionals such as BLYP.^[117]

Def2 Basis Sets

The naming of the *def2* family of basis sets developed by Ahlrichs and co-workers is due to their inclusion in the TURBO-MOLE program,^[120,121] where a *def* (for default) prefix was added to some basis sets to prevent confusion with earlier, similarly named sets. These more recent developments have established a family of basis sets that have a similar accuracy for (almost) all elements in the periodic table,^[122,123] which are now prefixed *def2* and should be used in preference over the *def* sets. Although they do not systematically approach the CBS limit in the same fashion as the *cc* or *pc-n* bases, the *def2* family has several qualities available: *def2-SV*, *def2-TZV*, and *def2-QZV*. There are two different sets of polarization functions available, P and PP (not to be confused with PPs), where the former is intended for HF and DFT calculations, and the latter for post-HF studies. These polarization prescriptions are added as a suffix, for example, *def2-QZVPP*. It should be noted that P and PP are identical for *p*-block elements.

The exponents describing the HF wavefunction are energy-optimized at the atomic HF level, along with segmented contraction coefficients. Many of the polarization functions are taken directly from the correlation consistent basis sets for *p*-block elements (and H, He), with those for other elements determined by a mixture of MP2 calculations and interpolation. For the elements Rb–Rn (excluding lanthanides), small-core relativistic PPs are used.^[56,58,124–127] Compared to the *cc* basis sets for the heavier *p*-block elements, the *def2-TZVP(P)* and *def2-QZVP(P)* bases contain additional tight *f* functions that produce significantly better results for complexes such as At₂, SbF₅, and SbF₃.^[122] Much of the *def2* family of basis sets is based on previous work by Ahlrichs and coworkers,^[128–132] with some redevelopment to ensure small-core ECPs are used consistently for the elements Rb–Rn and that a similar level of performance can be expected for all elements.

The *def2* family has recently been extended to the lanthanides,^[133] based on the segmented bases of Cao and Dolg and using small-core PPs.^[134,135] The majority of the primitives were energy optimized in HF calculations on the lowest 4*f*^{*n*} + 1*6s*² electronic state of a lanthanide, and augmented with HF

energy-optimized diffuse d and p functions for a $4f^0 5d^1 6s^2$ configuration. Six g functions were also added by fixing their exponents to be the same as the six smallest f functions, producing a (14s13p10d8f6g) set.^[136] The segmented contraction coefficients were then determined from an ANO-like scheme.^[135,136] The def2-TZVP(P) sets recontracted the p functions of Cao and Dolg, added additional diffuse p functions and reduced the number of g functions. The def2-SV(P) and def2-QZVP(P) sets were then constructed by removing and adding functions, including the addition of a single h function in the latter case. Calibration of the bases on a number of lanthanide complexes indicates that the basis set errors are comparable to those calculated previously for lighter elements,^[133] in-keeping with the overall philosophy of the def2 bases.

Two different schemes for the partial augmentation of the def2 bases with diffuse functions have been presented by Rapoport and Furche,^[137] and Zheng and coworkers.^[138] The former (denoted by a suffix D, e.g., def2-TZVPPD) takes a property-optimized approach to ensure the resulting sets are suitable for the calculation of response properties. By including only the minimal number of diffuse functions needed to obtain the desired accuracy in dipole polarizabilities, these bases avoid some of the problems associated with diffuse functions such as ill-conditioned overlap matrices and reducing the effectiveness of integral prescreening.^[137] The approach of Zheng et al. was to simply add additional s and p exponents equal to $\zeta_1/3$, where ζ_1 is the most diffuse (s or p) exponent in the parent def2 set. These minimally augmented (ma-) sets perform very well for the calculation of EAs with DFT methods, although little improvement is observed when evaluating interaction energies for noncovalent bonds with DFT methods and the ma-def2 basis sets.^[138]

Other Orbital Basis Sets

There are a plethora of Gaussian basis sets available, to the point where attempting to provide a broad coverage of all of them would be monumental task. These bases range from those optimized for a single atom and a single purpose to much more general and broadly applicable. A small number of these basis sets are briefly described below, yet no conclusions on quality or purpose should be inferred from the absence or presence of any particular bases.

Pantazis and Neese have developed a series of segmented all-electron relativistic contracted (SARC) basis sets, primarily for investigating properties that depend on core electrons, such as electron paramagnetic resonance, where the use of PPs is not advisable. The small size of these bases ensures they are competitive with PP sets in terms of computational efficiency, even though they require a relativistic Hamiltonian such as DKH or zeroth-order regular approximation (ZORA).^[139–141] The SARC primitives are extrapolated based on CASSCF results, before contraction based on scalar relativistic CASSCF calculations. This procedure was carried out for the $5d$ transition metal elements,^[142] the lanthanides,^[143] and the actinides,^[144] with a relativistic recontraction of the def2 basis sets performed for the elements of the first three rows of the

periodic table to ensure consistency with the SARC bases.^[142] DFT-based ZORA and DKH results using the SARC bases are in good agreement with experimental or high level *ab initio* results for ionization potentials, geometries, and bond dissociation energies.

The natural orbital-based segmented contracted Gaussian (NOSeC) basis sets consist of polarization exponents and contraction coefficients determined by minimizing deviation from accurate NOs generated by atomic CISD calculations. The functional used in this minimization is

$$\Delta = \sum_{k=1}^N v_k \int |\xi_k(\mathbf{r}) - \psi_k(\mathbf{r})|^2 w(\mathbf{r}) d\mathbf{r} \quad (4)$$

where ψ_k is the familiar linear combination of atomic orbitals (contracted basis functions), ξ_k are the reference NOs with occupation numbers v_k , and w is a weight function. These polarization functions can be combined with an arbitrary HF basis set and have been reported for most of the periodic table,^[145–149] including sets for core-valence correlation,^[150] and relativistic effects using a DKH Hamiltonian.^[151–155] The NOSeC bases typically recover more than 99% of the correlation energy calculated using the parent NOs, but are significantly more computationally efficient. Where necessary, such as the first row transition metal elements, the reference NOs were averaged over different electronic configurations, and can be combined with PPs for the heavier elements. For elements with a single valence electron, such as Li, a common approach is to optimize exponents for polarization functions using the homonuclear dimer.^[41] However, the approach taken for the NOSeC bases uses the virtual K orbital method of Feller and Davidson to produce reference orbitals rather than CISD NOs. These orbitals are produced by diagonalizing the G operator of Eq. (5) over the virtual space:

$$G = \sum_j K_j - \alpha F \quad (5)$$

where K_j is the exchange integral operator for the occupied orbital j , α is based on the virtual orbital occupation of an important orbital, and F is the Fock operator.^[156] The resulting correlating orbitals can be loosely compared to a molecular situation where an additional electron is provided by an adjacent atom.^[149] Recently, the Sapporo- n ZP ($n = D, T, Q$) have been developed using a similar fitting to reference NOs approach.^[157,158] Unlike the NOSeC bases, Sapporo- n ZP specify an underlying HF set and core correlating functions are available for all elements (with the obvious exceptions of H and He). For the heavier elements, relativistic effects are considered through a DKH Hamiltonian.

Many of basis sets for use with DKH Hamiltonians are simply the same exponents taken from the nonrelativistic sets recontracted with a relativistic treatment. In contrast, Hirao and coworkers have developed Gaussian basis sets specifically for use in third-order DKH calculations.^[159,160] The exponents were optimized by minimizing atomic HF energies (with a DKH Hamiltonian), with functions added until the incremental

energy change was less than 1 m E_h in each angular momentum symmetry. This approach ensures equal quality for all elements, at the expense of requiring a large number of primitives for heavier elements. To enable their use in molecular calculations, the primitives were generally contracted, and augmented with polarization functions and a number of even-tempered diffuse functions.

Auxiliary Basis Sets

The evaluation of electron repulsion integrals is often a bottleneck in both DFT and *ab initio* calculations. Focusing momentarily on pure DFT, that is, without exact exchange, the Coulomb (J) term can be separated from the exchange-correlation term as:

$$J = \frac{1}{2} \int \rho(r_1) \frac{1}{r_{12}} \rho(r_2) d\tau \quad (6)$$

where $\rho(r)$ is the molecular electron density. By expanding these molecular electron densities in an atom-centered auxiliary basis set (ABS), it is possible to bypass the evaluation of four-center two-electron integrals and hence reduce the computational effort of pure DFT calculations by up to an order of magnitude with negligible loss of accuracy. Due to the expressions resembling a resolution of the identity (RI), this method is often referred to as RI-J. The expansion of the electron density takes the form:

$$\rho(r) \approx \sum_{\alpha} c_{\alpha} \alpha(r) \quad (7)$$

where α represents a function in the ABS and c_{α} is an expansion coefficient.^[161–168] The approximation used means that the ABS used in Eq. (7) depends on the chosen orbital basis set [orbital basis set (OBS) will be used herein to reduce confusion between auxiliary and orbital bases], and the expansion coefficients are determined by minimizing

$$d_{\rho} = (\rho - \rho_{\text{RI-J}} | \rho - \rho_{\text{RI-J}}), \quad (8)$$

where $\rho_{\text{RI-J}}$ denotes the density approximated in Eq. (7).^[165,168] This leads to a set of linear equations expressed as:

$$\sum_{\beta} c_{\beta} (\alpha | \beta) = \sum_{\nu\mu} (\alpha | \nu\mu) D_{\nu\mu}, \quad (9)$$

where ν and μ are functions in the OBS and D denotes the density matrix. It can be shown that Eq. (8) is equivalent to

$$(\nu\mu | \kappa\tau) = \sum_{\alpha\beta} (\nu\mu | \alpha) (\alpha | \beta)^{-1} (\beta | \kappa\tau), \quad (10)$$

which appears similar to the insertion of a RI.^[168] This least-squares fitting and choice of Eq. (8) to specify the metric is not the only route to approximating J ,^[167,169–171] but has proven to be the most popular. Two methods of determining the ABS exponents specifically matched to an OBS have thus emerged; explicit optimization of ABS exponents,^[168] and algo-

ritms for the automatic generation of ABSs.^[172,173] Although the latter holds the advantage that no ABS development is required to fit new OBSs, the ABSs resulting from explicit optimization typically contain fewer functions for a given level of accuracy and are thus computationally more efficient.

The first systematic investigation into optimized ABSs was carried out by Ahlrichs and coworkers and produced bases specifically matched to the def-SVP OBS.^[168] The resulting ABSs were optimized such that difference between the fitted and conventional densities (at the HF level to prevent any bias towards a particular functional) were effectively minimized and the functions were contracted using a segmented pattern, with this work later complimented by ABSs matched to def-TZVP.^[174] Through fitting to def2-QZVPP OBSs and keeping the number of functions as small as possible, Weigend^[175] superseded this previous work by developing a single RI-J ABS (for the elements H–Rn) that is suitable for use with any of the def2 family of OBSs, showing negligible errors for atomization energies, dipole moments, geometries, and vibrational frequencies.

The reduction in the computational cost of evaluating the Coulomb integrals has promoted the wide adoption of RI-J fitting when pure DFT functionals are used. However, in hybrid DFT methods, where some amount of exact HF exchange is included, evaluating the exchange scales as ON^2M , where N and M are the number of functions in the OBS and ABS, respectively, and O is the number of occupied orbitals. This means that for systems with a large number of electrons, such as medium sized transition metal complexes, the fitting has no advantage over conventional integral evaluation algorithms. Some methods that alleviate the exchange fitting problem have been proposed, and will be discussed later.

RI-like approximations have also been used to improve the efficiency of methods such as HF,^[176–179] MP2,^[180] coupled cluster with singles and doubles (CCSD),^[181] and approximate coupled cluster singles-and-doubles model (CC2).^[182] Herein, the approximation will be referred to as density fitting (DF) due to the reasons detailed in Ref. [183], although several groups and program packages still use an RI prefix. It is important to note that the density refers to products of basis functions rather than the electronic density, and the DF moniker also eases potential confusion when discussing ABSs in the context of explicitly correlated methods. In addition to the Coulomb term, applying the DF approximation to HF theory requires fitting of the exchange part of the Fock matrix,^[178] and, as mentioned above in the context of hybrid DFT, the efficiency of the resulting algorithm is related to the ratio between the number of functions in the basis sets and the number of occupied orbitals. This indicates that DF-HF is best applied to calculations using relatively large basis sets on small to medium sized molecules and, in practice, is most often used in conjunction with DF-MP2. Compared to RI-J, where only the total electron density is approximated, DF-HF requires the fitting of products of individual basis functions and hence a larger number of auxiliary functions are required. The additional flexibility this affords means that ABSs optimized for exchange fitting are also appropriate for DF of the Coulomb

part.^[184] ABSs for use in DF-HF are thus often termed “JK-fit” and have been matched to def-TZVPP, cc-pVTZ, cc-pVQZ, and cc-pV5Z for the elements H, B–F, Al–Cl, and Ga–Br.^[178] A series of “universal” JK-fit sets for the elements H–Rn have also been published and although originally optimized for the def2 family of OBSs,^[133,184] they have also seen significant use when matched to the cc bases. DF-HF can be further simplified by using the Poisson method, and even-tempered ABSs for this purpose have been published for the elements H and C–F.^[179]

The overall goal of DF-MP2 is to ensure that the difference between the conventional MP2 correlation energy and that from DF-MP2 is as small as possible. However, minimizing such a quantity does not lead to optimal ABSs and in practice the optimization proceeds by minimizing the quantity

$$\delta\text{DF} = \frac{1}{4} \sum_{a|bj} \frac{[\langle ab||ij \rangle_{\text{DF}} - \langle ab||ij \rangle]^2}{\epsilon_a - \epsilon_i + \epsilon_b - \epsilon_j}, \quad (11)$$

where $\langle ab||ij \rangle = (a|b)(i|j) - (a|j)(b|i)$, with i, j denoting occupied orbitals, a, b virtual orbitals, and ϵ_x the HF orbital energies.^[131] Unlike the energy difference, δDF will always be positive and the fact that it only depends on the ABS functions (for a given OBS) makes it desirable from an optimization viewpoint. It has been demonstrated that accurate DF at the MP2 level requires the ABS to include functions with an angular momentum of at least $\ell_{\text{occ}} + \ell_{\text{OBS}}$, where ℓ_{occ} is the highest angular momentum occupied in the HF wavefunction and ℓ_{OBS} the highest angular momentum present in the OBS. This can result in the case where ℓ_{ABS} is greater than what is supported in some quantum chemical codes, despite ℓ_{occ} being within the allowed range. Unsurprisingly, simply truncating the ABS to the maximum supported ℓ results in large fitting errors,^[185] but there is some evidence that this error is somewhat constant and cancels for relative energies.^[186]

The development of ABSs for DF-MP2 was greatly streamlined by the introduction of a driver program that uses analytic gradients for the optimization of ABSs within the RICC2 module of the TURBOMOLE package.^[121,187] This means that MP2 fitting ABSs (often referred to as “MP2-fit”) have been produced that are specifically matched to correlation consistent and def2 OBSs for most of the periodic table,^[95,185,187–193] with the advent of explicitly correlated and local electron correlation methods (that use DF) a strong driving force for this development. The resulting ABSs typically contain between two and four times as many functions as the OBS they are matched to, and the fitting errors are two to four orders of magnitude smaller than the incompleteness errors in the OBS. To ensure a balanced description of the different oxidation states of transition metal elements, a number of ABS exponents are optimized for the cations corresponding to the highest oxidation state commonly found in complexes, along with some exponents optimized for intermediate oxidation states.^[187] Calibration of the resulting ABS on small to medium sized complexes with a variety of oxidation states indicates that negligible fitting errors are produced.^[185,187,190–192]

An alternative to DF using preoptimized ABSs are methods based on Cholesky decomposition (CD) techniques, and

although such methods have been the subject of recent reviews,^[194,195] they are briefly recapped below. A symmetric positive definite matrix \mathbf{V} can be decomposed into the product of a lower triangular matrix, \mathbf{L} , and its adjoint:

$$\mathbf{V} = \mathbf{L}\mathbf{L}^\dagger, \quad (12)$$

and when CDs are applied in a quantum chemical context, the matrix \mathbf{V} is usually taken to be that of the two-electron integrals, $V_{ij,kl} = (ij|kl)$.^[196] This allows the integrals to be expressed as

$$(ij|kl) = \sum_p^Q L_{ij}^p L_{kl}^p, \quad (13)$$

and in a similar fashion to Eq. (11), the error in the integral representation can be given as

$$D_{ij,kl} = (ij|kl) - (ij|kl)_{\text{CD}}. \quad (14)$$

Combining Eqs. (13) and (14) defines a residual matrix such that

$$(ij|kl) = \sum_{p=1}^Q L_{ij}^p L_{kl}^p + D_{ij,kl}^{(Q)}, \quad (15)$$

and this suggests the accuracy of the CD approximation can be controlled as

$$\max(D_{ij,kl}^{(Q)}) \leq \delta. \quad (16)$$

δ is then known as the decomposition threshold, with all integrals being reproduced to at least the value of δ via a recursive procedure. Beebe and Linderberg demonstrated that CD of two-electron integrals can also be thought of as the determination of a linearly dependent set of atomic orbital product functions,^[196] which are commonly referred to as the Cholesky basis, and it is very important to note that using a Cholesky basis as an ABS makes CD and DF formally equivalent.

Early CD implementations were significantly more expensive than DF as they still required four-center two-electron integrals to be evaluated, and the Cholesky basis had a strong dependence on molecular geometry. The one-center CD (1C-CD) approach improved computational efficiency by restricting the product functions within the Cholesky basis to be located on a single center, but retained the strong dependence on geometry.^[173] A significant advance in the application of CD to quantum chemical methods was the introduction of atomic CD (aCD), where the Cholesky basis is produced on-the-fly from CD of the atomic two-electron integral matrix for a given OBS, along with the addition of missing angular components.^[173] A hierarchy of aCD ABSs were generated by varying the threshold $\delta = 10^{-n}$, with the ABSs then denoted aCD- n^* . With $\delta = 10^{-4}$,^[173,197] these ABSs produce accurate results for a number of methods, including DFT (pure and hybrid), HF and MP2, but the number of functions in the set is often too large to make aCD truly competitive with preoptimized DF sets.^[198]

To reduce the total number of functions, the atomic compact CD (acCD) procedure performs a second CD on the aCD basis, before the remaining functions are fit to produce an accurate representation of that parent aCD set.^[199] For a large iron containing complex, this led to a 70% reduction in the number of ABS primitives, but with no loss of accuracy when compared to aCD. Although it is possible to generate acCD ABSs and store them for later use in a basis set library, in practice this is not carried out as the computational cost for on-the-fly generation is negligible. In addition to removing the need for preoptimization of ABSs, general applicability to a number of quantum chemical methods is a major strength of the CD approach, with recent applications to CASSCF and CASPT2 methods.^[200,201] The number of functions in an acCD ABS is still considerably larger than in pre-optimized ABSs, with a study on a test set of noncovalent interactions using TZ OBSs indicating that acCD-4 is around four times larger than the equivalent MP2-fit ABS, although, perhaps reassuring given the significantly larger number of functions, smaller errors relative to conventional methods were produced.^[202]

As mentioned above, fitting of the exchange integrals proves problematic when the number of occupied orbitals or size of the OBS becomes large, limiting the size of problem that may be tackled with DF-HF or DF of hybrid DFT methods. Two different solutions to this problem have emerged, both based on locality. The local DF-HF approach of Polly et al.^[179,183] localizes orbitals and then restricts the fitting basis for any occupied orbital to a spatially close "domain," which is asymptotically independent of system size and leads to linear scaling for evaluating the exchange matrix. The local K (or LK) method of Aquilante et al.^[203] also uses localized orbitals, but uses CD techniques and integral screening to achieve a similar result without variable domain definitions.

A large number of explicitly correlated methods, such as MP2-F12 and approximate CCSD(T)-F12,^[204–207] use DF in the construction of the Fock and exchange matrices (using JK-fit sets), and in the evaluation of the remaining two-electron integrals (using MP2-fit sets).^[208] But, the three- and four-electron integrals necessary to explicitly include interelectronic distances in the wavefunction can also be tackled in a superficially similar manner. Inserting an RI approximation allows these multielectron integrals to be simplified into products of no more than two-electron integrals,^[209–211] with a separate RI ABS used to allow smaller OBSs to be used.^[212] The complementary ABS (CABS) approach involves constructing the union of the OBS and ABS and the reduction in RI error (due to using an incomplete RI basis) from this approach means it has become something of a standard in production level F12 codes.^[213] A secondary benefit of the CABS approach is that the error in the HF energy can be significantly reduced by including single excitations into complementary auxiliary orbital space.^[205]

Peterson and coworkers have developed a series of ABSs specifically for use in the RI approximation in explicitly correlated methods, which are denoted OptRI. These sets are specifically matched to OBSs and are optimized using the functional

$$\delta \text{RI} = \sum_{ij} \frac{(V_{ij,ij}^{\text{RI}} - V_{ij,ij}^{\text{RI}_{\text{ref}}})^2}{V_{ij,ij}^{\text{RI}_{\text{ref}}}} + \frac{(B_{ij,ij}^{\text{RI}} - B_{ij,ij}^{\text{RI}_{\text{ref}}})^2}{B_{ij,ij}^{\text{RI}_{\text{ref}}}}, \quad (17)$$

where **V** and **B** contain the many-electron integrals approximated by the RI,^[91] and RI_{ref} is a very large reference set of functions.^[214] The resulting OptRI ABSs are compact and produce negligible RI errors compared to the basis set incompleteness error of the underlying OBS, both for correlation and relative energies. Although originally published to match the cc-pVnZ-F12 bases, the same methodology has been used to produce OptRI sets specifically matched to aug-cc-pVnZ bases,^[215] to cc-pCVnZ-F12 bases,^[51] for the light alkali and alkaline earth elements,^[95] and for the group 11 and 12 transition metal elements.^[186] It has been observed that while these OptRI basis sets are not optimal for producing the largest CABS singles energy corrections, if relative energies are considered the OptRI sets in conjunction with CABS singles corrections are sufficient to practically eliminate the HF basis set error.^[216]

Outlook

The main purpose of this review was to give an overview of the current state of atom-centered Gaussian basis set development, mainly focusing on families of energy-optimized basis sets. With the bases currently available, it is possible to carry out high quality calculations on every element in the periodic table and one may be tempted to think that there will be little in terms of future basis set development. However, basis set families that offer systematic routes toward the CBS limit, such as the correlation consistent and pc bases, do not yet span the entire periodic table. For the correlation consistent family (excluding the sets designed for use in fully relativistic four-component calculations that are not currently practical for many calculations of interest), heavier alkali and alkaline earth elements are one area that is not currently covered, along with the lanthanide and actinide elements. Such elements pose problems in terms of strong relativistic effects and multi-reference character, but the publication of a new accurate relativistic PP for uranium may be a precursor for further development in this area.^[217]

There may also be interesting developments in terms of the pc basis sets and how they treat transition metal elements. With the popularity of DFT methods for transition metal complexes, it will be intriguing to discover the intrinsic exchange-correlation functional error for such systems, and this, in turn, may aid the development of more advanced functionals. It is safe to assume that some difficult design decisions will have to be considered in the development of such bases, with questions relating to the use of PPs and ensuring a balanced description of all important oxidation states.

Of course, it is also impossible to predict the basis set development opportunities brought about by future advances in quantum chemical methods. The obvious recent example of this is the explicitly correlated methods and their requirement for accurate RI ABSs. The development of OptRI or similar

auxiliary bases for transition metal elements may herald a new era of high-accuracy calculations on transition metal complexes, and new orbital bases specifically designed for the explicitly correlated methods are currently only available for a small section of the periodic table. It is not inconceivable that further fundamental theory advances may take us back to square one in terms of basis set requirements.

Acknowledgment

The author would like to thank Dr. Kirk A. Peterson for his careful reading of the manuscript prior to submission.

Keywords: basis sets · correlation consistent · density fitting · polarization consistent · Cholesky decomposition

How to cite this article: J. Grant Hill, *Int. J. Quantum Chem.* **2013**, *113*, 21–34. DOI: 10.1002/qua.24355

- [1] D. Feller. *J. Comput. Chem.* **1996**, *17*, 1571.
- [2] K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li, T. L. Windus. *J. Chem. Inf. Model.* **2007**, *47*, 1045.
- [3] EMSL Basis Set Exchange. Available at: <https://bse.pnl.gov/bse/portal>. Accessed on 19 October 2012.
- [4] F. Jensen. *Introduction to Computational Chemistry*, 2nd ed.; Wiley, Chichester, **2007**.
- [5] S. F. Boys. *Proc. R. Soc. Lond. A* **1950**, *200*, 542.
- [6] S. F. Boys, G. B. Cook, C. M. Reeves, I. Shavitt. *Nature* **1956**, *178*, 1207.
- [7] E. Clementi, D. R. Davis. *J. Comput. Phys.* **1967**, *1*, 223.
- [8] A. Baranowska, A. J. Sadlej. *J. Comput. Chem.* **2010**, *31*, 552.
- [9] W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling. *Numerical Recipes in FORTRAN 77: The Art of Scientific Computing*, 2nd ed.; Cambridge University Press: Cambridge, **1992**.
- [10] G. A. Petersson, S. Zhong, J. A. Montgomery, Jr., M. J. Frisch. *J. Chem. Phys.* **2003**, *118*, 1101.
- [11] R. C. Raffanetti. *J. Chem. Phys.* **1973**, *58*, 4452.
- [12] E. R. Davidson, D. Feller. *Chem. Rev.* **1986**, *86*, 681.
- [13] J. Almlöf, P. R. Taylor. *J. Chem. Phys.* **1987**, *86*, 4070.
- [14] J. Almlöf, T. Helgaker, P. R. Taylor. *J. Phys. Chem.* **1988**, *92*, 3029.
- [15] J. Almlöf, P. R. Taylor. *Adv. Quantum Chem.* **1991**, *22*, 301.
- [16] F. B. van Duijneveldt. IBM Research Report RJ: San Jose, California 945, **1971**.
- [17] C. W. Bauschlicher, Jr., P. R. Taylor. *Theor. Chim. Acta* **1993**, *86*, 13.
- [18] C. W. Bauschlicher, Jr. *Theor. Chim. Acta* **1995**, *92*, 183.
- [19] P.-O. Widmark, P.-Å. Malmqvist, B. O. Roos. *Theor. Chim. Acta* **1990**, *77*, 291.
- [20] P.-O. Widmark, B. J. Persson, B. O. Roos. *Theor. Chim. Acta* **1991**, *79*, 419.
- [21] R. Pou-Amérgio, M. Merchán, I. Nebot-Gil, P.-O. Widmark, B. O. Roos. *Theor. Chim. Acta* **1995**, *92*, 149.
- [22] K. Pierloot, B. Dumez, P.-O. Widmark, B. O. Roos. *Theor. Chim. Acta* **1995**, *90*, 87.
- [23] B. O. Roos, V. Veryazov, P.-O. Widmark. *Theor. Chem. Acc.* **2004**, *111*, 345.
- [24] M. Douglas, N. M. Kroll. *Ann. Phys. (N.Y.)* **1974**, *82*, 89.
- [25] G. Jansen, B. A. Hess. *Phys. Rev. A* **1989**, *39*, 6016.
- [26] K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej, K. Wolinski. *J. Phys. Chem.* **1990**, *94*, 5483.
- [27] K. Andersson, P.-Å. Malmqvist, B. O. Roos. *J. Chem. Phys.* **1992**, *96*, 1218.
- [28] B. O. Roos, R. Lindh, P.-Å. Malmqvist, V. Veryazov, P.-O. Widmark. *J. Phys. Chem. A* **2004**, *108*, 2851.
- [29] B. O. Roos, R. Lindh, P.-Å. Malmqvist, V. Veryazov, P.-O. Widmark. *J. Phys. Chem. A* **2005**, *109*, 6575.
- [30] B. O. Roos, R. Lindh, P.-Å. Malmqvist, V. Veryazov, P.-O. Widmark. *Chem. Phys. Lett.* **2005**, *409*, 295.
- [31] B. O. Roos, R. Lindh, P.-Å. Malmqvist, V. Veryazov, P.-O. Widmark, A. C. Borin. *J. Phys. Chem. A* **2008**, *112*, 11431.
- [32] F. Aquilante, L. De Vico, N. Ferré, G. Ghigo, P.-Å. Malmqvist, P. Neogrady, T. B. Pedersen, M. Pitonak, M. Reiher, B. O. Roos, L. Serrano-Andrés, M. Urban, V. Veryazov, R. Lindh. *J. Comput. Chem.* **2010**, *31*, 224.
- [33] F. Neese, E. F. Valeev. *J. Chem. Theory Comput.* **2011**, *7*, 33.
- [34] R. J. Gdanitz, R. Ahlrichs. *Chem. Phys. Lett.* **1988**, *143*, 413.
- [35] T. H. Dunning, Jr. *J. Chem. Phys.* **1989**, *90*, 1007.
- [36] R. J. Bartlett, M. Musiał. *Rev. Mod. Phys.* **2007**, *79*, 291.
- [37] D. Feller, K. A. Peterson, J. G. Hill. *J. Chem. Phys.* **2011**, *135*, 044102.
- [38] D. E. Woon, T. H. Dunning, Jr. *J. Chem. Phys.* **1994**, *100*, 2975.
- [39] T. van Mourik, A. K. Wilson, T. H. Dunning, Jr. *J. Mol. Struct. (THEO-CHEM)* **1996**, *388*, 339.
- [40] D. Feller, K. A. Peterson, T. D. Crawford. *J. Chem. Phys.* **2006**, *124*, 054107.
- [41] B. P. Prascher, D. E. Woon, K. A. Peterson, T. H. Dunning, Jr., A. K. Wilson. *Theor. Chem. Acc.* **2011**, *128*, 69.
- [42] D. E. Woon, T. H. Dunning, Jr. *J. Chem. Phys.* **1993**, *98*, 1358.
- [43] T. van Mourik, T. H. Dunning, Jr. *Int. J. Quantum Chem.* **2000**, *76*, 205.
- [44] A. K. Wilson, D. E. Woon, K. A. Peterson, T. H. Dunning, Jr. *J. Chem. Phys.* **1999**, *110*, 7667.
- [45] T. H. Dunning, Jr., K. A. Peterson, A. K. Wilson. *J. Chem. Phys.* **2001**, *114*, 9244.
- [46] K. A. Peterson. *Ann. Reports Comput. Chem.* **2007**, *3*, 195.
- [47] R. A. Kendall, T. H. Dunning, Jr., R. J. Harrison. *J. Chem. Phys.* **1992**, *96*, 6796.
- [48] D. E. Woon, T. H. Dunning, Jr. *J. Chem. Phys.* **1995**, *103*, 4572.
- [49] K. A. Peterson, T. H. Dunning, Jr. *J. Chem. Phys.* **2002**, *117*, 10548.
- [50] N. J. DeYonker, K. A. Peterson, A. K. Wilson. *J. Phys. Chem. A* **2007**, *111*, 11383.
- [51] J. G. Hill, S. Mazumder, K. A. Peterson. *J. Chem. Phys.* **2010**, *132*, 054108.
- [52] W. A. de Jong, R. J. Harrison, D. A. Dixon. *J. Chem. Phys.* **2001**, *114*, 48.
- [53] M. Dolg, X. Cao. *Chem. Rev.* **2012**, *112*, 403.
- [54] J. M. L. Martin, A. Sundermann. *J. Chem. Phys.* **2001**, *114*, 3408.
- [55] A. Bergner, M. Dolg, W. Küchle, H. Stoll, H. Preuß. *Mol. Phys.* **1993**, *80*, 1431.
- [56] B. Metz, H. Stoll, M. Dolg. *J. Chem. Phys.* **2000**, *113*, 2563.
- [57] K. A. Peterson. *J. Chem. Phys.* **2003**, *119*, 11099.
- [58] K. A. Peterson, D. Figgen, E. Goll, H. Stoll, M. Dolg. *J. Chem. Phys.* **2003**, *119*, 11113.
- [59] K. A. Peterson, K. E. Yousaf. *J. Chem. Phys.* **2010**, *132*, 174116.
- [60] C. W. Bauschlicher, Jr. *J. Chem. Phys. Lett.* **1999**, *305*, 446.
- [61] K. A. Peterson, B. C. Shepler, D. Figgen, H. Stoll. *J. Phys. Chem. A* **2006**, *110*, 13877.
- [62] K. G. Dyall. *Theor. Chim. Acta* **1998**, *99*, 366.
- [63] K. G. Dyall. *Theor. Chem. Acc.* **2002**, *108*, 365.
- [64] K. G. Dyall. *Theor. Chem. Acc.* **2002**, *108*, 335.
- [65] K. G. Dyall. *Theor. Chem. Acc.* **2006**, *115*, 441.
- [66] K. G. Dyall. *Theor. Chem. Acc.* **2012**, *131*, 1217.
- [67] K. G. Dyall. *Theor. Chem. Acc.* **2012**, *131*, 1217.
- [68] C. W. Bauschlicher, Jr. *Theor. Chem. Acc.* **1999**, *103*, 141.
- [69] A. Ricca, C. W. Bauschlicher, Jr. *Theor. Chem. Acc.* **2001**, *106*, 314.
- [70] H. Partridge. *J. Chem. Phys.* **1989**, *90*, 1043.
- [71] N. B. Balabanov, K. A. Peterson. *J. Chem. Phys.* **2005**, *123*, 064107.
- [72] N. B. Balabanov, K. A. Peterson. *J. Chem. Phys.* **2006**, *125*, 074110.
- [73] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, T. Korona, R. Lindh, A. Mitrushchenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobson, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, D. P. O'Neill, P. Palmieri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, and M. Wang. MOLPRO, Version 2010.1, a Package of Ab Initio Programs. Available at: <http://www.molpro.net>. Accessed on 19 October 2012.
- [74] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz. *WIREs Comput. Mol. Sci.* **2012**, *2*, 242.
- [75] D. Figgen, G. Rauhut, M. Dolg, H. Stoll. *Chem. Phys.* **2005**, *311*, 227.

- [76] K. A. Peterson, C. Puzzarini. *Theor. Chem. Acc.* **2005**, 114, 283.
- [77] K. A. Peterson, D. Figgen, M. Dolg, H. Stoll. *J. Chem. Phys.* **2007**, 126, 124101.
- [78] D. Figgen, K. A. Peterson, M. Dolg, H. Stoll. *J. Chem. Phys.* **2009**, 130, 164108.
- [79] K. G. Dyall. *Theor. Chem. Acc.* **2004**, 112, 403.
- [80] K. G. Dyall. *Theor. Chem. Acc.* **2007**, 117, 483.
- [81] J. Koput, K. A. Peterson. *J. Chem. Phys.* **2002**, 106, 9595.
- [82] C. W. Bauschlicher, Jr., S. R. Langhoff, H. Partridge. *J. Chem. Phys.* **1986**, 84, 901.
- [83] K. G. Dyall. *J. Phys. Chem. A* **2009**, 113, 12638.
- [84] A. S. P. Gomes, K. G. Dyall, L. Visscher. *Theor. Chem. Acc.* **2010**, 127, 369.
- [85] K. G. Dyall. *Theor. Chem. Acc.* **2007**, 117, 491.
- [86] D. P. Tew, C. Hättig, R. A. Bachorz, W. Klopper. In *Recent Progress in Coupled Cluster Methods: Theory and Applications*, P. Čársky, J. Paldus, J. Pittner, Eds.; Springer: Berlin, **2010**; p. 535.
- [87] H. -J. Werner, T. B. Adler, G. Knizia, F. R. Manby. In *Recent Progress in Coupled Cluster Methods: Theory and Applications*, P. Čársky, J. Paldus, J. Pittner, Eds.; Springer: Berlin, **2010**; p. 573.
- [88] C. Hättig, W. Klopper, A. Köhn, D. P. Tew. *Chem. Rev.* **2012**, 112, 4.
- [89] L. Kong, F. A. Bischoff, E. F. Valeev. *Chem. Rev.* **2012**, 112, 75.
- [90] D. P. Tew, W. Klopper, C. Neiss, C. Hättig. *Phys. Chem. Chem. Phys.* **2007**, 9, 1921.
- [91] H. -J. Werner, T. B. Adler, F. R. Manby. *J. Chem. Phys.* **2007**, 126, 164102.
- [92] G. Knizia, H. -J. Werner. *J. Chem. Phys.* **2008**, 128, 154103.
- [93] K. A. Peterson, T. B. Adler, H. -J. Werner. *J. Chem. Phys.* **2008**, 128, 084102.
- [94] J. G. Hill, K. A. Peterson, G. Knizia, H. -J. Werner. *J. Chem. Phys.* **2009**, 131, 194105.
- [95] J. G. Hill, K. A. Peterson. *Phys. Chem. Chem. Phys.* **2010**, 12, 10460.
- [96] J. A. Platts, J. G. Hill, K. E. Riley, J. Řezáč, P. Hobza (in press).
- [97] S. F. Boys, F. Bernardi. *Mol. Phys.* **1970**, 19, 553.
- [98] C. Schwartz. *Phys. Rev.* **1962**, 126, 1015.
- [99] R. N. Hill. *J. Chem. Phys.* **1985**, 83, 1173.
- [100] W. Kutzelnigg, J. D. Morgan, III. *J. Chem. Phys.* **1992**, 96, 4484.
- [101] W. Kutzelnigg, J. D. Morgan, III. *J. Chem. Phys.* **1992**, 97, 8821.
- [102] F. Jensen. *Theor. Chem. Acc.* **2000**, 104, 484.
- [103] K. A. Christensen, F. Jensen. *Chem. Phys. Lett.* **2000**, 317, 400.
- [104] W. J. Hehre, R. Ditchfield, J. A. Pople. *J. Chem. Phys.* **1972**, 56, 2257.
- [105] F. Jensen. *J. Chem. Phys.* **2001**, 115, 9113.
- [106] F. Jensen. *J. Chem. Phys.* **2002**, 116, 3502.
- [107] T. Koga, S. Watanabe, K. Kanayama, R. Yasuda, A. J. Thakkar. *J. Chem. Phys.* **1995**, 103, 3000.
- [108] A. D. Becke. *Phys. Rev. A* **1988**, 38, 3098.
- [109] C. Lee, W. Yang, R. G. Parr. *Phys. Rev. B* **1988**, 37, 785.
- [110] F. Jensen. *J. Chem. Phys.* **2002**, 116, 7372.
- [111] F. Jensen. *J. Chem. Phys.* **2002**, 117, 9234.
- [112] O. A. Vydrov, G. E. Scuseria, J. P. Perdew. *J. Chem. Phys.* **2007**, 126, 154109.
- [113] A. J. Cohen, P. Mori-Sánchez, W. Yang. *Science* **2008**, 321, 792.
- [114] F. Jensen. *J. Chem. Theory Comput.* **2006**, 2, 1360.
- [115] F. Jensen. *Theor. Chem. Acc.* **2010**, 126, 371.
- [116] F. Jensen, T. Helgaker. *J. Chem. Phys.* **2004**, 121, 3463.
- [117] F. Jensen. *J. Chem. Phys.* **2003**, 118, 2459.
- [118] F. Jensen. *J. Phys. Chem. A* **2007**, 111, 11198.
- [119] F. Jensen. *J. Chem. Phys.* **2012**, 136, 114107.
- [120] R. Ahlrichs, M. Bär, M. Häser, H. Horn, C. Kölmel. *Chem. Phys. Lett.* **1989**, 162, 165.
- [121] TURBOMOLE V6.4 2012, A Development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH, **2007**; Available at <http://www.turbomole.com>. Accessed on 19 October 2012.
- [122] F. Weigend, R. Ahlrichs. *Phys. Chem. Chem. Phys.* **2005**, 7, 3297.
- [123] F. Weigend, A. Baldes. *J. Chem. Phys.* **2010**, 133, 174102.
- [124] M. Dolg, H. Stoll, A. Savin, H. Preuß. *Theor. Chim. Acta* **1989**, 75, 173.
- [125] D. Andrae, U. Häußermann, M. Dolg, H. Stoll, H. Preuß. *Theor. Chim. Acta* **1990**, 77, 123.
- [126] M. Kaupp, P. v. R. Schleyer, H. Stoll, H. Preuß. *Chem. Phys. Lett.* **1991**, 94, 1360.
- [127] T. Leininger, A. Nicklass, W. Küchle, H. Stoll, M. Dolg, A. Bergner. *Chem. Phys. Lett.* **1996**, 255, 274.
- [128] A. Schäfer, H. Horn, R. Ahlrichs. *J. Chem. Phys.* **1992**, 97, 2571.
- [129] A. Schäfer, C. Huber, R. Ahlrichs. *J. Chem. Phys.* **1994**, 100, 5829.
- [130] K. Eichkorn, F. Weigend, O. Treutler, R. Ahlrichs. *Theor. Chem. Acc.* **1997**, 97, 119.
- [131] F. Weigend, M. Häser, H. Patzelt, R. Ahlrichs. *Chem. Phys. Lett.* **1998**, 294, 143.
- [132] F. Weigend, F. Furche, R. Ahlrichs. *J. Chem. Phys.* **2003**, 119, 12753.
- [133] R. Gulde, P. Pollak, F. Weigend. *J. Chem. Theory Comput. (in press)*, DOI:10.1021/ct300302u.
- [134] M. Dolg, H. Stoll, H. Preuss. *J. Chem. Phys.* **1989**, 90, 1730.
- [135] X. Cao, M. Dolg. *J. Mol. Struct. (THEOCHEM)* **2002**, 581, 139.
- [136] X. Cao, M. Dolg. *J. Chem. Phys.* **2001**, 115, 7348.
- [137] D. Rappoport, F. Furche. *J. Chem. Phys.* **2010**, 133, 134105.
- [138] J. Zheng, X. Xu, D. G. Truhlar. *Theor. Chem. Acc.* **2011**, 128, 295.
- [139] E. van Lenthe, E. J. Baerends, J. G. Snijders. *J. Chem. Phys.* **1993**, 99, 4597.
- [140] E. van Lenthe, E. J. Baerends, J. G. Snijders. *J. Chem. Phys.* **1994**, 101, 9783.
- [141] E. van Lenthe, J. G. Snijders, E. J. Baerends. *J. Chem. Phys.* **1996**, 105, 6505.
- [142] D. A. Pantazis, X. -Y. Chen, C. R. Landis, F. Neese. *J. Chem. Theory Comput.* **2008**, 4, 908.
- [143] D. A. Pantazis, F. Neese. *J. Chem. Theory Comput.* **2009**, 5, 2229.
- [144] D. A. Pantazis, F. Neese. *J. Chem. Theory Comput.* **2011**, 7, 667.
- [145] T. Noro, M. Sekiya, T. Koga. *Theor. Chem. Acc.* **1997**, 98, 25.
- [146] M. Sekiya, T. Noro, T. Koga, H. Matsuyama. *J. Mol. Struct. (THEOCHEM)* **1998**, 451, 51.
- [147] T. Noro, M. Sekiya, T. Koga, H. Matsuyama. *Theor. Chem. Acc.* **2000**, 104, 146.
- [148] M. Sekiya, T. Noro, Y. Osanai, T. Koga. *Theor. Chem. Acc.* **2001**, 106, 297.
- [149] T. Noro, M. Sekiya, T. Koga. *Theor. Chem. Acc.* **2003**, 109, 85.
- [150] T. Noro, M. Sekiya, T. Koga. *Theor. Chem. Acc.* **2008**, 121, 289.
- [151] T. Noro, M. Sekiya, Y. Osanai, E. Miyoshi, T. Koga. *J. Chem. Phys.* **2003**, 119, 5142.
- [152] M. Sekiya, T. Noro, E. Miyoshi, Y. Osanai, T. Koga. *J. Comput. Chem.* **2006**, 27, 463.
- [153] T. Noro, M. Sekiya, Y. Osanai, T. Koga, H. Matsuyama. *J. Comput. Chem.* **2007**, 28, 2511.
- [154] T. Noro, M. Sekiya, T. Koga, S. L. Saito. *Chem. Phys. Lett.* **2009**, 481, 229.
- [155] M. Sekiya, T. Noro, T. Koga, S. L. Saito. *J. Comput. Chem.* **2010**, 31, 497.
- [156] D. Feller, E. R. Davidson. *J. Chem. Phys.* **1981**, 74, 3977.
- [157] T. Noro, M. Sekiya, T. Koga. *Theor. Chem. Acc.* **2012**, 131, 1124.
- [158] M. Sekiya, T. Noro, T. Koga, T. Shimazaki. *Theor. Chem. Acc.* **2012**, 131, 1247.
- [159] T. Tsuchiya, M. Abe, T. Nakajima, K. Hirao. *J. Chem. Phys.* **2001**, 115, 4463.
- [160] T. Nakajima, K. Hirao. *J. Chem. Phys.* **2002**, 116, 8270.
- [161] D. M. Schrader, S. Prager. *J. Chem. Phys.* **1962**, 37, 1456.
- [162] J. L. Whitten. *J. Chem. Phys.* **1973**, 58, 4496.
- [163] B. I. Dunlap, J. W. D. Connolly, J. R. Sabin. *J. Chem. Phys.* **1979**, 71, 4993.
- [164] B. I. Dunlap. *J. Chem. Phys.* **1983**, 78, 3140.
- [165] O. Vahtras, J. Almlöf, M. Feyereisen. *Chem. Phys. Lett.* **1993**, 213, 514.
- [166] R. A. Kendall, H. A. Frúchtl. *Theor. Chem. Acc.* **1997**, 97, 158.
- [167] B. I. Dunlap. *J. Mol. Struct. (THEOCHEM)* **2000**, 529, 37.
- [168] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, R. Ahlrichs. *Chem. Phys. Lett.* **1995**, 240, 283.
- [169] B. I. Dunlap. *J. Mol. Struct. (THEOCHEM)* **2000**, 501–502, 221.
- [170] B. I. Dunlap. *J. Phys. Chem. A* **2003**, 107, 10082.
- [171] S. Reine, E. Tellgren, A. Krapp, T. Kjærgaard, T. Helgaker, B. Jansik, S. Høst, P. Salek. *J. Chem. Phys.* **2008**, 129, 104101.
- [172] R. Yang, A. P. Rendell, M. J. Frisch. *J. Chem. Phys.* **2007**, 127, 074102.
- [173] F. Aquilante, R. Lindh, T. B. Pedersen. *J. Chem. Phys.* **2007**, 127, 114107.
- [174] K. Eichkorn, F. Weigend, O. Treutler, R. Ahlrichs. *Theor. Chem. Acc.* **1997**, 97, 119.
- [175] F. Weigend. *Phys. Chem. Chem. Phys.* **2006**, 8, 1057.

- [176] E. J. Baerends, D. E. Ellis, P. Ros. *Chem. Phys.* **1973**, 2, 41.
- [177] S. Ten-no, S. Iwata. *Chem. Phys. Lett.* **1995**, 240, 578.
- [178] F. Weigend. *Phys. Chem. Chem. Phys.* **2002**, 4, 4285.
- [179] R. Polly, H. -J. Werner, F. R. Manby, P. J. Knowles. *Mol. Phys.* **2004**, 102, 2311.
- [180] M. Feyereisen, G. Fitzgerald, A. Komornicki. *Chem. Phys. Lett.* **1993**, 208, 359.
- [181] A. P. Rendell, T. J. Lee. *J. Chem. Phys.* **1994**, 101, 400.
- [182] C. Hättig, F. Weigend. *J. Chem. Phys.* **2000**, 113, 5154.
- [183] H. -J. Werner, F. R. Manby, P. J. Knowles. *J. Chem. Phys.* **2003**, 118, 8149.
- [184] F. Weigend. *J. Comput. Chem.* **2008**, 29, 167.
- [185] J. G. Hill, J. A. Platts. *J. Chem. Theory Comput.* **2009**, 5, 500.
- [186] J. G. Hill, K. A. Peterson. *J. Chem. Theory Comput.* **2012**, 8, 518.
- [187] C. Hättig. *Phys. Chem. Chem. Phys.* **2005**, 7, 59.
- [188] F. Weigend, A. Köhn, C. Hättig. *J. Chem. Phys.* **2002**, 116, 3175.
- [189] A. Hellweg, C. Hättig, S. Höfener, W. Klopper. *Theor. Chem. Acc.* **2007**, 117, 587.
- [190] J. G. Hill, J. A. Platts. *J. Chem. Phys.* **2008**, 128, 044104.
- [191] J. G. Hill, J. A. Platts. *J. Chem. Phys.* **2008**, 129, 134101.
- [192] J. G. Hill. *J. Chem. Phys.* **2011**, 135, 044105.
- [193] C. Hättig, G. Schmitz, J. Kößmann. *Phys. Chem. Chem. Phys.* **2012**, 14, 6549.
- [194] T. B. Pedersen, F. Aquilante, R. Lindh. *Theor. Chem. Acc.* **2009**, 124, 1.
- [195] F. Aquilante, L. Boman, J. Boström, H. Koch, R. Lindh, A. Sánchez de Merás, T. B. Pedersen. In *Linear-Scaling Techniques in Computational Chemistry and Physics*, Vol. 13; R. Zalesny, M. G. Papadopoulos, P. G. Mezey, J. Leszczynski, Eds.; Springer: Berlin, **2011**, p. 301.
- [196] N. H. F. Beebe, J. Linderberg. *Int. J. Quantum Chem.* **1977**, 12, 683.
- [197] J. Boström, F. Aquilante, T. B. Pedersen, R. Lindh. *J. Chem. Theory Comput.* **2009**, 5, 1545.
- [198] F. Weigend, M. Kattannek, R. Ahlrichs. *J. Chem. Phys.* **2009**, 130, 164106.
- [199] F. Aquilante, L. Gagliardi, T. B. Pedersen, R. Lindh. *J. Chem. Phys.* **2009**, 130, 154107.
- [200] F. Aquilante, T. B. Pedersen, R. Lindh, B. O. Roos, A. Sánchez de Merás, H. Koch. *J. Chem. Phys.* **2008**, 129, 024113.
- [201] J. Boström, M. G. Delcey, F. Aquilante, L. Serrano-Andrés, T. B. Pedersen, R. Lindh. *J. Chem. Theory Comput.* **2010**, 6, 747.
- [202] J. Boström, M. Pitoňák, F. Aquilante, P. Neogrády, T. B. Pedersen, R. Lindh. *J. Chem. Theory Comput.* **2012**, 8, 1921.
- [203] F. Aquilante, T. B. Pedersen, R. Lindh. *J. Chem. Phys.* **2007**, 126, 194106.
- [204] K. Raghavachari, G. W. Trucks, J. A. Pople, M. Head-Gordon. *Chem. Phys. Lett.* **1989**, 157, 479.
- [205] T. B. Adler, G. Knizia, H. -J. Werner. *J. Chem. Phys.* **2007**, 127, 221106.
- [206] G. Knizia, T. B. Adler, H. -J. Werner. *J. Chem. Phys.* **2009**, 130, 054104.
- [207] C. Hättig, D. P. Tew, A. Köhn. *J. Chem. Phys.* **2010**, 132, 231102.
- [208] F. R. Manby. *J. Chem. Phys.* **2003**, 119, 4607.
- [209] W. Kutzelnigg. *Theor. Chim. Acta* **1985**, 68, 445.
- [210] W. Kutzelnigg, W. Klopper. *J. Chem. Phys.* **1991**, 94, 1985.
- [211] W. Klopper, W. Kutzelnigg. *J. Chem. Phys.* **1991**, 94, 2020.
- [212] W. Klopper, C. C. M. Samson. *J. Chem. Phys.* **2002**, 116, 6397.
- [213] E. F. Valeev. *Chem. Phys. Lett.* **2004**, 395, 190.
- [214] K. E. Yousaf, K. A. Peterson. *J. Chem. Phys.* **2008**, 129, 184108.
- [215] K. E. Yousaf, K. A. Peterson. *Chem. Phys. Lett.* **2009**, 476, 303.
- [216] A. Köhn, D. P. Tew. *J. Chem. Phys.* **2010**, 132, 024101.
- [217] M. Dolg, X. Cao. *J. Phys. Chem. A* **2009**, 113, 12573.

Received: 30 August 2012
Revised: 28 September 2012
Accepted: 2 October 2012
Published online on 29 October 2012