

Theor Chem Acc (2010) 127:369–381
DOI 10.1007/s00214-009-0725-7

REGULAR ARTICLE

Relativistic double-zeta, triple-zeta, and quadruple-zeta basis sets for the lanthanides La–Lu

André S. P. Gomes · Kenneth G. Dyall ·
Lucas Visscher

Received: 17 November 2009 / Accepted: 31 December 2009 / Published online: 23 January 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract Relativistic basis sets of double-zeta, triple-zeta, and quadruple-zeta quality have been optimized for the lanthanide elements La–Lu. The basis sets include SCF exponents for the occupied spinors and for the 6p shell, exponents of correlating functions for the valence shells (4f, 5d and 6s) and the outer core shells (4d, 5s and 5p), and diffuse functions, including functions for dipole polarization of the 4f shell. A finite nuclear size was used in all optimizations. The basis sets are illustrated by calculations on YbF. Prescriptions are given for constructing contracted basis sets. The basis sets are available as an internet archive and from the Dirac program web site, <http://dirac.chem.sdu.dk>.

Keywords Gaussian basis sets · Relativistic basis sets · Lanthanide elements · Double zeta · Triple zeta · Quadruple zeta · Correlating functions · Ytterbium fluoride · Spectroscopic constants

1 Introduction

The increasing use of all-electron four-component methodology for relativistic effects in molecular structure calculations brings with it a need for high-quality basis sets. Several other authors have generated basis sets at the SCF level for part or all of the periodic table, either with the Dirac Hamiltonian or with the Douglas-Kroll-Hess Hamiltonian [1–19]. While some of these basis sets include correlating functions, a number of them lack the functions needed for valence and outer core correlation and polarization of the SCF sets. In a series of papers [20–27], Dyall has focused on providing consistent basis sets of double-zeta (dz), triple-zeta (tz), and quadruple-zeta (qz) quality for the heavy elements, including correlating functions for the valence and outer core orbitals in the style of the correlation consistent basis sets [28–31]. So far these basis sets have covered the 4s, 5s, 6s, and 7s blocks; the 4p, 5p, and 6p blocks; the 4d and 5d transition series; and the actinides. This paper continues the series by presenting basis sets for the lanthanides. Work is in progress on the 3d and 6d transition series and the 7p block.

The lanthanides have similar challenges to the actinides, but also have some unique features, as far as basis sets are concerned. As for the actinides, the 5d orbital is occupied for several members of the series (La, Ce, Gd, and Lu), and is low-lying for many of the early members of the series. It is in fact doubly occupied in some excited states early in the series. It must therefore be included in the basis set

Electronic supplementary material The online version of this article (doi:10.1007/s00214-009-0725-7) contains supplementary material, which is available to authorized users.

A. S. P. Gomes
Laboratoire PhLAM, CNRS UMR 8523, Université de Lille 1,
Bât P5, 59655 Villeneuve d'Ascq Cedex, France
e-mail: gomes@phlam.univ-lille1.fr

K. G. Dyall (✉)
Schrödinger, Inc, 101 SW Main Street Suite 1300,
Portland, OR 97204, USA
e-mail: ken.dyall@schrodinger.com

L. Visscher
Department of Theoretical Chemistry, Faculty of Sciences,
Amsterdam Center for Multiscale Modeling,
VU University Amsterdam, De Boelelaan 1083,
1081 HV Amsterdam, The Netherlands
e-mail: visscher@chem.vu.nl

optimization. The 6p orbital is not occupied, but is needed for polarization of the 5d and 6s orbitals. However, unlike the actinides, this orbital must be doubly occupied in the exponent optimization to prevent it from becoming too tight and thus not serve its purpose as a polarization orbital. Correlation of the 5s and 5p shells is critical as it is for the actinides. Obviously, correlation of the 4f is critical, but since the 4f is the first shell of its symmetry, it does not have the orthogonality constraint that the 5f has in the actinides, and so is much more compact relative to the other shells. This influences the extent to which the ranges of the correlating functions can overlap. Because of the compactness of the 4f, correlation of the 4d is also important.

2 SCF basis sets

The methods used for the SCF basis set optimization have been described previously [20, 22, 32, 33]. The basis sets were optimized in Dirac-Hartree-Fock calculations using the Dirac-Coulomb Hamiltonian with the standard Gaussian nuclear charge distribution [34] for the most abundant (or stable) isotope. As for previous basis sets in this series, ℓ -optimization was employed. In the SCF optimizations, the exponents are varied only within a given angular space, i.e. for a given ℓ value, with all other exponents fixed. This is because the exponents in each angular space are to a large degree independent of those in the other angular spaces. The angular spaces are cycled through the optimization process until there is no significant change in the total energy and the gradient with respect to the logarithms of the exponents.

The electron configuration used for the basis set optimization was as follows. For the elements in which the 5d is occupied (La, Ce, Gd, Lu), the ground configuration was used for all angular symmetries. For the remaining elements, the ground $4f^n 6s^2$ configuration was used for all symmetries except d, where the $4f^{n-1} 5d^1 6s^2$ configuration was used. For the optimization of the 6p function, the configuration in which $6s^2$ was replaced with $6p^2$ in the ground configuration was used.

The size of the basis sets was determined for consistency with basis sets across the row, taking into account the basis set sizes of the 6s, 6p and 5d blocks, and for the quality of the representation of the outermost maxima. As a consequence of this consistency criterion, some revisions were deemed necessary for the 5d block: these revisions are described elsewhere [35].

The size of the dz basis set optimized for Ba [27] is 24s16p10d, and the revised 5d basis set size is 24s17p12d7f [35]. The s set size was therefore chosen to be 24s. For the p set, the choices are 16p or 17p. Optimization of both 16p

and 17p sets across the lanthanide series revealed that the 16p sets were too tight by the end of the block, so the larger 17p sets were selected. Because the 5d is more diffuse for the lanthanides than for the 5d transition series, the d set size chosen for the lanthanides was 13d. These choices yielded a 24s17p13d7f dz basis set. Two p functions were then added for the 6p, and the outermost 9 functions were reoptimized. These functions describe the outermost maxima of the 3p through 6p orbitals. The final SCF basis set size including the 6p functions is 24s19p13d7f.

For triple-zeta, the size of the basis set for Ba is 30s21p13d, the revised 5d set size is 30s21p15d10f, and the 6p set size is 30s26p16d10f. The s and p set sizes were therefore taken to be 30s and 21p. As for the dz basis sets, an extra d function was added for the lanthanides relative to the 5d series because the 5d is more diffuse in the lanthanides. These choices yielded a basis set size of 30s21p16d10f. Three p functions were added for the 6p, and the outermost 12 exponents were reoptimized. The final SCF basis set size including the 6p functions is 30s24p16d10f.

For quadruple-zeta, the basis set size for Ba is 35s26p17d, and the 5d basis set size is 34s26p19d12f, with 34s31p20d12f for the 6p series. After optimizing both a 34s and a 35s set, the 35s set was chosen for the lanthanides. Addition of an extra d function to make a 20d set was found to yield 6 functions for the 5d. For reasonable balance, 5 functions are required, which was obtained with a 19d set. For the p set, 4 functions were added to the 26p set for the 6p, and the outermost 14 functions were reoptimized. The final SCF basis set size including the 6p functions is 35s30p19d12f.

Total energies for the ground configuration are given in Table 1. The basis sets used in these calculation include the 6p set and the f functions added for 5d and 6s correlation in the double-zeta and triple-zeta basis sets, as described below. The energies for the quadruple-zeta basis sets fall below the numerical energy, due to the phenomenon of “prolapse” in relativistic basis set calculations [3].

3 Correlating and dipole polarizing functions

Correlating functions for the 4d, 4f, 5s, 5p, 5d, and 6s shells were optimized in MR-SDCI calculations, using the RAMCI program [33] modified for basis set optimization. All exponents are fully optimized. Exponents for the double-zeta basis sets are given in Table 2. Exponents for the triple-zeta basis sets are given in Tables 3, 4, 5, 6, 7. The f functions reported for 5d6s correlation are the outermost SCF f function and the even-tempered extension. Exponents for the quadruple-zeta basis sets are given in Tables 8, 9, 10, 11. The procedure for the various shells is described below.

Table 1 Total SCF configuration average energies in E_h for uncontracted basis set and numerical calculations on the ground configuration

| Element | Double-zeta | Triple-zeta | Quadruple-zeta | Numerical |
|---------|----------------|----------------|----------------|----------------|
| La | -8,493.639341 | -8,493.645370 | -8,493.645719 | -8,493.645716 |
| Ce | -8,861.064082 | -8,861.071103 | -8,861.071492 | -8,861.071488 |
| Pr | -9,238.139031 | -9,238.148090 | -9,238.148530 | -9,238.148527 |
| Nd | -9,625.124869 | -9,625.135267 | -9,625.135758 | -9,625.135755 |
| Pm | -10,022.082957 | -10,022.094819 | -10,022.095364 | -10,022.095359 |
| Sm | -10,429.149086 | -10,429.162518 | -10,429.163098 | -10,429.163114 |
| Eu | -10,846.489254 | -10,846.504417 | -10,846.505085 | -10,846.505076 |
| Gd | -11,274.214430 | -11,274.229516 | -11,274.230220 | -11,274.230204 |
| Tb | -11,712.525421 | -11,712.544506 | -11,712.545321 | -11,712.545308 |
| Dy | -12,161.516977 | -12,161.538275 | -12,161.539172 | -12,161.539155 |
| Ho | -12,621.377725 | -12,621.401439 | -12,621.402427 | -12,621.402408 |
| Er | -13,092.251084 | -13,092.277414 | -13,092.278502 | -13,092.278479 |
| Tm | -13,574.286755 | -13,574.315900 | -13,574.317094 | -13,574.317066 |
| Yb | -14,067.643816 | -14,067.675985 | -14,067.677294 | -14,067.677260 |
| Lu | -14,572.500730 | -14,572.531976 | -14,572.533305 | -14,572.533253 |

Table 2 Exponents of correlating and dipole polarizing functions for the double-zeta basis sets

| Element | 4f corr g | 4f pol g | 5d6s corr f | 5d pol f | 4d corr f |
|---------|-------------|-------------|-------------|-------------|-------------|
| La | | | 8.45270E-02 | 9.44513E-02 | 3.77622E+00 |
| Ce | 2.41155E+00 | 4.84872E-01 | 9.02210E-02 | 9.06142E-02 | 5.07203E+00 |
| Pr | 2.73613E+00 | 3.29341E-01 | 9.05129E-02 | 9.26754E-02 | 5.19558E+00 |
| Nd | 3.03289E+00 | 3.68511E-01 | 9.26514E-02 | 9.25793E-02 | 5.52414E+00 |
| Pm | 3.32315E+00 | 4.02319E-01 | 9.71568E-02 | 9.44735E-02 | 5.85160E+00 |
| Sm | 3.61136E+00 | 4.35774E-01 | 9.99005E-02 | 9.31816E-02 | 6.18377E+00 |
| Eu | 3.89400E+00 | 4.69244E-01 | 1.02652E-01 | 9.40343E-02 | 6.51097E+00 |
| Gd | 4.23840E+00 | 6.91846E-01 | 1.05171E-01 | 9.21335E-02 | 7.02533E+00 |
| Tb | 4.49742E+00 | 5.06910E-01 | 1.08638E-01 | 9.27688E-02 | 7.18146E+00 |
| Dy | 4.79772E+00 | 5.25552E-01 | 1.12412E-01 | 9.01388E-02 | 7.52277E+00 |
| Ho | 5.09790E+00 | 5.41688E-01 | 1.15743E-01 | 8.87705E-02 | 7.86869E+00 |
| Er | 5.40093E+00 | 5.56074E-01 | 1.19546E-01 | 8.68166E-02 | 8.21940E+00 |
| Tm | 5.70767E+00 | 5.71565E-01 | 1.22281E-01 | 8.68192E-02 | 8.58081E+00 |
| Yb | 6.01858E+00 | 5.88486E-01 | 1.25887E-01 | 8.48782E-02 | 8.94219E+00 |
| Lu | 6.57288E+00 | 8.62210E-01 | 1.30709E-01 | 7.88095E-02 | 9.48737E+00 |

3.1 4f correlation

In the basis set optimizations for other blocks, including the 5f, it has been assumed that it is sufficient to add a single function with one unit higher in angular momentum to correlate the main shell for the dz basis set, then increase the maximum correlating angular momentum by one and add one function for each correlating angular momentum, going up the basis set series to tz and qz. Owing to the nodeless nature of the 4f shell, this assumption was tested in a series of calculations. These tests assessed the effect of adding functions in a single symmetry (g, h, i, k) in CI calculations involving double excitations from the 4f shell into shells of

that symmetry. The tests showed that the 1g and 2g1h sets are fairly well balanced. The first i function has about the same energy gain as the fourth g function and the third h function, making 4g3h1i fairly well balanced. The next well-balanced set is 5g4h3i2k (k was the highest angular momentum in the study). Balance between the symmetries is not the only consideration, however: it is also important to be able to extrapolate the basis sets smoothly within symmetries. It was considered that the additional g and h functions in the 4g3h1i set might unduly perturb the series used for extrapolation. Therefore, a 3g2h1i set was chosen for the qz basis set, and the scheme used for other basis sets is deemed appropriate for the 4f shell.

Table 3 Exponents of 4f correlating 2g1h functions for the triple-zeta basis sets

| Element | g | g | h |
|---------|-------------|-------------|-------------|
| Ce | 4.60187E+00 | 1.26645E+00 | 2.46227E+00 |
| Pr | 5.21894E+00 | 1.44452E+00 | 2.80259E+00 |
| Nd | 5.73909E+00 | 1.60617E+00 | 3.31802E+00 |
| Pm | 6.25199E+00 | 1.76176E+00 | 3.70020E+00 |
| Sm | 6.76566E+00 | 1.91412E+00 | 4.10229E+00 |
| Eu | 7.27747E+00 | 2.06410E+00 | 4.63184E+00 |
| Gd | 7.86364E+00 | 2.27042E+00 | 4.44128E+00 |
| Tb | 8.30901E+00 | 2.37782E+00 | 5.06931E+00 |
| Dy | 8.83334E+00 | 2.53297E+00 | 5.45901E+00 |
| Ho | 9.35928E+00 | 2.68663E+00 | 5.82653E+00 |
| Er | 9.89086E+00 | 2.84057E+00 | 6.14555E+00 |
| Tm | 1.04311E+01 | 2.99619E+00 | 6.54057E+00 |
| Yb | 1.09797E+01 | 3.15409E+00 | 7.07280E+00 |
| Lu | 1.20253E+01 | 3.52552E+00 | 7.72327E+00 |

Table 4 Exponents of 5s5p correlating 2d1f functions for the triple-zeta basis sets

| Element | d | d | f |
|---------|-------------|-------------|-------------|
| La | 8.83728E-01 | 4.98922E-01 | 7.34543E-01 |
| Ce | 9.42741E-01 | 5.21011E-01 | 7.50120E-01 |
| Pr | 7.97131E-01 | 3.49380E-01 | 7.20805E-01 |
| Nd | 8.34188E-01 | 3.63175E-01 | 7.51011E-01 |
| Pm | 8.70921E-01 | 3.76720E-01 | 7.81409E-01 |
| Sm | 9.07631E-01 | 3.90152E-01 | 8.12042E-01 |
| Eu | 9.44635E-01 | 4.03585E-01 | 8.42992E-01 |
| Gd | 1.08505E+00 | 5.01588E-01 | 9.03196E-01 |
| Tb | 1.01938E+00 | 4.30475E-01 | 9.05937E-01 |
| Dy | 1.05714E+00 | 4.43959E-01 | 9.37989E-01 |
| Ho | 1.09580E+00 | 4.57666E-01 | 9.70600E-01 |
| Er | 1.13446E+00 | 4.71338E-01 | 1.00361E+00 |
| Tm | 1.17434E+00 | 4.85335E-01 | 1.03730E+00 |
| Yb | 1.21501E+00 | 4.99537E-01 | 1.07159E+00 |
| Lu | 1.44770E+00 | 6.91863E-01 | 1.18257E+00 |

For the 4f shell the functions were optimized on the ground f^ns^2 or $f^{n-1}d^1s^2$ configuration. The configuration space included all double excitations from the 4f shell into the correlating space, which consisted of 1g for the double-zeta basis sets, 2g1h for triple-zeta, and 3g2h1i for quadruple-zeta. Only configuration state functions (CSF) coupling to the ground state J value were included in the optimization, mainly due to the large number of such CSF. Due to the number of CSF generated for elements in the middle of the block with the quadruple-zeta basis sets, it was necessary to perform separate CI calculations, splitting the correlating set into a 3g set and a 2h1i set. Since the

Table 5 Exponents of 5d6s correlating 2f1g functions for the triple-zeta basis sets

| Element | f | f | g |
|---------|-------------|-------------|-------------|
| La | 4.53557E-01 | 8.94200E-02 | 1.39474E-01 |
| Ce | 4.11579E-01 | 9.32460E-02 | 1.74529E-01 |
| Pr | 4.94792E-01 | 9.26257E-02 | 1.54023E-01 |
| Nd | 5.11550E-01 | 9.49640E-02 | 1.55532E-01 |
| Pm | 5.47907E-01 | 9.81182E-02 | 1.68651E-01 |
| Sm | 5.84265E-01 | 1.01272E-01 | 1.81771E-01 |
| Eu | 6.05560E-01 | 1.04331E-01 | 1.73664E-01 |
| Gd | 6.25271E-01 | 1.21678E-01 | 1.11212E-01 |
| Tb | 6.54741E-01 | 1.10432E-01 | 1.81414E-01 |
| Dy | 6.88647E-01 | 1.14012E-01 | 1.89141E-01 |
| Ho | 7.14003E-01 | 1.17401E-01 | 1.93543E-01 |
| Er | 7.39580E-01 | 1.21146E-01 | 2.01880E-01 |
| Tm | 7.62879E-01 | 1.24412E-01 | 1.97891E-01 |
| Yb | 7.89083E-01 | 1.28222E-01 | 2.02366E-01 |
| Lu | 8.08137E-01 | 1.32066E-01 | 1.98590E-01 |

Table 6 Exponents of 4d correlating 2f1g functions for the triple-zeta basis sets

| Element | f | f | g |
|---------|-------------|-------------|-------------|
| La | 6.38586E+00 | 2.38265E+00 | 4.28068E+00 |
| Ce | 6.89973E+00 | 3.20714E+00 | 4.53039E+00 |
| Pr | 7.26815E+00 | 3.25513E+00 | 4.74863E+00 |
| Nd | 7.65872E+00 | 3.46087E+00 | 5.01983E+00 |
| Pm | 8.05469E+00 | 3.66441E+00 | 5.29475E+00 |
| Sm | 8.45823E+00 | 3.87223E+00 | 5.57332E+00 |
| Eu | 8.86541E+00 | 4.07065E+00 | 5.85778E+00 |
| Gd | 9.27153E+00 | 4.40581E+00 | 6.18852E+00 |
| Tb | 9.70613E+00 | 4.47951E+00 | 6.43933E+00 |
| Dy | 1.01380E+01 | 4.68651E+00 | 6.73803E+00 |
| Ho | 1.05778E+01 | 4.89576E+00 | 7.04224E+00 |
| Er | 1.10257E+01 | 5.10727E+00 | 7.35206E+00 |
| Tm | 1.14818E+01 | 5.32143E+00 | 7.66765E+00 |
| Yb | 1.19465E+01 | 5.53865E+00 | 7.98901E+00 |
| Lu | 1.23938E+01 | 5.87922E+00 | 8.35247E+00 |

angular spaces are fairly well decoupled, this is not a serious approximation.

3.2 5d and 6s correlation

For the 5d and 6s shells, correlating functions were optimized on the $f^{n-1}d^1s^2$ configuration. The configuration space included all double excitations from these two shells into the correlating space, which consisted of 1f for the double-zeta basis sets, 2f1g for triple-zeta, and 3f2g1h for quadruple-zeta. For the double-zeta and triple-zeta sets, the

Table 7 Exponents of 4f dipole polarization functions for the triple-zeta basis sets

| Element | g | h |
|---------|-------------|-------------|
| Ce | 2.54635E-01 | 4.95069E-01 |
| Pr | 1.73873E-01 | 3.37341E-01 |
| Nd | 1.95157E-01 | 4.03155E-01 |
| Pm | 2.13288E-01 | 4.47966E-01 |
| Sm | 2.30971E-01 | 4.95012E-01 |
| Eu | 2.48734E-01 | 5.58158E-01 |
| Gd | 3.70606E-01 | 7.24963E-01 |
| Tb | 2.68007E-01 | 5.71368E-01 |
| Dy | 2.77466E-01 | 5.97990E-01 |
| Ho | 2.85473E-01 | 6.19109E-01 |
| Er | 2.92462E-01 | 6.32740E-01 |
| Tm | 3.00038E-01 | 6.54971E-01 |
| Yb | 3.08402E-01 | 6.91566E-01 |
| Lu | 4.62466E-01 | 1.01311E+00 |

energy was averaged over all possible coupled angular momenta for the CSF. For the quadruple-zeta basis sets, the 4f shell was coupled to the highest possible angular momentum to reduce the number of configurations to a minimum, and the average taken over all possible couplings to this angular momentum. Since the f shell is frozen in these optimizations, this approximation has little impact on the results.

Because the f functions have some overlap with the SCF functions, it is only necessary to add to the basis set those functions that are not already represented. For the double-zeta basis sets, the f function was added to the SCF f set. For the quadruple-zeta basis sets, the SCF set covers the range of the correlating f functions, so the optimized

functions are not needed. The choice for the triple-zeta sets was more involved, and is discussed next.

As for the 5f block, two minima were found for the outer d correlating 2f1g functions in the triple-zeta basis sets: one in which the f functions were tighter and one in which the f functions were more diffuse. The g functions in both sets were about the same. The tighter set exists for all elements except Pm; the looser set exists from Pm through Lu. It is likely that the orthogonalization to the 4f influences the optimization. The exponents in both sets are only partly covered by the outer SCF f functions, but an even-tempered extension of one function serves to cover both the correlating f space and a reasonable representation of the dipole polarizing f space, so the even-tempered extension has been used to supplement the SCF f set instead.

3.3 5s and 5p correlation

Correlating functions for the 5s and 5p shells were optimized in MR-SDCI calculations on the ground configuration, for the state with the maximum J value. Single and double excitations out of the 5s and 5p shells into the correlating space were coupled to $J = 0$, to ensure that only configurations representing 5s and 5p correlation were included. The correlating space consisted of 1d for the double-zeta basis sets, 2d1f for triple-zeta, and 3d2f1g for quadruple-zeta. These optimizations are not strictly necessary for the dz and tz basis sets, because the range of correlating functions is covered by the SCF occupied set. However, they are a useful guide to the functions that need to be uncontracted. The d and f functions are not included in the final basis set.

For the qz basis sets, it is necessary to at least examine the g function to see if it is covered by any of the other

Table 8 Exponents of 4f correlating 3g2h1i functions for the quadruple-zeta basis sets

| Element | g | g | g | h | h | i |
|---------|-------------|-------------|-------------|-------------|-------------|-------------|
| Ce | 7.43139E+00 | 2.42945E+00 | 7.79990E-01 | 4.52559E+00 | 1.31866E+00 | 3.26586E+00 |
| Pr | 8.40821E+00 | 2.75921E+00 | 8.90537E-01 | 5.16524E+00 | 1.51223E+00 | 3.65619E+00 |
| Nd | 9.17466E+00 | 3.04563E+00 | 9.92588E-01 | 6.00969E+00 | 1.82116E+00 | 4.04447E+00 |
| Pm | 9.93735E+00 | 3.32463E+00 | 1.08959E+00 | 6.63026E+00 | 2.03632E+00 | 4.44271E+00 |
| Sm | 1.07128E+01 | 3.60113E+00 | 1.18364E+00 | 7.27481E+00 | 2.26242E+00 | 4.81035E+00 |
| Eu | 1.13610E+01 | 3.86902E+00 | 1.28127E+00 | 7.38091E+00 | 2.26252E+00 | 5.10757E+00 |
| Gd | 1.23567E+01 | 4.24876E+00 | 1.42714E+00 | 7.95057E+00 | 2.45644E+00 | 5.52707E+00 |
| Tb | 1.28900E+01 | 4.41968E+00 | 1.46645E+00 | 8.60433E+00 | 2.67914E+00 | 5.88158E+00 |
| Dy | 1.36920E+01 | 4.69687E+00 | 1.55651E+00 | 9.34572E+00 | 2.95329E+00 | 6.26361E+00 |
| Ho | 1.44860E+01 | 4.98053E+00 | 1.65040E+00 | 9.96413E+00 | 3.16239E+00 | 6.64896E+00 |
| Er | 1.53152E+01 | 5.27543E+00 | 1.74794E+00 | 1.06318E+01 | 3.39463E+00 | 7.06079E+00 |
| Tm | 1.61159E+01 | 5.55717E+00 | 1.83927E+00 | 1.12566E+01 | 3.60686E+00 | 7.44285E+00 |
| Yb | 1.69300E+01 | 5.84306E+00 | 1.93167E+00 | 1.20159E+01 | 3.89459E+00 | 7.82630E+00 |
| Lu | 1.86329E+01 | 6.53462E+00 | 2.22348E+00 | 1.31505E+01 | 4.34768E+00 | 8.58032E+00 |

Table 9 Exponents of 5d6s correlating 3f2g1h functions for the quadruple-zeta basis sets

| Element | f | f | f | g | g | h |
|---------|-------------|-------------|-------------|-------------|-------------|-------------|
| La | 5.80878E-01 | 1.93168E-01 | 8.21761E-02 | 4.84954E-01 | 1.30103E-01 | 1.69562E-01 |
| Ce | 5.96654E-01 | 2.42778E-01 | 9.36315E-02 | 6.20372E-01 | 1.83737E-01 | 3.54283E-01 |
| Pr | 5.34349E-01 | 1.93241E-01 | 8.38357E-02 | 6.07091E-01 | 1.57983E-01 | 2.12586E-01 |
| Nd | 5.46379E-01 | 1.46165E-01 | 7.84318E-02 | 5.36163E-01 | 1.46121E-01 | 1.89756E-01 |
| Pm | 5.69845E-01 | 1.49345E-01 | 8.05277E-02 | 7.16390E-01 | 1.60845E-01 | 1.93460E-01 |
| Sm | 5.91201E-01 | 1.53332E-01 | 8.25137E-02 | 7.54269E-01 | 1.64648E-01 | 1.97858E-01 |
| Eu | 6.15880E-01 | 1.58010E-01 | 8.41734E-02 | 5.34114E-01 | 1.52409E-01 | 2.05083E-01 |
| Gd | 6.34338E-01 | 1.61318E-01 | 8.58553E-02 | 3.99425E-01 | 1.39819E-01 | 2.10143E-01 |
| Tb | 6.57778E-01 | 1.65916E-01 | 8.79757E-02 | 3.87717E-01 | 1.38906E-01 | 2.15892E-01 |
| Dy | 6.77737E-01 | 1.70916E-01 | 9.04537E-02 | 3.71397E-01 | 1.36158E-01 | 2.20838E-01 |
| Ho | 6.97697E-01 | 1.75916E-01 | 9.29317E-02 | 3.55076E-01 | 1.33410E-01 | 2.25784E-01 |
| Er | 7.19860E-01 | 1.80051E-01 | 9.53712E-02 | 3.58547E-01 | 1.34373E-01 | 2.29797E-01 |
| Tm | 7.42833E-01 | 1.86636E-01 | 9.78682E-02 | 4.03143E-01 | 1.43112E-01 | 2.35933E-01 |
| Yb | 7.47181E-01 | 1.92033E-01 | 1.02161E-01 | 4.66918E-01 | 1.51294E-01 | 2.35228E-01 |
| Lu | 7.82848E-01 | 1.95506E-01 | 1.04425E-01 | 3.42997E-01 | 1.29027E-01 | 2.41330E-01 |

Table 10 Exponents of 4d correlating 3f2g1h functions for the quadruple-zeta basis sets

| Element | f | f | f | g | g | h |
|---------|-------------|-------------|-------------|-------------|-------------|-------------|
| La | 1.50709E+01 | 5.10166E+00 | 2.07185E+00 | 5.96061E+00 | 2.77155E+00 | 4.94917E+00 |
| Ce | 9.41576E+00 | 5.31807E+00 | 2.87443E+00 | 6.36594E+00 | 2.93436E+00 | 5.59503E+00 |
| Pr | 1.08644E+01 | 5.64827E+00 | 2.89254E+00 | 6.67623E+00 | 3.05153E+00 | 5.85079E+00 |
| Nd | 1.12015E+01 | 5.93605E+00 | 3.07609E+00 | 7.04605E+00 | 3.22356E+00 | 6.18198E+00 |
| Pm | 1.15838E+01 | 6.22441E+00 | 3.25395E+00 | 7.42159E+00 | 3.39717E+00 | 6.51650E+00 |
| Sm | 1.19793E+01 | 6.50331E+00 | 3.42768E+00 | 7.80434E+00 | 3.57327E+00 | 6.85558E+00 |
| Eu | 1.24323E+01 | 6.80218E+00 | 3.59873E+00 | 8.19258E+00 | 3.75128E+00 | 7.19942E+00 |
| Gd | 1.20011E+01 | 6.96116E+00 | 3.84074E+00 | 8.63359E+00 | 3.97948E+00 | 7.60134E+00 |
| Tb | 1.33864E+01 | 7.39064E+00 | 3.93549E+00 | 8.99096E+00 | 4.11513E+00 | 7.90341E+00 |
| Dy | 1.38843E+01 | 7.68712E+00 | 4.10235E+00 | 9.40153E+00 | 4.30162E+00 | 8.26447E+00 |
| Ho | 1.43933E+01 | 7.98545E+00 | 4.26875E+00 | 9.81998E+00 | 4.49138E+00 | 8.63195E+00 |
| Er | 1.49168E+01 | 8.28732E+00 | 4.43527E+00 | 1.02464E+01 | 4.68443E+00 | 9.00596E+00 |
| Tm | 1.54537E+01 | 8.59202E+00 | 4.60201E+00 | 1.06808E+01 | 4.88088E+00 | 9.38666E+00 |
| Yb | 1.60011E+01 | 8.90009E+00 | 4.76931E+00 | 1.11235E+01 | 5.08082E+00 | 9.77421E+00 |
| Lu | 1.55742E+01 | 8.91844E+00 | 4.89285E+00 | 1.16141E+01 | 5.32797E+00 | 1.02104E+01 |

correlating sets. It turns out that the the g function has a fair overlap with the outer g for 4f correlation, but the gradient with Z is quite different. A compromise must therefore be made in the selection to avoid linear dependence. Some of the options are to use the outer g from the 4f correlating set for the 5s5p correlation, and to substitute the outer g from the 4f correlating set with the g from the 5s5p correlating set, with or without reoptimization of the other g functions. All of these options result in a loss of correlation energy of a few millihartrees, so none of them is to be preferred. The final choice was to use the outer g from the 4f correlating set for the 5s5p correlation.

3.4 4d correlation

Correlating functions for the 4d shell were optimized on the ground configuration for the state with the maximum J value, with single and double excitations out of the 4d shell into the correlating space coupled to $J = 0$. The correlating space consisted of 1f for the double-zeta basis sets, 2f1g for triple-zeta, and 3f2g1h for quadruple-zeta. Again, the f functions are not strictly necessary and are not included in the basis set, but serve as a guide to the contraction pattern. The 4d correlating set is needed for La, but for the other elements, 4d correlation is

Table 11 Exponents of 4f dipole polarization functions for the quadruple-zeta basis sets

| Element | g | h | i |
|---------|-------------|-------------|-------------|
| Ce | 1.56827E-01 | 2.65133E-01 | 6.56642E-01 |
| Pr | 1.07192E-01 | 1.82024E-01 | 4.40083E-01 |
| Nd | 1.20604E-01 | 2.21280E-01 | 4.91423E-01 |
| Pm | 1.31911E-01 | 2.46529E-01 | 5.37859E-01 |
| Sm | 1.42827E-01 | 2.73000E-01 | 5.80453E-01 |
| Eu | 1.54399E-01 | 2.72644E-01 | 6.15486E-01 |
| Gd | 2.32955E-01 | 4.00971E-01 | 9.02198E-01 |
| Tb | 1.65285E-01 | 3.01969E-01 | 6.62920E-01 |
| Dy | 1.70504E-01 | 3.23509E-01 | 6.86128E-01 |
| Ho | 1.75367E-01 | 3.36026E-01 | 7.06498E-01 |
| Er | 1.79966E-01 | 3.49507E-01 | 7.26972E-01 |
| Tm | 1.84184E-01 | 3.61190E-01 | 7.45325E-01 |
| Yb | 1.88876E-01 | 3.80806E-01 | 7.65243E-01 |
| Lu | 2.91669E-01 | 5.70315E-01 | 1.12554E+00 |

adequately covered by the SCF f functions and the 4f correlating functions.

3.5 Dipole polarization

Functions for dipole polarization of the 4f shell were determined as follows. For the double-zeta basis sets, a single g function was determined for the 4f dipole polarization by maximizing the polarizability calculated by second-order perturbation theory. The basis states for the perturbation theory consisted of the eigenfunctions of the Dirac Hamiltonian for the configurations generated by a single $f \rightarrow g$ excitation from the ground configuration. For the triple-zeta and quadruple-zeta basis sets, the exponents of the dipole polarization functions were determined by multiplying the outermost correlating exponent for each angular momentum by a factor obtained from the double-zeta results. This factor is the exponent of the double-zeta dipole polarizing g function divided by the exponent of the double-zeta correlating g function. In this way 1g1h and 1g1h1i sets for 4f dipole polarization were generated for the triple-zeta and quadruple-zeta basis sets.

For the 5d dipole polarization, a single f function was determined for the dz basis sets in the same manner, using the eigenfunctions of the Dirac Hamiltonian for the configurations generated by a single $d \rightarrow f$ excitation from the $f^{n-1}d^1s^2$ configuration for the perturbation basis. The functions so generated for the dipole polarization were sufficiently similar to the correlating functions that it was not considered necessary to generate functions for the triple-zeta and quadruple-zeta basis sets.

Table 12 Configuration average total SCF energies in E_h for uncontracted basis set calculations on the weighted average of the configurations used for the contractions

| Element | Double-zeta | Triple-zeta | Quadruple-zeta |
|---------|----------------|----------------|----------------|
| La | -8,493.624083 | -8,493.630096 | -8,493.630444 |
| Ce | -8,861.048401 | -8,861.055407 | -8,861.055795 |
| Pr | -9,238.122189 | -9,238.131109 | -9,238.131553 |
| Nd | -9,625.099353 | -9,625.109635 | -9,625.110130 |
| Pm | -10,022.049991 | -10,022.061750 | -10,022.062300 |
| Sm | -10,429.109635 | -10,429.122980 | -10,429.123565 |
| Eu | -10,846.444130 | -10,846.459218 | -10,846.459892 |
| Gd | -11,274.196122 | -11,274.211188 | -11,274.211891 |
| Tb | -11,712.470974 | -11,712.490007 | -11,712.490828 |
| Dy | -12,161.458736 | -12,161.479991 | -12,161.480895 |
| Ho | -12,621.316197 | -12,621.339880 | -12,621.340875 |
| Er | -13,092.186736 | -13,092.213047 | -13,092.214143 |
| Tm | -13,574.220021 | -13,574.249160 | -13,574.250362 |
| Yb | -14,067.575100 | -14,067.607279 | -14,067.608596 |
| Lu | -14,572.478843 | -14,572.510056 | -14,572.511382 |

4 Contraction patterns

Contraction coefficients for the occupied spinors, including the 5d and 6p, were taken from SCF calculations on a weighted average of the valence configurations, as follows:

- For La, Ce, Gd, and Lu, where the 5d is occupied in the ground state, the $4f^{n-1} 5d^1 6s^2$ and $4f^{n-1} 5d^1 6p^2$ configurations were used with a 9:1 weight ratio (90% s, 10% p). Note that for La the 4f is empty: this element is treated as a transition metal.
- For the rest (Pr–Eu, Tb–Yb), the $4f^n 6s^2$, $4f^{n-1} 5d^1 6s^2$, and $4f^{n-1} 6p^1 6s^2$ configurations were used with a weight ratio of 6:3:1.

These ratios were chosen so that the primary configuration is not greatly perturbed by the inclusion of orbitals that are empty in the ground state, but still give a reasonable representation of these orbitals. Thus, a weight of 10% was chosen for 6p configurations, and a 2:1 ratio was chosen for the $4f^n 6s^2$ and $4f^{n-1} 5d^1 6s^2$ configurations. These choices are somewhat arbitrary. The higher weight of the $4f^{n-1} 5d^1 6s^2$ configuration reflects the greater importance of the 5d than the 6p.

The f functions that were added for 5d and 6s correlation have been included in the SCF set, and therefore in the contraction calculations. References to the SCF set include these functions.

The total energies for the weighted average used for generating the basis sets are given in Table 12.

To determine which primitive functions should be uncontracted, a sequence of MR-SDCI calculations was

performed on several of the elements across the row, in which different primitive functions were included in the correlating space. For each basis set size, the appropriate number of primitive functions was used in the MRCI calculations. For example, for 5s5p correlation in the triple-zeta basis sets, the correlating set was 2s2p2d1f. Excitations into the 4f shell were not considered in these calculations. The large and small component coefficients of these correlating functions were determined by diagonalizing the Fock matrix in the space that consists of the DHF occupied spinors used as contracted functions and the additional primitive functions that are to be used for correlation. The virtual functions from this diagonalization were then orthogonalized to the original DHF occupied functions, to ensure strict orthogonality. (This procedure also allows the elimination of any linearly dependent functions, but no linear dependence was observed in this work.)

The contraction pattern is described first for the outer valence shells, then the additional functions required for the inner valence are described, and finally the functions for correlation of the 4d. To any of these contractions, the relevant polarization functions listed in Tables 2, 7 and 11 can be added. In the descriptions, functions are counted by increasing exponent size, from the smallest. In cases where linear dependence might be a problem, alternative prescriptions are given.

4.1 Double-zeta basis sets

- *Outer valence 5d/6s/6p* To the SCF functions, add the second s primitive, the second p primitive, the first d primitive, and the first f primitive (the 5d/6s correlating f function for La).
- *Inner valence 4f/5s/5p* To the outer valence set, add the fourth s primitive, the third p primitive, the third d primitives, the second and third f primitives, and the correlating g function for the 4f shell.
- *Core 4d* To the inner valence set, add the fifth s primitive, the fifth p primitive, the sixth d primitive, and the fifth f primitive.

4.2 Triple-zeta basis sets

- *Outer valence 5d/6s/6p* To the SCF functions, add the first and third s and p primitives, and the first and second d primitives. For La add the correlating 2f1g functions for the 5d/6s shells; for Ce–Lu add the first and third f primitives and the 5d/6s correlating g function.
- *Inner valence 4f/5s/5p* To the outer valence set, add the fourth and sixth s primitives and the fifth and sixth p

primitives. For La add the fourth and fifth d primitives, and the 5s5p correlating f function; for Ce–Lu add the third through fifth d and f primitives, and the 4f correlating 2g1h set.

- *Core 4d* To the inner valence set, add the seventh and ninth s and p primitives, and the seventh and eighth d primitives. For La, add the correlating 2f1g functions for the 4d shell; for Ce–Lu add the sixth and seventh f primitives from the SCF set. The g function for 4d correlation is covered by the 4f correlating set.

4.3 Quadruple-zeta basis sets

- *Outer valence* To the SCF functions, add the first, second, and fourth s and p primitives, the first through third d primitives. For La add the correlating 3f2g1h for the 5d and 6s shells; for Ce–Lu add the first through third f primitives and the correlating 2g1h functions for the 5d and 6s shells.
- *Inner valence* To the outer valence set add the fifth, sixth, and eighth s primitives and the sixth, seventh and eighth p primitives. For La add the the third, fourth, and fifth d primitives and the correlating 2f1g set for the 5s and 5p shells; for Ce–Lu add the third through sixth d and f primitives, and the 4f correlating 3g2h1i set.

If linear dependence problems are encountered in the d space, do not include the second d primitive.

- *Core 4d* To the inner valence set, add the tenth through 12th s and p primitives, and the seventh through ninth d primitives. For La add the 4d correlating 3f2g1h set; for Ce–Lu add the seventh and eighth f primitives. The 4d correlating 2g1h set is well enough represented by the 4f correlating functions to be omitted.

5 Application

To exemplify the performance of the basis sets developed here in molecular calculations we have investigated the YbF molecule. This system has received a fair amount of theoretical attention because of its potential for the observation of parity-violating interactions [36, 37], in connection with the determination of the electric dipole moment of the electron (see, for instance, [38–42] and references therein).

YbF is an open-shell molecule with a $^2\Sigma^+$ ground state. Experimental [43, 44] and theoretical [45, 46] investigations indicate that the unpaired electron is located in a σ orbital with dominant contributions from the 6s orbital of Yb, corresponding to a Yb($f^{14}s\sigma$)F configuration.

5.1 Computational details

The calculations were performed with the DIRAC program suite [47], using the Dirac-Coulomb Hamiltonian. The valence double-zeta, triple-zeta, and quadruple-zeta sets from the present work were used for Yb, including the correlating functions and the functions for dipole polarization of the 4f and the 5d. The matching augmented correlation-consistent basis sets of Dunning [28] (aug-cc-pVnZ, $n = 2, 3, 4$) were used for F. All basis sets were kept uncontracted, with the small component basis generated by restricted kinetic balance. Furthermore, all two-electron integrals over small component (S) basis sets (the so-called (SSISS)-type integrals) were replaced by a simple correction [48].

In addition to the potential energy curves for individual basis sets, we have extrapolated the points on the potential energy curves for the triple- and quadruple-zeta sets to obtain a curve for the complete basis set limit (E_∞), using the relation [49]

$$E_\infty(\mathbf{R}) = \frac{4^3 E_4(\mathbf{R}) - 3^3 E_3(\mathbf{R})}{4^3 - 3^3} \quad (1)$$

where the subscripts denote the cardinal numbers for the basis sets and $E_n(\mathbf{R})$ is the potential energy for a given geometry and electronic structure method for a basis of cardinal number n ($=2, 3, 4, \infty$). It is possible to extrapolate the potential energy rather than the individual energies because the expression used to fit the energy as a function of the cardinal number n is linear in the parameter of the fit, $A(\mathbf{R})$:

$$E_n(\mathbf{R}) = E_\infty(\mathbf{R}) + A(\mathbf{R})/n^3 \quad (2)$$

Any combination of energies merely produces the same combination of the fit parameters, which can then be treated as a single parameter and the combination of energies can be extrapolated.

The $E_n(\mathbf{R})$ energies were obtained at the SCF, MP2, CCSD, and CCSD(T) levels of theory for the ground state [50], using as reference wave function that of the Yb($s\sigma_{1/2}$)F configuration. For the active occupied space, orbitals arising from the combination of the Yb 4f, 5s, 5p, 6s and the F 2s and 2p atomic orbitals were explicitly considered in the calculations. Due to computational constraints, we have truncated the active virtual space so that 117, 230 and 296 orbitals were used in the double-zeta, triple-zeta and quadruple-zeta calculations.

Spectroscopic constants (r_e , D_e , ω_e and $\omega_e x_e$) were determined from a fifth degree polynomial fit in the vicinity of the potential energy minima, which corresponded to bond lengths between 1.96 and 2.10 Å, spaced by 0.01 Å between 1.98 and 2.06 Å, and by 0.02 Å for the outer regions. In the calculation of D_e the asymptotic

dissociation limit is calculated from the energies of the isolated neutral atoms, fluorine in the $^2P_{3/2}$ state and Yb in the 1S state.

5.2 Results and discussion

The results of our calculations for the $^{174}\text{Yb}^{19}\text{F}$ isotopomer are presented in Table 13, along with the experimental results [43, 44, 51–53] and other theoretical results [39, 54–56]. The overall qualitative trend clearly shows that improvement of the basis set and correlation treatment leads to progressively better agreement with experiment for the properties under consideration, as expected from basis sets constructed in such a systematic fashion.

For the main spectroscopic constants (r_e , ω_e , and $\omega_e x_e$) the extrapolated CCSD results agree remarkably well with experiment: within 0.001 Å for the bond length, 1 cm^{-1} for the harmonic frequency, and 0.1 cm^{-1} for the anharmonicity. The agreement is better than that of the CCSD(T) results, for which the results differ from experiment by about 0.01 Å, 20 cm^{-1} , and 0.3 cm^{-1} respectively. The agreement of the CCSD results must be considered fortuitous, however, because it is clear from the perturbative triples results that the calculations are not converged. The current calculations do not include correlation of the 4d shell, and it is possible that 4d–4f correlation is significant. Truncation of the virtual space might also contribute to the discrepancy. In addition, the size of the triples correction suggests that a better treatment of the higher-order excitations is necessary to approach the experimental results for the right reasons.

Nevertheless, the current results are comparable to or better than previous theoretical predictions. The all-electron results of Nayak and Chaudhuri [39] do not include any higher angular momentum functions and therefore are not very accurate. The all-electron results of Su et al. [56] do not include higher angular momentum functions on the Yb atom but use the aug-cc-pVQZ basis set for F, which leads to a very unbalanced, and consequently inaccurate, calculation. The remaining calculations all include g functions on Yb, but no higher, and use either triple-zeta or quadruple-zeta basis sets on F. Although these calculations are also somewhat unbalanced, the bulk of the correlation of the Yb 4f is included with the inclusion of g functions. The best of these calculations appears to be the CISD + Q calculations of Cao et al. [54], which include the correlation of the 4d shell. However, given the size of the triples correction in our results, it is likely that these results also represent a fortuitous cancellation of errors. Our results are the only results to include angular momentum beyond g on the Yb atom, and since both the triple-zeta and the quadruple-zeta basis sets include higher angular momentum

Table 13 Spectroscopic properties for the ground state ($^2\Sigma^+$) of ^{174}YbF , obtained with double-, triple- and quadruple-zeta basis sets, and extrapolated to the complete basis set

| Property | Method | dz | tz | qz | Extrapolated |
|------------------------------------|-------------------------------|--------|--------|------------------------|--|
| r_e (Å) | SCF | 2.0627 | 2.0592 | 2.0585 | 2.0579 |
| | MP2 | 2.0078 | 2.0010 | 1.9959 | 1.9922 |
| | CCSD | 2.0188 | 2.0228 | 2.0196 | 2.0174 |
| | CCSD(T) | 2.0756 | 2.0399 | 2.0338 | 2.0289 |
| | Exp. | | | | 2.0161 ^e (2.0165 ^f) |
| | Cao et al. ^a | | | | 2.022 (2.034) |
| | Heiberg et al. ^b | | | | 2.03 |
| | Nayak et al. ^c | | | | 2.051 |
| ω_e (cm ⁻¹) | Su et al. ^d | | | | (i) 2.0127; (ii) 2.185 |
| | SCF | 489.5 | 491.8 | 491.9 | 492.0 |
| | MP2 | 509.8 | 512.9 | 516.9 | 519.8 |
| | CCSD | 475.1 | 497.4 | 503.2 | 507.6 |
| | CCSD(T) | 610.9 | 539.8 | 534.1 | 528.2 |
| | Exp. | | | | 506.6674 ^g |
| | Cao et al. ^a | | | | 514 (502) |
| | Nayak et al. ^c | | | | 529 |
| $\omega_e x_e$ (cm ⁻¹) | Su et al. ^d | | | | (i) 566.8; (ii) 447.98 |
| | SCF | 2.087 | 2.109 | 2.133 | 2.390 |
| | MP2 | 2.261 | 2.413 | 2.314 | 2.086 |
| | CCSD | 4.014 | 2.525 | 2.452 | 2.357 |
| | CCSD(T) | 0.488 | 1.258 | 1.269 | 1.939 |
| | Exp. | | | | 2.2451 ^g |
| | Su et al. ^d | | | | (i) 3.7885; (ii) 1.8995 |
| | D_e (kJ mol ⁻¹) | SCF | 383.68 | 382.11 | 381.41 |
| MP2 | | 503.72 | 512.91 | 515.50 | 519.81 |
| CCSD | | 473.42 | 478.56 | 484.78 | 489.32 |
| CCSD(T) | | 467.96 | 480.84 | 487.46 | 492.33 |
| Exp. | | | | | 520.59 ± 9.62 ^h |
| | | | | | 524.60 ± 9.57 ⁱ |
| Cao et al. ^a | | | | | 517.16 (503.65) |
| Heiberg et al. ^b | | | | | 465 (455) |
| Su et al. ^d | | | | (i) 665.7; (ii) 581.71 | |

^a Ref. [54], CISD + Q with a small-core PP (14s13p10d8f6g)/[6s6p5d4f3g] basis on Yb and aug-cc-pVQZ on F; BSSE-corrected in parenthesis

^b Ref. [55], CCSD(T) with a (24s21p15d10f3g)/[9s8p6d4f3g] basis on Yb and cc-pVTZ on F; BSSE-corrected in parenthesis

^c Ref. [39], RASCI with 27s27p12d8f basis on Yb and 15s10p on F; values at ground-state equilibrium geometry

^d Ref. [56], UCCSD(T) with (i) small-core PP (12s11p9d8f4g)/[9s8p6d5f4g], (ii) all-electron (29s22p16d13f)/[6s4p2d1f] bases on Yb; aug-cc-pVQZ on F

^e Ref. [43]

^f Ref. [51]

^g Ref. [44]

^h Ref. [53]; D_e calculated from the experimental D^0_0 and ω_e , $\omega_e x_e$ from Ref. [44]

ⁱ Ref. [52]

functions, it is difficult to compare directly with the other results.

The current double-zeta results should be treated with caution. Both the Yb basis and the F basis are too small for accurate calculations, so that a large part of the correlation

is missing. This is true of double-zeta calculations in general, which are usually only suitable for qualitative work. However, for YbF there is another reason. Dolg et al. [45] found that SCF calculations pointed to an $f^{13} s^2$ ground state for YbF whereas in MRCI calculation the

$f^{14} s^1$ configuration was preferentially stabilized and became the ground state. The inadequate correlating space in the present double-zeta calculations could therefore lead to a poor description of the ground state. This possibility is supported by the large triples corrections for the predicted properties, not only for the double-zeta basis, but also for the larger basis sets.

The existence of a perturbing state is verified by calculations done with the triple-zeta basis set with two different virtual spaces. The first included 198 virtuals, and the second included 230 virtuals, which corresponds to an additional s, d, f, g, and h shell. The potential energy curves relative to the minimum for the CCSD and CCSD(T) calculations are shown in Fig. 1, along with the T_1 diagnostic, as a function of r . The CCSD curve for the smaller virtual space shows an abrupt variation at around 1.9 Å that indicates the presence of another state, which is manifested by the large values of the T_1 diagnostic in the vicinity of this point. The T_1 diagnostic for the larger virtual space at these geometries is in the normal range for a single reference. It is clear from the figures, however, that the same abrupt change in the CCSD potential observed for the smaller virtual space occurs for the larger virtual space, but at a shorter bond length (about 1.78 Å).

As this perturbation occurs relatively far from the CCSD minimum and without significantly changing the shape of the potential, there should only be a small effect on the calculated spectroscopic constants. This is not the case for CCSD(T) due to the sharp rise in energy in the repulsive region, which explains why the CCSD(T) constants are further from experiment than the CCSD constants. This observation also helps to explain the poor accuracy of other calculations, notably those of Su et al. [56], who observe a shoulder between 1.7 and 2.0 Å in their pseudopotential calculations (which no doubt influences the fitting of the curve and consequently the quality of the spectroscopic constants).

These results confirm the idea that there is more than one state to consider, and that the position of these two states is strongly dependent on the degree of correlation included. The results of single-point calculations with several virtual spaces, given in Table 14, show that the bulk of the extra correlation, and the biggest change in the T_1 diagnostic, comes from the addition of a d and a g shell. Moreover, it is clear that the perturbative treatment for the triple excitations is not reliable in the regions of large T_1 values, and a more thorough treatment of higher excitations and of the reference space is mandated.

For the dissociation energies, the extrapolated CCSD(T) values are the closest to experiment, but are still 20–30 kJ mol^{-1} too low. Treatment of higher-order excitations, and a better treatment of the perturbing state, are likely to increase the binding and thus bring the results into

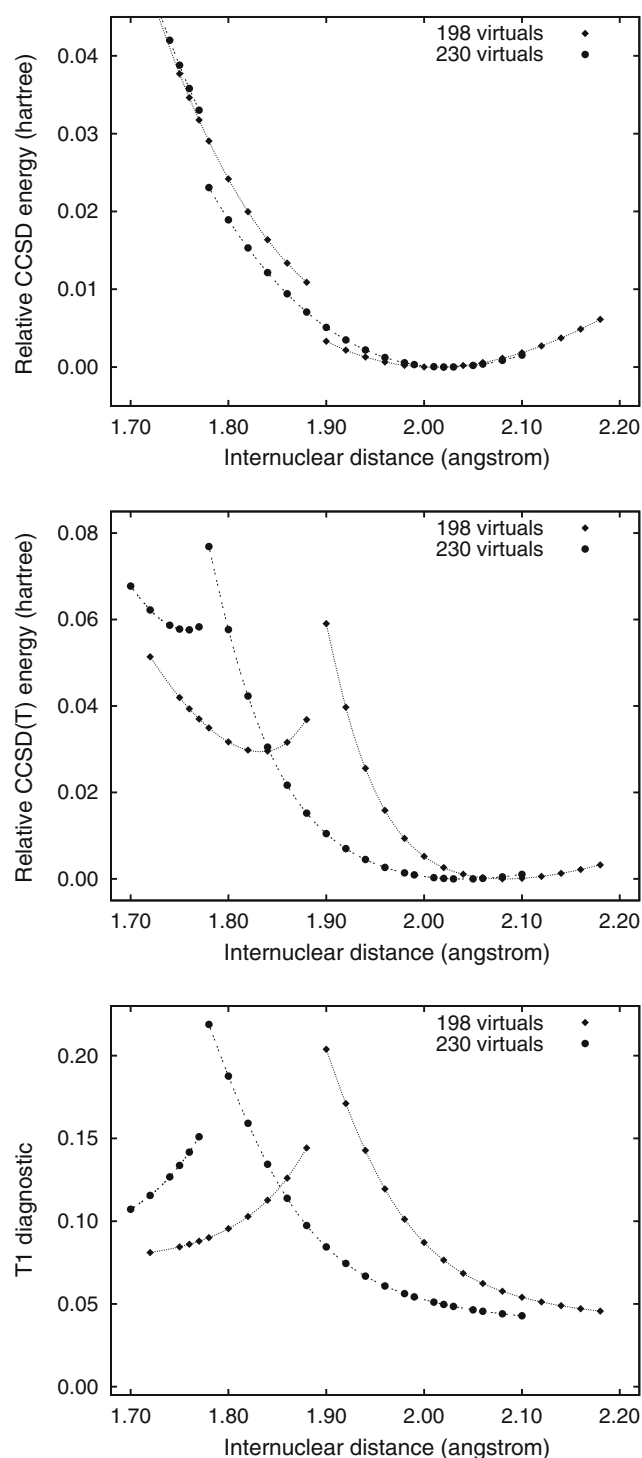


Fig. 1 CCSD and CCSD(T) potential energy curves (relative to the minima) and T_1 diagnostic in the triple-zeta basis set, using different numbers of active virtuals (198 and 230)

better agreement with experiment. The results of Cao et al. are closer to experiment than ours, which might be due to the correlation of the 4d shell in their calculations. However, improvements in the basis set and the correlation treatment are likely to move their results to a value that is

Table 14 Contributions to the electron correlation energy for the coupled cluster methods

| r (Å) | Basis | Virtual space size | CCSD | CCSD(T) |
|---------|-------|--------------------|-----------------|---------------------------|
| 1.780 | tz | 198 | −1.3485 (0.090) | −0.0507 (−0.0972, 0.0465) |
| | | +1s1d1f1g1h (=230) | −1.5212 (0.219) | −0.0133 (−0.2087, 0.1954) |
| 1.900 | tz | 198 | −1.3490 (0.204) | −0.0008 (−0.1854, 0.1846) |
| | | +1s1d1f1g1h (=230) | −1.5126 (0.067) | −0.0618 (−0.0991, 0.0373) |
| 2.000 | tz | 198 | −1.3442 (0.087) | −0.0509 (−0.0958, 0.0448) |
| | | +1h | −1.3832 (0.080) | −0.0546 (−0.0931, 0.0385) |
| | | +1d1g1h | −1.4966 (0.054) | −0.0654 (−0.0861, 0.0206) |
| | | +1s1d1f1g1h (=230) | −1.5109 (0.053) | −0.0667 (−0.0865, 0.0197) |
| | | 230 + 1s1p1d1f | −1.5183 (0.053) | −0.0673 (−0.0869, 0.0196) |

For CCSD the value of the T_1 diagnostic is given in parenthesis; for CCSD(T) the values for the fourth and fifth order triples corrections are given in parentheses (for details see Ref. [50])

too high. The fact that the BSSE in their results is about 14 kJ mol^{-1} indicates that neither the 1-particle nor the N -particle space is saturated. Although the current calculations do not account for BSSE, the effect of BSSE on our results should be smaller than on the results of Cao et al. or of Heiberg et al. Our quadruple-zeta basis set is larger than any of the others used for Yb and includes higher angular momentum functions, which will reduce the size of the BSSE relative to calculations that lack these higher angular momentum functions. The PP values of Su and coworkers are too high and are yet another indication of the poor quality of their calculations.

The experimental numbers are, however, obtained by indirect methods: the results of Ref. [52] are derived from molecular beam chemiluminescence studies, whereas those of Ref. [53] are based on thermochemical data and a ligand-field model. They are therefore rather dependent upon assumptions and approximations in the underlying models. This, together with the variations in the theoretical values, leaves some uncertainty about the magnitude of the error in the current calculations.

6 Conclusion

Basis sets of double-zeta, triple-zeta, and quadruple-zeta quality have been optimized for the lanthanide elements La–Lu, including functions for correlation of the 6s and 5d, the 4f, the 5s and 5p, and the 4d shells. The full tables of basis sets including spin-free relativistic SCF [57] and Dirac-Fock SCF coefficients are available in ASCII format from the Dirac web site, <http://dirac.chem.sdu.dk>. The spin-free relativistic SCF coefficients include the Foldy–Wouthuysen transformed large component coefficients that can be used in the scalar one-electron NESC approximation [58].

These basis sets are shown to perform well in the determination of the ground-state spectroscopic constants

of YbF, and extrapolations to the basis set limit using the triple- and quadruple-zeta basis yields results close to experiment for r_e , ω_e and $\omega_e x_e$. For D_e , the results are outside the experimental limits by some tens of kJ mol^{-1} , but should be improved by the inclusion of more outer core orbitals and a better treatment of higher excitations. The double-zeta basis sets provide qualitatively correct results, but cannot be used for high accuracy calculations, especially in difficult cases such as YbF where electron correlation effects are extremely important for correctly describing the ground-state wave function.

The importance of using proper, flexible correlating basis sets in order to allow single-reference based perturbative approaches, such as CCSD(T), to work properly in cases where dynamical correlation can strongly influence the nature of the wavefunction was shown here. In light of our results, we therefore suggest the use of at least the triple-zeta basis set, and including in the virtual space all correlation functions, in order to obtain a qualitative and quantitatively correct picture.

7 Internet archive

This paper includes an internet archive in ASCII format. The archive contains the Dirac-Fock SCF coefficients and the spin-free relativistic SCF coefficients, including the Foldy–Wouthuysen transformed large component coefficients, and the correlating and polarizing functions. Prescriptions are given in the archive for the construction of various basis sets.

Acknowledgments The authors wish to thank The Netherlands Organization for Scientific Research (NWO) for financial support via the TOP and the Vici programs. We acknowledge computer time provided by the Dutch National Computing Facilities (NCF), as well as by CINES (Centre Informatique National de l'Enseignement Supérieur) under grant ph12531.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Malli GL, Da Silva ABF, Ishikawa Y (1993) *Phys Rev A* 47:143
- Malli GL, Da Silva ABF, Ishikawa Y (1994) *J Chem Phys* 101:6829
- Fægri K Jr (2001) *Theor Chem Acc* 105:252
- Fægri K Jr (2005) *Chem Phys* 311:25
- Tsuchiya T, Abe M, Nakajima T, Hirao K (2001) *J Chem Phys* 115:4463
- Nakajima T, Hirao K (2002) *J Chem Phys* 116:8270
- Koga T, Tatewaki H, Matsuoka O (2001) *J Chem Phys* 115:3561
- Koga T, Tatewaki H, Matsuoka O (2002) *J Chem Phys* 117:7813
- Koga T, Tatewaki H, Matsuoka O (2003) *J Chem Phys* 119:1279
- Tatewaki H, Watanabe Y (2004) *J Chem Phys* 121:4528
- Osanai Y, Noro T, Miyoshi E (2002) *J Chem Phys* 117:9623
- Noro T, Sekiya M, Osanai Y, Miyoshi E, Koga T (2003) *J Chem Phys* 119:5142
- Osanai Y, Noro T, Miyoshi E, Sekiya M, Koga T (2004) *J Chem Phys* 120:6408
- Sekiya M, Noro T, Miyoshi E, Osanai Y, Koga T (2006) *J Comp Chem* 27:463
- Noro T, Sekiya M, Osanai Y, Koga T, Matsuyama H (2007) *J Comp Chem* 28:2511
- Noro T, Sekiya M, Koga T (2008) *Theor Chem Acc* 121:289
- Balabanov NB, Peterson KA (2005) *J Chem Phys* 123:064107
- Peterson KA, Puzzarini C (2005) *Theor Chem Acc* 114:283
- Gomes ASP, Custodio R, Visscher L (2006) *Theor Chem Acc* 115:398
- Dyall KG (1998) *Theor Chem Acc* 99:366
- Dyall KG (2002) *Theor Chem Acc* 108:365
- Dyall KG (2002) *Theor Chem Acc* 108:335
- Dyall KG (2006) *Theor Chem Acc* 115:441
- Dyall KG (2004) *Theor Chem Acc* 112:403
- Dyall KG (2007) *Theor Chem Acc* 117:483
- Dyall KG (2007) *Theor Chem Acc* 117:491
- Dyall KG (2009) *J Phys Chem A* 113:12638
- Dunning TH Jr (1989) *J Chem Phys* 90:1007
- Kendall RA, Dunning TH Jr, Harrison RJ (1992) *J Chem Phys* 96: 6769
- Woon DE, Dunning TH Jr (1993) *J Chem Phys* 98:1358
- Wilson AK, Woon DE, Peterson KA, Dunning TH Jr (1999) *J Chem Phys* 110:7667
- Dyall KG, Fægri K Jr (1996) *Theor Chim Acta* 94:39
- Seth M, Shepard R, Wagner A, Dyall KG (2001) *J Phys B* 34:2383
- Visscher L, Dyall KG (1997) *At Data Nucl Data Tables* 67:207
- Dyall KG, Gomes ASP (2010) *Theor Chem Acc* 125:97
- Sapirstein J (2002) Parity violation. In: Schwerdtfeger P (ed) *Relativistic electronic structure theory*. Elsevier, Amsterdam
- Berger R (2002) Parity-violation effects in molecules. In: Schwerdtfeger P (ed) *Relativistic electronic structure theory*. Elsevier, Amsterdam
- Parpia FA (1998) *J Phys B* 31: 1409
- Nayak MK, Chaudhuri RK (2006) *Chem Phys Lett* 419(1–3):191
- Titov AV, Mosyagin NS, Petrov AN, Isaev TA (2005) *Int J Quantum Chem* 104:223
- Hudson JJ, Sauer BE, Tarbutt MR, Hinds EA (2002) *Phys Rev Lett* 89:023003
- Steimle TC, Tongmei M, Linton C (2007) *J Chem Phys* 127:234316
- Huber KP, Herzberg G (2005) Constants of diatomic molecules, (data prepared by Gallagher JW and Johnson RD, III). In: Linstrom PJ, Mallard WG (eds) *NIST chemistry webBook*, NIST standard reference database number 69, June. National Institute of Standards and Technology, Gaithersburg, 20899 (<http://webbook.nist.gov>)
- Dunfield KL, Linton C, Clarke TE, McBride J, Adam AG, Peers JRD (1995) *J Mol Spect* 174: 433
- Dolg M, Stoll H, Preuss H (1992) *Chem Phys* 165:21
- Liu W, Dolg M, Li L (1998) *J Chem Phys* 108:2886
- Visscher L, Jensen HJ Aa, Saue T, with new contributions from Bast R, Dubillard S, Dyall KG, Ekström U, Eliav E, Fleig T, Gomes ASP, Helgaker TU, Henriksson J, Iliáš M, Jacob ChR, Knecht S, Norman P, Olsen J, Pernpointner M, Ruud K, Sałek P, Sikkema J (2008) DIRAC, a relativistic ab initio electronic structure program, Release DIRAC08 (see <http://dirac.chem.sdu.dk>)
- Visscher L (1997) *Theor Chem Acc* 98:68
- Helgaker T, Klopper W, Koch H, Noga J (1997) *J Chem Phys* 106:9639
- Visscher L, Lee TJ, Dyall KG (1996) *J Chem Phys* 105:8769
- Dickinson CS, Coxon JA, Walker NR, Gerry CL (2001) *J Chem Phys* 115:6979
- Yokozeiki A, Menzinger M (1976) *Chem Phys* 14:427
- Kaledin LA, Heaven MC, Field RW (1999) *J Mol Spect* 193: 285
- Cao X, Liu W, Dolg M (2002) *Sci China B* 45:91
- Heiberg H, Gropen O, Laerdahl JK, Swang O, Wahlgren U (2003) *Theor Chem Acc* 110:118
- Su T, Yang C, Wang X, Bai F, Wang M (2009) *Chem Phys Lett* 467:265
- Dyall KG (1994) *J Chem Phys* 100:2118
- Dyall KG (2001) *J Chem Phys* 115:9136