REGULAR ARTICLE

Optimized accurate auxiliary basis sets for RI-MP2 and RI-CC2 calculations for the atoms Rb to Rn

Arnim Hellweg · Christof Hättig · Sebastian Höfener · Wim Klopper

Abstract The introduction of the resolution-of-the-identity (RI) approximation for electron repulsion integrals in quantum chemical calculations requires in addition to the orbital basis so-called auxiliary or fitting basis sets. We report here such auxiliary basis sets optimized for second-order Møller–Plesset perturbation theory for the recently published (Weigend and Ahlrichs *Phys Chem Chem Phys*, 2005, 7, 3297–3305) segmented contracted Gaussian basis sets of split, triple-ζ and quadruple-ζ valence quality for the atoms Rb–Rn (except lanthanides). These basis sets are designed for use in connection with small-core effective core potentials including scalar relativistic corrections. Hereby accurate resolution-of-the-identity calculations with second-order Møller–Plesset perturbation theory (MP2) and related methods can now be performed for molecules containing elements from H to Rn. The error of the RI approximation has been evaluated for a test set of 385 small and medium sized molecules, which represent the common oxidation states of each element, and is compared with the one-electron basis set error, estimated based on highly accurate explicitly correlated MP2–R12 calculations. With the reported auxiliary basis sets the RI error for MP2 correlation energies is typically two orders of magnitude smaller than the one-electron basis set error, independent on the position of the atoms in the periodic table.

A. Hellweg · W. Klopper
Forschungszentrum Karlsruhe,
Institute of Nanotechnology,
P. O. Box 3640, 76021 Karlsruhe, Germany

A. Hellweg (✉) · C. Hättig
Lehrstuhl für Theoretische Chemie,
Ruhr-Universität Bochum,
44780 Bochum, Germany
e-mail: arnim.hellweg@theochem.ruhr-uni-bochum.de

S. Höfener · W. Klopper
Institut für Physikalische Chemie,
Universität Karlsruhe (TH), 76128 Karlsruhe, Germany

1 Introduction

The resolution-of-the-identity (RI) approximation [5, 23,32]—also denoted as density fitting (DF)—is a very efficient way to reduce the computational costs for evaluating two-electron four-center integrals over Gaussian basis functions for second-order Møller–Plesset perturbation theory (MP2) and related methods. Typically reductions by an order of magnitude or more are obtained without significant loss of accuracy. [12,29] Furthermore, the bottlenecks concerning memory and disk space demands are widened.

The basic idea behind the RI approximation is, that products of orbital basis functions χ_μ can be approximated by a linear expansion of so-called auxiliary basis functions φ^Q

$$\chi_\mu \chi_\nu \approx \sum_Q \varphi^Q c_{\mu\nu}^Q, \tag{1}$$

with expansion coefficients $c_{\mu\nu}^Q$.

An additional error Δ_{RI} is hereby introduced in the theory. It should be smaller than the basis set incompleteness error Δ_{BS}, which, again, should be smaller than the methodic error. The ascending order $\Delta_{RI} < \Delta_{BS} < \Delta_{method}$ should be preserved to allow an appropriate discussion of the results. In particular should the accuracy not be limited by Δ_{RI} to ensure that the well-established model chemistry of MP2 or CC2 theory is

left unaffected. To achieve this, auxiliary basis sets have to be designed and optimized for the conjoint use with orbital basis sets. In the last decade optimized auxiliary basis sets which fulfill the requirement $\Delta_{RI} < \Delta_{BS}$ have been reported for several basis sets including highly accurate correlation consistent basis sets from the work of Dunning and coworkers [8,11,26,27,30,31].

The recent development of balanced and accurate segmented contracted Gaussian type orbital (GTO) basis sets [28] makes ab initio and density functional theory (DFT) calculations from hydrogen to radon (excluding the lanthanides) possible with the same quality across the periodic table. These basis sets are all-electron bases from H to Kr and for the elements from Rb to Rn the basis sets were optimized to be used with small core scalar relativistic effective core potentials (ECPs). Our goal here was to put auxiliary basis sets for these orbital basis sets at disposal that keep the balance between speed and accuracy.

2 Optimization of the auxiliary basis sets

The theoretical background to optimize auxiliary basis sets for RI-MP2 calculations is given in Ref. [30]. The error of the RI approximation is the difference between a RI-MP2 and a conventional MP2 calculation

$$\Delta_{RI} = E_{RI\text{-}MP2} - E_{MP2}. \qquad (2)$$

It consists of a linear and a quadratic contribution. The latter is negative semi-definite and is used for a variational optimization process. The functional that is minimized is

$$\delta_{RI} = \frac{1}{4} \sum_{aibj} \frac{(\langle ab||ij \rangle_{RI} - \langle ab||ij \rangle)^2}{\varepsilon_a - \varepsilon_i + \varepsilon_b - \varepsilon_j}, \qquad (3)$$

with $\langle ab||ij \rangle = (ai|bj) - (aj|bi)$, where i, j denote the occupied, and a, b the virtual SCF spin orbitals and ε_p are the SCF orbital energies. The functional δ_{RI} can be understood as the sum of the squared errors in the electron repulsion integrals $\langle ab||ij \rangle$, weighted by the SCF orbital energy differences to ensure that the auxiliary basis reproduces in particular those integrals with high accuracy which contribute most to the first-order wavefunction and the MP2 correlation energy.

In Ref. [11] the implementation of an analytic gradient for δ_{RI} with respect to exponents and contraction coefficients of the auxiliary basis in the `ricc2` module [12] in the TURBOMOLE package [1] was reported. This program has been used here together with the optimization module `relax`.

We optimized auxiliary basis sets for the new developed or modified Gaussian basis sets by Weigend and Ahlrichs [28] denoted as def2-SVP, def2-TZVP, def2-TZVPP, and def2-QZVPP (where the latter is synonymous with QZVPP). No auxiliary basis sets were designed for orbital basis sets without polarization functions, because they are not suitable for correlated calculations. The def2-SVP and def2-TZVP basis sets have quite a large remaining basis set incompleteness error (BSIE) in MP2 or CC2 calculations, since they only cover about half of the correlation energy, and should thus only be used for qualitative investigations. With def2-TZVPP and def2-QZVPP the BSIE should be small enough to obtain quantitative results at MP2 and CC2 level.

The number of basis functions per angular momentum and the exponents of the functions depend strongly on the orbital basis set. Following the specifications of previously developed auxiliary basis sets [11,30,31] a number of 2–4 times more primitive GTOs than for the orbital bases was found reasonable. The optimization process itself is nonlinear and the quality of the auxiliary bases has to be examined carefully. The highest angular momentum has to be at least $l_{occ} + l_{bas}$, where l_{occ} is the highest angular momentum occupied in the SCF reference wavefunction for the respective atom and l_{bas} the highest angular momentum of the underlying orbital basis set. The sizes of the auxiliary basis sets are assorted in Table 1 and compared with those of the orbital basis sets. For the def2-SVP basis set the ratio N_{aux}/N_{orb} lies between 2.4 and 4.7, while the values for the def2-QZVPP basis set are much closer together, between 2.9 and 3.4. For smaller basis sets one needs, due to the incompleteness of the spanned space, more auxiliary functions per orbital basis function to resolve appropriately the identity.

For the def2-TZVPP and def2-QZVPP auxiliary bases no contractions were used, because it was found that they were mostly redundant. For the two smaller auxiliary basis sets some contracted functions are used, since here auxiliary bases from Ref. [30] for similar orbital basis sets have been taken as starting point and were just roughly reoptimized, keeping the contraction scheme unchanged, if possible.

As in previous publications [11,31] the frozen core approximation was employed in the optimization process with relatively small cores: from Na to Ar a helium core was used and a neon core from K to Kr. For the elements beyond Kr no orbitals have been frozen in addition to those included in the ECP. A compilation of the ECPs used from Rb to Rn is given in Table 2. In a typical MP2 or CC2 calculation usually larger frozen cores are used for efficiency reasons. By the conservative

Table 1 The number of GTOs and their contractions in the orbital bases and the new developed auxiliary bases

	Orbital basis	Auxiliary basis	N_{aux}/N_{orb}
def2-SVP			
Li	(7s3p)/[3s2p]	(9s5p3d)/[5s5p1d]	2.8
Na	(10s6p1d)/[4s2p1d]	(10s8p5d2f)/[7s5p3d2f]	3.4
Mg	(10s7p1d)/[4s3p1d]	(10s8p5d1f)/[7s5p3d1f]	2.4
K	(14s9p2d)/[5s3p2d]	(11s8p4d2f)/[8s6p4d2f]	2.5
Ca	(14s9p4d)/[5s3p2d]	(11s8p6d2f)/[8s8p6d2f]	3.2
Sc–Zn	(14s9p5d1f)/[5s3p2d1f]	(9s9p6d4f3g2h)	4.6
Rb	(7s6p2d)/[5s3p2d]	(9s8p5d2f)	3.0
Y–Cd	(7s6p5d1f)/[5s3p2d1f]	(8s7p6d5f3g1h)	4.3
In–I	(10s7p6d)/[4s4p2d]	(9s8p7d4f2g)	4.4
Cs	(7s6p2d)/[5s3p2d]	(9s8p4d2f)	2.8
La, Hf–Hg	(7s6p5d1f)/[6s3p2d1f]	(8s8p7d5f3g1h)	4.4
Tl–At	(10s7p6d)/[4s4p2d]	(9s8p7d5f2g)	4.7
def2-TZVP			
Al–Ar	(14s9p3d1f)/[5s5p2d1f]	(10s7p7d5f2g)/[9s7p6d5f2g]	3.1
Sc–Cu	(17s11p7d1f)/[6s4p4d1f]	(11s9p8d6f5g2h)/[9s8p8d6f5g2h]	4.0
Zn	(17s12p7d1f)/[6s5p4d1f]	(11s10p8d6f5g2h)/[9s9p8d6f5g2h]	3.9
Rb, Sr	(7s7p3d)/[6s4p3d]	(9s8p5d3f)	2.4
Y–Cd	(7s7p5d1f)/[6s4p3d1f]	(8s7p6d5f3g2h)	3.6
Cs	(7s6p3d)/[5s3p3d]	(9s8p4d2f)	2.3
La, Hf–Au	(7s7p5d1f)/[6s4p3d1f]	(8s8p7d5f3g2h)	3.8
Hg	(8s8p6d1f)/[6s5p3d1f]	(8s8p7d5f3g2h)	3.5
def2-TZVPP			
Li, Be	(11s3p1d)/[5s3p1d]	(9s6p3d1f)	2.6
Na	(14s8p3d)/[5s4p3d]	(10s8p6d5f)	3.1
Mg	(14s8p3d)/[5s4p3d]	(10s8p6d4f)	2.9
Al–Ar	(14s9p3d1f)/[5s5p3d1f]	(11s8p7d5f2g)	2.9
K	(17s11p3d)/[6s4p3d]	(11s8p7d4f)	3.0
Ca	(17s12p4d)/[6s5p3d]	(11s8p7d5f)	3.0
Sc–Cu	(17s12p7d2f1g)/[6s5p4d2f1g]	(9s9p8d7f5g3h1i)	3.4
Zn	(17s12p7d2f1g)/[6s5p4d2f1g]	(9s10p8d7f5g3h1i)	3.4
Rb, Sr	(7s7p3d1f)/[6s4p3d1f]	(9s8p6d4f2g)	2.7
Y–Cd	(7s7p5d2f1g)/[6s4p3d2f1g]	(8s7p6d5f4g3h1i)	3.1
In–Xe, Tl–Rn	(10s9p8d2f)/[6s5p3d2f]	(9s8p8d6f4g3h)	3.7
Cs	(7s6p3d1f)/[5s3p3d1f]	(9s8p6d4f2g)	3.0
Ba	(7s7p5d1f)/[6s4p3d1f]	(9s8p6d4f2g)	2.7
La, Hf–Au	(7s7p5d2f1g)/[6s4p3d2f1g]	(8s8p7d5f4g3h1i)	3.3
Hg	(8s8p6d2f1g)/[6s5p3d2f1g]	(8s8p7d5f4g3h1i)	3.1
def2-QZVPP			
Rb, Sr	(8s8p5d3f)/[7s5p4d3f]	(10s9p8d7f6g)	2.9
Y–Nb	(8s8p7d3f2g)/[7s5p4d3f2g]	(10s9p8d7f5g5h2i)	3.1
Mo	(9s8p7d3f2g)/[7s5p4d3f2g]	(10s9p8d7f5g5h2i)	3.1
Tc–Cd	(9s8p7d4f2g)/[7s5p4d4f2g]	(10s9p8d7f5g5h2i)	2.9
In–Xe	(17s14p11d4f1g)/[7s6p4d4f1g]	(10s10p9d8f6g5h2i)	3.4
Cs	(8s7p5d3f)/[6s5p3d3f]	(10s9p8d7f5g)	3.0
Ba	(8s8p5d3f)/[7s5p3d3f]	(10s9p8d7f5g)	2.9
La	(9s8p6d3f2g)/[7s5p4d3f2g]	(10s9p8d7f5g5h2i)	3.1
Hf–W	(10s8p6d3f2g)/[7s5p4d3f2g]	(10s9p8d7f5g5h2i)	3.1
Re–Hg	(10s8p6d4f2g)/[7s5p4d4f2g]	(10s9p8d7f5g5h2i)	2.9
Tl–Rn	(17s14p10d4f1g)/[7s6p4d4f1g]	(10s10p9d8f6g4h2i)	3.2

settings used here, we want to ensure that the auxiliary basis sets also work for special requests. For the test sets of molecules used for benchmarking the final basis sets more realistic frozen cores are employed as listed in Table 3.

2.1 Auxiliary split valence basis sets

Weigend and Ahlrichs [28] reported new orbital basis sets of split valence quality for the elements Li, Na, Mg, K, Ca, Sc–Zn, Rb, Y–I, Cs, La, and Hf–At.

Table 2 Sizes and types of the ECPs for the def2-SVP, def2-TZVP, def2-TZVPP and def2-QZVPP basis sets

Elements	Electrons	Configuration	Theoretical level	References
Rb–Cd	28	[Ar] $3d^{10}$	MWB	[2,13,18]
In–Xe	28	[Ar] $3d^{10}$	MDF	[20,21]
Cs–La	46	[Pd]	MWB	[4,13,18]
Hf–Hg	60	[Pd] $4f^{14}$	MWB	[2]
Tl–Rn	60	[Pd] $4f^{14}$	MDF	[20,21]

The theoretical level refers either to Wood–Boring (WB) quasi-relativistic or Dirac–Fock (DF) relativistic pseudopotentials, the letter M indicates that the ECPs have been optimized for the neutral atoms

Table 3 Electrons included in the frozen core for the benchmark calculations

Elements	Electrons	Configuration
H–Be	0	–
B–Mg	2	He
Al–Zn	10	Ne
Ga–Kr	18	Ar
Rb–Cd	28	[Ar] $3d^{10}$
In–Xe	36	[Kr]
Cs–La	46	[Pd]
Hf–Hg	60	[Pd] $4f^{14}$
Tl–Rn	68	[Xe] $4f^{14}$

For the alkali and alkaline earth metals Li, Na, Mg, K, and Ca previously published [30] auxiliary SVP basis sets were used as starting point and the contraction schemes for the s and p functions were kept. For Mg a d function was decontracted and one f function was added, for Na a d function was decontracted and two f functions were added, and for K and Ca all d functions were decontracted and two f functions were added.

For the other atoms all exponents from the previously published [30] auxiliary SVP basis sets have been decontracted. From Sc–Zn two redundant steep s functions were removed and two new h functions added. For Rb one d and two f functions were added. From Y–Cd one g function was removed and one h function added.

For the main group elements In–I the new orbital basis sets use ECPs with a smaller core and thus require auxiliary basis sets with some more steep functions to account for the inner shells: two steep s, three steep p, two steep d, and two new g functions were added. The same approach was used for the heavier homologous, Tl–Bi, but with one more f function. No auxiliary basis sets have yet been available for Po and At. Here, the optimization was started with exponents extrapolated from those for Tl–Bi.

The def2-SVP auxiliary set for Cs differs from the SVP set by two additional f functions and for La and Hf–Hg one d and one h function were added and one g function removed.

2.2 Auxiliary triple-ζ valence basis set

New def2-TZVP basis sets have been reported [28] for the elements: Al–Ar, Sc–Zn, Rb–Cd, Cs, La, and Hf–Hg and new def2-TZVPP basis sets for: Li, Be, Na–Zn, Rb–Rn (excluding the lanthanides).

For the elements from Al to Ar the def2-TZVPP and def2-TZVP orbital basis sets in Ref. [28] were put together from the s and p functions of the TZVPP and TZVP basis sets [22], while the d and f polarization functions were taken from the cc-pV(T + d)Z basis set, [6,7] with the two steepest d functions contracted for the def2-TZVP basis sets. Following a similar prescription, we constructed the respective auxiliary basis set from the s, p, and d functions of the TZVPP and TZVP auxiliary sets reported in Ref. [30] and the f and g functions of the cc-pV(T + d)Z auxiliary sets from refs. [11,31] without any further reoptimization.

Also for the auxiliary def2-TZVP basis sets for Sc–Zn, the s and p kernels of the TZVP sets from Ref. [30] were taken with unchanged contractions, while one d, two f, and two h functions were added and one g function was removed. Only for Zn one additional diffuse p functions is used.

For the remaining atoms the new triple-ζ auxiliary basis sets are uncontracted. The auxiliary basis used as starting points [30] for the optimization have been decontracted or have already been uncontracted.

For the alkali and alkaline earth metals Li, Be, Na, Mg, and Ba as well as for the main group elements In–Xe no def2-TZVP orbital basis sets were developed, as they are identical with the def2-TZVPP (for Li with the TZVPP) basis sets.

For the auxiliary triple-ζ valence basis set from Sc–Rn mainly diffuse functions with high angular momenta were needed.

Again, the new orbital basis sets for the main group elements of the fourth row, In–I, are optimized for ECPs with a smaller core. Thus, additional steep s, p, and d functions had to be added for the respective auxiliary basis sets. For Tl–Rn the exponents of the fourth

row homologes were used as starting point for the optimization.

2.3 Auxiliary quadruple-ζ valence basis set

For the def2-QZVPP or QZVPP basis sets new auxiliary sets were developed for the atoms Rb–Rn; for H–Kr these are available from Ref. [11]. Since the inner core electrons are described by ECPs, it was found, that a few steep s, p, and d functions are sufficient in the auxiliary basis. In Fig. 1 the exponents in the orbital and auxiliary basis of Rn are plotted to illustrate the typical range and number of exponents in the orbital and auxiliary basis sets with ECPs. It can be seen, that not the whole space that would be spanned by the primitives of the orbital basis functions is covered here by auxiliary basis functions. Yet, this has no influence on the RI error, since these steep primitive GTOs contribute only with very small coefficients to the contracted GTOs.
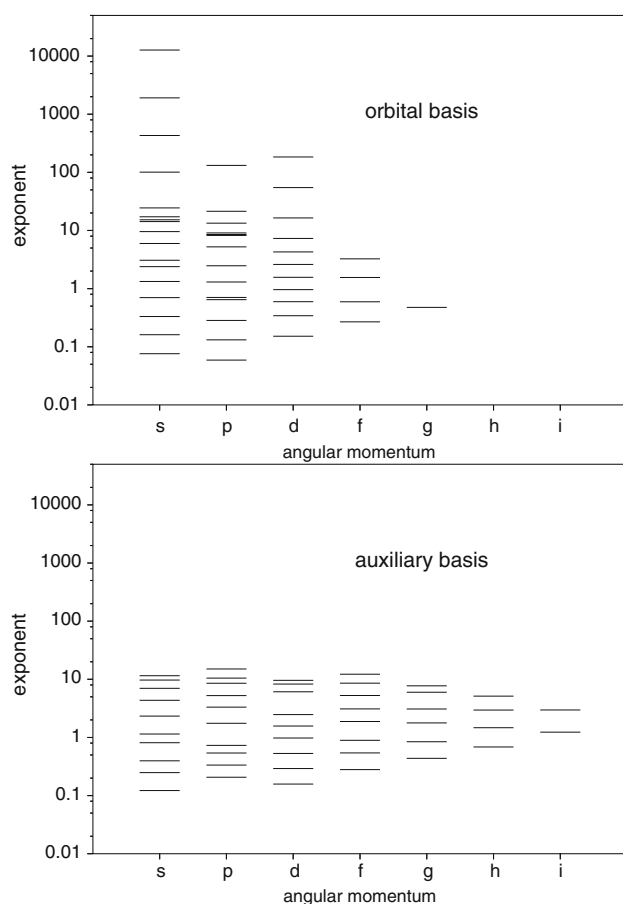

Fig. 1 The exponents for the def2-QZVPP orbital (*upper*) and auxiliary basis set (*lower*) of Rn

3 Evaluation of the auxiliary basis sets

3.1 Auxiliary basis set (RI) errors

It is of prime importance that the error Δ_{RI} is kept small not only for isolated atoms but also for molecules. To benchmark the accuracy in calculations on molecules, a test set of diatomic and small polyatomic molecules has been collected in Ref. [30] to represent all common oxidation states for each element. The structures were optimized with density functional theory on B-P86/SV(P) level. We extended this collection by a few transition metal complexes and metal and rare gas dimers. The test set of 385 molecules is listed in Table 4. When Δ_{RI} was found unacceptably large in a test case, the auxiliary basis for the respective atom was changed and reoptimized.

The frozen core approximation was applied in these benchmark calculations with the sizes of the cores corresponding up to the third row to the configuration of the previous rare gas atom. In the fourth and fifth row the frozen cores are almost identical to the sizes of the ECPs, only for the main group elements In–Xe and Tl–Rn we have included in addition the sub-valence s and p shells in the core. For full details see Table 3.

The HF wavefunctions have been calculated with the `dscf` program. [10] Tight convergence thresholds were applied to exclude numerical noise. The MP2 calculations were performed with the `mpgrad` program [9] and the RI-MP2 calculations with the `ricc2` program. [12]

In Table 5 the relative RI errors of the benchmark calculations are summarized. It can be seen, that the error decreases with the increase of the basis set as indented. Also, the errors for one auxiliary basis set are in the same order of magnitude across one row of the periodic table. Besides calculations of orbital basis sets together with their specific optimized auxiliary basis set, we also investigated the error of the def2-TZVP orbital basis set with the def2-TZVPP auxiliary basis set and the def2-QZVP orbital basis set with the def2-QZVPP auxiliary basis set. The first case was investigated to check if the quality of the def2-TZVPP calculation can be reached with the higher quality basis set, the second case, because no auxiliary basis set has been developed for the def2-QZVP orbital basis set. Normally the def2-TZVP or def2-QZVP orbital basis set are not good choices for a correlated calculation, but if it is needed for certain applications, the RI error changes systematically and fits properly into the basis set hierarchy. This hierarchy is illustrated in Fig. 2, where the distributions of RI errors for the test set are plotted for the four basis sets def2-SVP, def2-TZVP, def2-TZVPP, and def2-QZVPP.

Table 4 The test set of molecules

Atoms	Molecules
H, He	H_2, He_2
Li, Be	Be_2F_4, Be_2H_4, Be_4, $Be(CH_3)_2$, $BeF_2(OH_2)_2$, BeH_2, BeS, PLi_3, Li_2O, $Li_4(CH_3)_4$, Li_4Cl_4, Li_4H_4, Li_8, $LiBH_4$, $LiCl$, LiF, LiH, $LiSLi$, Li_2
B–Ne	B_2H_6, $B_3N_3H_6$, B_4H_4, BF_3, BH_3, BH_3CO, BH_3NH_3, C_2H_2, C_2H_3N, C_2H4, C_2H_6, C_4H_4, C_6H_6, CF_4, CH_2O, CH_2O_2, CH_3N, CH_3OH, CH_4, CO, CO_2, F_2, H_2, H_2CO_3, H_2O, H_2O_2, HCN, HF, HNC, HNO, HNO_2, HNO_3, He_2, N_2, N_2H_2, N_2H_4, N_4, NF_3, NH_3, NH_4F, Ne_2, OF_2
Na, Mg	Mg_4, $MgCl_2$, MgF, MgF_2, MgH_2, Na_2, Na_2O, Na_2S, Na_3N, Na_3P, $NaCl$, NaF, NaH
Al–Ar	Al_2O_3, Al_2S_3, $AlCl_3$, AlF_3, AlH_3, AlN, Ar_2, CS_2, Cl_2, ClF, ClF_3, H_2SO_4, H_3PO_4, HCP, HCl, H_2S, H_2S_2, P_2, PF_3, PF_5, PH_3, S_2, S_5, SF_2, SF_4, SF_6, $SiCl_2$, SiF_4, SiH_4, SiO_2, SiS_2
K, Ca	Ca_4, $CaCl_2$, CaF_2, CaH_2, K_2, K_2S, K_3P, KBr, KCl, KF, KH
Sc–Zn	$CoCl_2$, $CoCl_3$, CoF, CoF_2, CoF_3, $(CoF_6)^{3-}$, $(CoO_4)^{3-}$, CoO, $CrCl_3$, $Cr(CO_6)$, CrF_3, CrH, CrO_3, Cu_2, Cu_2O, Cu_2S, $CuCl$, $CuCl_2$, $CuCN$, CuF, CuH, $Fe(CO)_5$, FeF_2, FeF_3, FeO, Ferrocene, MnF_2, MnF_3, MnO, MnO_2, MnO_3F, MnO_4^-, MnO_3^-, MnO_4^{2-}, MnS, Ni_2, $NiCl_2$, $Ni(CO)_4$, NiF_2, NiF_3, NiO, NiS, $ScCl_3$, ScF, ScF_3, ScH_3, ScO, $TiCl_4$, $Ti(CO)_4$, TiF_3, TiF_4, TiH_4, TiO, TiO_2, TiS_2, VH_5, VO, VOF_3, $ZnCl_2$, Zn_2, ZnF_2, ZnH_2, $Zn(CH_3)_2$
Ga–Kr	As_4, $AsCl_3$, $AsCl_6^-$, AsH_3, As_4S_4, Br_2, $BrCl$, BrO^-, BrO_2^-, BrO_3^-, BrO_4^-, $GaCl$, $GaCl_3$, GaF, GaF_2, GaH_3, GaO, $GeCl_4$, GeF_3, GeF_4, GeH_4, GeO, GeO_2, Kr_2, Se_8, SeH_2, SeO, SeO_2, HBr, $HCBr_3$
Rb, Sr	Rb_2, RbF, RbH, RbO, Sr_4, SrF, SrF_2, SrH_2, SrO, SrS
Y–Cd	YF, YF_3, YO, ZrF, ZrF_3, ZrO, ZrO_2, NbF_3, NbH, NbO, NbO_2, NbO_2F, Mo_2, $MoCl_2$, $Mo(CO)_6$, MoF_3, MoH, MoO_2, MoO_3, Tc_2O_7, TcH, TcO, TcO_2, TcO_3, TcO_3F, $RuCl_2$, $RuCl_3$, $Ru(CO)_5$, RuF, RuF_2, RuF_3, RuF_6, RuO, RuO_2, RuO_4, $RhBr_3$, RhF, RhF_4, RhF_5, RhF_6, RhO, Pd_2, $PdBr_2$, $PdCl_4^{2-}$, $Pd(CO)_4$, PdF, PdO_2, PdO_3, Ag_2, $AgCl$, AgF_2, AgF_4^-, Cd_2, Cd_2Cl_2, CdF_2, $Cd(CH_3)_2$
In–Xe	$InCl$, $InCl_3$, InH, InH_3, InO, SnF_3, SnH_4, SnO, SnO_2, $SbCl_6^-$, SbF, SbF_3, SbH_3, SbO_2, TeF_3, TeH_2, TeO, TeO_2, I_2, IH, IO_4^-, ICl, IK, IO_3^-, XeF_2, XeF_4, $XeOF_4$
Cs, Ba	Cs_2, CsF, CsH, CsO, Ba_4, BaF, BaF_2, BaH_2, BaO, BaS
La, Hf–Hg	$LaCl_3$, LaF, LaF_3, LaH_3, LaO, HfF, HfF_3, HfO, HfO_2, TaF, TaF_3, TaO_2F, $W(CO)_6$, WF_3, WH, WO, WO_2, WO_3, Re_2O_7, ReH, ReO, ReO_2, ReO_3, ReO_3F, $Os(CO)_5$, OsO_2, OsO_3, OsO_4, $OsOF_5$, $IrCl$, $IrCl_3$, IrF_4, IrF_6, $IrOF_3$, Pt_2, $Pt(CO)_4$, PtO, PtO_2, Au_2, Au_3^-, $AuCl$, $AuCl_3$, AuF_5, AuO, Hg_2, Hg_2Cl_2, HgF_2, $Hg(CH_3)_2$
Tl–Rn	$TlCl$, $TlCl_3$, TlH, TlH_3, TlO, PbF_3, PbH_4, PbO, PbO_2, $BiCl_6^-$, BiF, BiF_3, BiH_3, BiO_2, $PoBr_2$, $PoCl_4$, PoF_2, PoF_6, PoO_2, At_2, $AtCl$, AtF_3, $AtNb$, Rn_2, RnF_2

In Fig. 3 the standard deviations of the RI error are divided into blocks of element groups. It can be seen from this graph, that the RI errors for the alkali and alkaline earth metal compounds with small basis sets are by far the largest. The reason for this are relatively large RI errors for the metal dimers, clusters and hydrides included in the test set. As already pointed out in Ref. [11] these compounds are particularly difficult cases for the RI approximation. Their unusual diffuse charge distribution would require more diffuse auxiliary basis sets to obtain the same accuracy as for the other molecules. However, it should be kept in mind that also the respective orbital basis sets have not been optimized for these oxidation states. Furthermore, in particular the smaller def2-SVP and def2-TZVP basis sets do not cover a big proportion of the correlation energy for

these atoms, and thus the RI errors are still orders of magnitudes smaller than the remaining BSIEs. For the def2-TZVPP and def2-QZVPP basis sets the errors are systematically smaller throughout the periodic table.

3.2 Orbital basis set error in MP2 valence correlation energies

For an analysis of correlation effects high quality basis sets are essential. The one-electron basis set errors in the MP2 valence correlation energy has been investigated in refs. [11,31] for main group elements in rows 1–3 by estimating the basis set limits by extrapolation of the correlation energies obtained with the hierarchies of correlation consistent basis sets developed by Dunning and coworkers. [6,14,33,34] The split valence and double-ζ

Table 5 The relative RI errors (in %) obtained with different auxiliary basis sets subclassified for the rows in the periodic table

Auxiliary basis set		H–Ne	Na–Ar	K–Kr	Rb–Xe	Cs–Rn	All
def2-SVP							
	Δ_{std}	0.079	0.143	0.046	0.021	0.068	0.071
	$|\Delta|$	0.052	0.071	0.019	0.012	0.022	0.029
	$|\Delta|_{max}$	0.499	0.599	0.285	0.104	0.557	0.599
def2-TZVP							
	Δ_{std}	0.042	0.033	0.018	0.026	0.024	0.028
	$|\Delta|$	0.029	0.018	0.008	0.010	0.010	0.013
	$|\Delta|_{max}$	0.222	0.123	0.153	0.162	0.147	0.222
def2-TZVP/def2-TZVPP							
	Δ_{std}	0.015	0.013	0.010	0.012	0.014	0.013
	$|\Delta|$	0.012	0.009	0.006	0.006	0.006	0.007
	$|\Delta|_{max}$	0.083	0.041	0.074	0.074	0.080	0.083
def2-TZVPP							
	Δ_{std}	0.014	0.014	0.009	0.010	0.011	0.011
	$|\Delta|$	0.013	0.010	0.006	0.006	0.006	0.007
	$|\Delta|_{max}$	0.053	0.051	0.060	0.054	0.059	0.060
def2-QZVP/def2-QZVPP							
	Δ_{std}	0.004	0.006	0.005	0.004	0.004	0.004
	$|\Delta|$	0.006	0.004	0.004	0.004	0.004	0.004
	$|\Delta|_{max}$	0.021	0.025	0.016	0.013	0.017	0.025
def2-QZVPP							
	Δ_{std}	0.004	0.006	0.002	0.003	0.004	0.004
	$|\Delta|$	0.006	0.004	0.001	0.002	0.002	0.002
	$|\Delta|_{max}$	0.021	0.025	0.004	0.014	0.028	0.028

Listed are the standard deviation $\Delta_{std} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\Delta_i - \overline{\Delta})^2}$, the mean absolute error $|\Delta| = (1/n)\sum_{i=1}^{n}\sqrt{\Delta_i^2}$, and the maximum of the absolute error $|\Delta|_{max} = \max_i|\Delta_i|$ with $\Delta_i = (E^{corr} - E^{corr,RI})/E^{corr} \times 100$

valence basis sets were found to cover for the atoms H, He, B–Ne and Al–Ar about 65%, while basis sets of triple-ζ or quadruple-ζ quality cover, respectively, about 85 and 95% of the correlation energy. For compounds with the third row elements Ga–Kr the errors are somewhat larger, which is related to a less accurate description of the 3d shell.

For the elements Rb–Rn this approach is not possible, since for most of these atoms systematic hierarchies of correlation consistent basis sets, which would allow an accurate extrapolation of the basis set limits, are not yet available. A further complication is that for strict comparison the basis set limits would be needed for the same ECPs. Therefore, we have estimated the remaining basis set errors of the segmented contracted basis sets for the elements Rb–Rn by carrying out highly accurate explicitly correlated MP2-R12 calculations for the MP2 valence correlation energies of the atoms. [15,16] (For a recent review of the MP2-R12 method, see Ref. [17].) For these calculations the same ECPs have been used and the same number of electrons have been frozen as in the molecular benchmark calculations and detailed in Tables 2 and 3.

The calculations have been carried out with a local implementation of RI-MP2-R12/A [24,25] theory into the ricc2 program for both closed shells [25] and open shells [3]. In this local version of TURBOMOLE, the RI-MP2-R12/A method has been implemented in exactly the same manner as proposed by Manby. [19] In the R12 standard approximation A, the commutator of the Fock operators $\hat{f}_1$ and $\hat{f}_2$ (for electrons 1 and 2) with the interelectronic distance r_{12} is approximated by considering only the commutator with the kinetic energy operators ($\hat{T}_1$ and $\hat{T}_2$),

$$\left[\hat{f}_1 + \hat{f}_1, r_{12}\right] \approx \left[\hat{T}_1 + \hat{T}_1, r_{12}\right]. \tag{4}$$

In the present work, the standard approximation A implies that the commutator of the ECPs (contained in $\hat{f}_1$ and $\hat{f}_2$) with r_{12} is neglected. Since the error introduced by the standard approximation A is much smaller than the error of the orbital basis set, the percentage of the recovered correlation energy given in Table 6 is not affected. Preliminary test calculations on the Zn atom, in which the ECP commutator had been included in the same manner as the exchange operator in standard approximation B of R12 theory, have indicated that the error introduced by neglecting the ECP commutator is substantially smaller than the error from neglecting the commutator with the exchange operator. In the
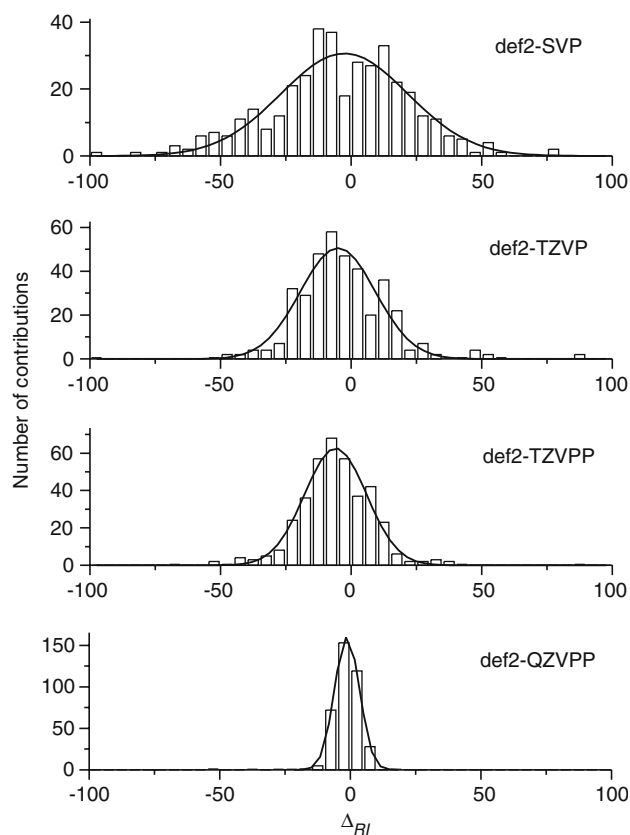
Fig. 2 Distribution of the error Δ_{RI} in μ Hartree with the different auxiliary basis sets. A Gaussian function $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$ is fitted to the distribution

large basis sets used in the present work, the difference between the standard approximations A and B is small and the ECP commutators can be safely neglected. Nevertheless, the basis set dependence of this approximation will be investigated in detail in further work.

The orbital basis sets used for the calculations of the RI-MP2-R12/A energies were derived from the

def2-QZVPP orbital basis sets for the atom of interest. The ECP remained unchanged in order to keep comparability. A criterion for reliable R12 calculations is the ratio of the conventional MP2 energy and the R12 energy correction that is added to it. High quality results are usually obtained when the ratio R12/MP2-R12 is below 10%. The basis sets are adjusted to fulfill this criterion. The def2-QZVPP basis set was fully decontracted and close-lying exponents were removed. This caused the basis sets to have different size for different atoms, but the number of basis functions only varied within a small range. After decontracting, several functions were added, mostly f, g and h type Gaussian basis functions for p elements, as well as some s, p and d functions for s and d elements. For 5s, 4d, 6s and 5d elements, the supplementary functions are [11s5p5d9f8g8h], for 5p and 6p elements they are [4f5g6h]. Most added basis functions have diffuse character since steep functions should vanish in the ECP. Plots of Gaussian exponents on a logarithmic scale were used for an overview over the space which is spanned by the exponents. If some regions were not saturated satisfyingly, additional functions were added, especially s, p, and d type Gaussians. In some cases, no symmetry could be imposed to obtain the selected atomic state. In these cases, the basis was decreased a little and the selected state was reached without significant loss in accuracy.

A complete uncontracted auxiliary basis (unpublished) from an all electron calculation of Kr [4] [40s29 p24d19f19g15h13i11k] was chosen as auxiliary basis for all atoms from Rb–Rn. Since, the number of basis functions in the latter was larger than in the calculations using ECPs, the auxiliary basis was not reoptimized.

The MP2-R12 correlation energies are listed in Table 6. Comparing these values with the correlation energies from MP2 calculations with the def2-SVP, def2-TZVPP, and def2-QZVPP basis sets, their one-electron basis set errors an be determined. The def2-SVP basis

Fig. 3 The standard deviations of the relative Δ_{RI} of the auxiliary basis sets for main group elements, transition metals, and alkali, alkaline earth metals
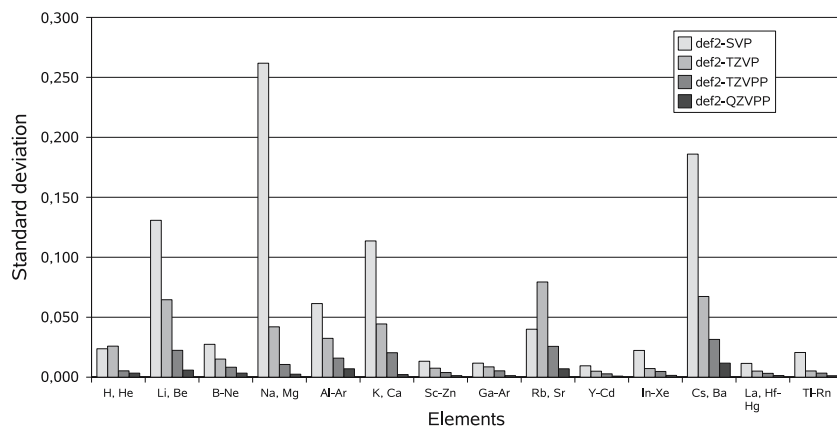
Table 6 The correlation energies of the atoms from Rb–Rn for the def2-SVP, def2-TZVPP, and def2-QZVPP basis sets compared to their basis sets limits

Element	def2-SVP		def2-TZVPP		def2-QZVPP		MP2-R12/A
	mH	%	mH	%	mH	%	mH
Rb	−53.68	22.7	−95.17	40.2	−161.27	68.1	−236.97
Sr	−79.52	26.3	−207.50	68.6	−239.34	79.1	−302.52
Y	−136.55	39.0	−214.59	61.4	−262.33	75.0	−349.69
Zr	−164.10	41.4	−258.16	65.1	−307.20	77.5	−396.36
Nb	−175.66	39.9	−303.34	69.0	−347.26	79.0	−439.74
Mo	−200.21	40.2	−354.29	71.1	−397.88	79.8	−498.37
Tc	−240.61	42.3	−414.09	72.8	−460.69	81.0	−568.42
Ru	−259.27	39.4	−473.57	72.0	−525.03	79.8	−657.78
Rh	−263.62	35.4	−527.05	70.8	−594.88	79.9	−744.15
Pd	−328.49	36.6	−638.19	71.1	−718.72	80.1	−897.72
Ag	−317.87	34.2	−653.87	70.3	−735.89	79.1	−930.44
Cd	−363.62	36.4	−709.62	71.0	−793.43	79.4	−999.42
In	−49.96	8.6	−248.44	42.8	−434.57	74.8	−580.91
Sn	−56.14	9.0	−264.63	42.2	−472.86	75.4	−626.80
Sb	−65.32	9.6	−285.45	42.1	−519.50	76.7	−677.62
Te	−77.34	10.4	−410.18	55.3	−564.75	76.1	−741.75
I	−97.35	11.9	−467.09	57.3	−635.08	77.9	−815.24
Xe	−525.25	58.6	−525.25	58.6	−702.04	78.4	−895.91
Cs	−38.64	16.4	−140.32	59.4	−166.00	70.2	−236.31
Ba	−75.47	27.6	−161.37	59.0	−186.38	68.1	−273.59
La	−134.12	41.2	−207.60	63.7	−235.21	72.2	−325.82
Hf	−146.40	38.3	−244.44	64.0	−290.34	76.0	−382.10
Ta	−168.79	40.0	−270.12	64.1	−329.75	78.2	−421.45
W	−189.22	39.4	−320.99	66.9	−373.23	77.7	−480.05
Re	−212.70	40.4	−361.00	68.6	−418.14	79.4	−526.30
Os	−254.76	42.0	−415.52	68.6	−476.81	78.7	−605.90
Ir	−279.19	41.1	−466.68	68.7	−533.22	78.5	−679.63
Pt	−286.44	37.0	−526.71	68.0	−599.82	77.5	−774.32
Au	−329.53	38.5	−582.23	68.0	−659.03	76.9	−856.83
Hg	−372.45	39.7	−653.40	69.7	−716.46	76.5	−936.99
Tl	−51.79	9.3	−344.50	61.6	−424.54	75.9	−559.34
Pb	−55.19	9.4	−351.54	60.1	−445.51	76.1	−585.11
Bi	−62.45	10.0	−369.90	59.1	−486.81	77.8	−626.07
Po	−73.63	10.8	−437.47	64.2	−521.29	76.5	−681.76
At	−87.93	11.8	−493.92	66.2	−566.73	76.0	−745.78
Rn	−540.64	66.0	−540.64	66.0	−629.56	76.9	−818.87

set has an uneven quality for these atoms. For the main group elements it covers under 12% of the correlation energies, while for the transition metals over 34% of the correlation energies are obtained. The def2-TZVPP and def2-QZVPP basis sets have for all atoms evenly distributed basis set errors. Yet they are notable larger then for lighter atoms. The remaining basis set error is by two orders of magnitude larger then the error introduced by the resolution-of-the-identity.

3.3 Computational performance

Typically, the RI approximation accelerates (in terms of computer time) MP2 calculations by a factor of 10–100. To illustrate this, the speedup of the computational time is shown in Fig. 4, where the timings for MP2 and RI-MP2 energy calculations of acene chains with a
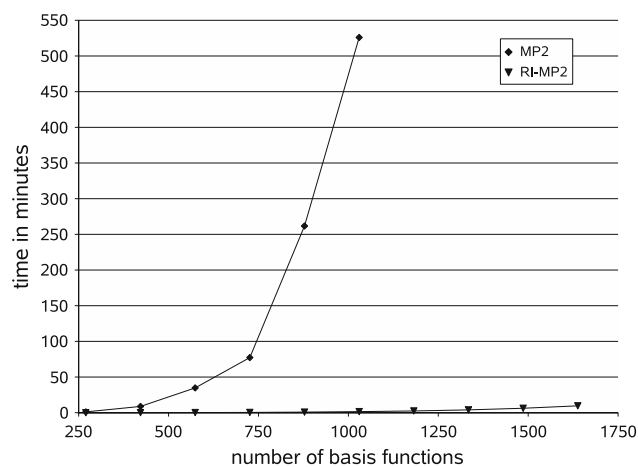
Fig. 4 The timings of MP2 and RI-MP2 ground state energies with the TZVPP basis set for acene chains with different size (from benzene to decacene)

Table 7 The difference in computational time for general and segmented contracted basis sets

Prim. GTOs and cont. GTOs are the number of primitive and contracted Gaussians, respectively. HF, N^5, and CPHF denote the time in minutes for the calculation of the HF wavefunction, the N^5 steps in MP2 and the solution of the CPHF equations. All calculations were performed with TURBOMOLE V5.8 on Intel Xeon PCs with 2.4 GHz

Molecule	Basis set	Prim. GTOs	Cont. GTOs	HF	N^5	CPHF
5-Fluorouracil						
	cc-pVTZ	426	312	25.5	0.9	11.0
	TZVPP	462	321	24.9	0.9	10.2
SiS$_2$						
	cc-pVQZ	261	117	1.8	0.02	3.9
	QZVPP	315	210	1.3	0.02	2.7
Se$_8$						
	cc-pVQZ	1216	544	1908	92	766
	QZVPP	1368	712	602	201	302
I$_2$						
	cc-pVQZ-PP	260	134	0.88	0.06	1.57
	def2-QZVPP	302	164	0.92	0.13	1.70
At$_2$						
	cc-pVQZ-PP	250	134	3.9	0.06	1.38
	def2-QZVPP	292	164	4.5	0.13	1.63

TZVPP basis set are plotted next to each other. It can be seen, that the RI approximation works very efficient here. In the limit the difference in computational time will be given by the factor $N^2/(N_x * n)$ with the number of basis functions N, of auxiliary basis functions N_x, and the number of electrons n. This factor is for reasonable basis sets always bigger than 1. The contraction scheme used for the atomic basis sets has also a large impact on the overall computational time. In particular, general contracted GTO basis sets can slow down integral direct calculations if the computer code for evaluating the electron repulsion integrals does not explicitly account for this kind of contraction, as it is the case in several program packages. On the other hand, general contracted basis sets need usually less contracted GTOs compared to the segmented contracted basis sets of the same quality, which reduces the time for correlated wavefunction methods. E.g. the time for $O(N^5)$ steps in MP2 and RI-MP2 calculations depends only on the number of contracted basis functions and can thus be expected to be shorter with the cc-pVTZ and cc-pVQZ basis sets than with the corresponding TZVPP or QZVPP basis sets. But for MP2 geometry optimizations with large basis sets the evaluation of the HF reference wavefunction and the solution of CPHF equations (both $O(N^4)$ steps) become often the actual bottleneck. In integral direct calculations (that were performed here) the time in these steps is determined by the number of primitive basis functions. In programs where general contractions can be exploited, the bottlenecks will be considerably different.

In Table 7 the timings for the largest contributions to geometry optimizations are listed. One example of a main group compound is given for each row. All calculations were performed in C_1 symmetry with the same thresholds and the same count of SCF cycles. The general observation is that RI-MP2 geometry optimizations with first row elements take nearly the same time with segmented and general contracted basis sets, for second row elements QZVPP is faster than cc-pVQZ and for third row elements QZVPP is significantly faster than cc-pVQZ. The overall wall time for a RI-MP2 gradient calculation for a Se$_8$ ring was here 13 h with QZVPP and 22 h with cc-pVQZ. For this basis sets the advantage in computational costs is striking.

4 Summary and conclusion

We have presented optimized auxiliary basis sets for RI-MP2 and RI-CC2 calculations with the segmented contracted def2-SVP, def2-TZVP, def2-TZVPP, and def2-QZVPP orbital basis sets developed recently by Weigend and Ahlrichs. [28] With these auxiliary basis sets it becomes now possible to carry out accurate RI-MP2 and RI-CC2 calculations for molecules with atoms beyond Kr using small core ECPs and basis sets up to quadruple-ζ quality. The accuracy of auxiliary basis set has been investigated by evaluating the error of the resolution-of-the-identity approximation for a test set of in total 385 molecules. Furthermore, we evaluated the atoms Rb–Rn the remaining basis set incompleteness error in the MP2 valence correlation energies with these one-electron basis sets by comparing with highly accurate explictly correlated RI-MP2-R12/A calculations. The error introduced by the resolution-of-the-identity is found to be in all cases orders of magnitudes smaller than the remaining basis set incompleteness error. E.g. with the def2-QZVPP basis the BSIE is found to be for Rb–Rn about $20 − 30\%$ for calculations on the isolated

atoms, while the largest RI error in the molecular test set is 0.06%. Thus, with these optimized auxiliary basis sets the RI approximation will only introduce errors which are completely negligible in practical applications and do not diminish the accuracy of the MP2 and CC2 model chemistries.

As a side result it is found that the def2-TZVPP and QZVPP (or def2-QZVPP) valence basis sets for Rb–Rn account for a smaller amount of correlation then for lighter atoms. The necessity to use at least the def2-QZVPP basis sets for Rb–Rn to keep the BSIE in MP2 calculations below 30% is indispensable. However, this basis set is today the only available one-electron basis which provides this accuracy consistently for all atoms from Rb to Rn. Due to the efficiency of the RI approximation, especially for large one-electron basis sets, the RI-MP2 approach—with optimized auxiliary basis sets as presented in this work—will thus be an important tool to decrease the computational costs for MP2 or CC2 calculations on compounds with heavy atoms and enhance their feasibility.

5 Availability

The optimized auxiliary basis sets presented here, as well as other auxiliary basis sets, are available by ftp at `ftp://ftp.chemie.uni-karlsruhe.de/pub/cbasen/`. In the header of each auxiliary basis set the correlation energy and the RI error of the atomic energy is documented. The Cartesian coordinates of the molecules in the test set are also available by ftp at `ftp://ftp.chemie.uni-karlsruhe.de/pub/structures/`.

References

1. Ahlrichs R, Bär M, Häser M, Horn H, Kömel C (1989) Chem Phys Lett 162:164–169
2. Andrae D, Häussermann U, Dolg M, Stoll H, Preuss H (1990) Theor Chim Acta 77:123–141
3. Bachorz R, Klopper W, Gutowski M (2007) (to be published)
4. Dolg M, Stoll H, Preuss H (1989) Theor Chim Acta 75:173–194
5. Dunlap BI, Connolly JWD, Sabin JR (1979) J Chem Phys 71:3396–3402
6. Dunning TH (1989) J Chem Phys 90:1007–1023
7. Dunning TH, Peterson KA, Wilson AK (2001) J Chem Phys 114:9244–9253
8. Eichkorn K, Weigend F, Treutler O, Ahlrichs R (1997) Theor Chem Acc 97:119–124
9. Haase F, Ahlrichs R (1993) J Comp Chem 14:907–912
10. Häser M, Ahlrichs R (1989) J Comput Chem 10:104–111
11. Hättig C (2005) Phys Chem Chem Phys 7:59–66
12. Hättig C, Weigend F (2000) J Chem Phys 113:5154–5162
13. Kaupp M, Schleyer PV, Stoll H, Preuss H (1991) J Chem Phys 94:1360–1366
14. Kendall RA, Dunning TH, Harrison RJ (1992) J Chem Phys 96:6796–6806
15. Klopper W (1991) Chem Phys Lett 186:583–585
16. Klopper W, Kutzelnigg W (1987) Chem Phys Lett 134:17–22
17. Klopper W, Manby FR, Ten-no S, Valeev EF (2006) Int Rev Phys Chem 25:427–468
18. Leininger T, Nicklass A, Küchle W, Stoll H, Dolg M, Bergner A (1996) Chem Phys Lett 255:274–280
19. Manby FR (2003) J Chem Phys 119:4607–4613
20. Metz B, Stoll H, Dolg M (2000) J Chem Phys 113:2563–2569
21. Peterson KA, Figgen D, Goll E, Stoll H, Dolg M (2003) J Chem Phys 119:11, 113–11, 123
22. Schäfer A, Huber C, Ahlrichs R (1994) J Chem Phys 100:5829–5835
23. Vahtras O, Almlöf JE, Feyereisen MW (1993) Chem Phys Lett 213:514–518
24. Villani C, Klopper W (2005) J Phys B 38:2555–2567
25. Villani C, A Glöß, Hättig C, Klopper W (2007) (to be published)
26. Weigend F (2002) Phys Chem Chem Phys 4:4285–4291
27. Weigend F (2006) Phys Chem Chem Phys 8:1057–1065
28. Weigend F, Ahlrichs R (2005) Phys Chem Chem Phys 7:3297–3305
29. Weigend F, Häser M (1997) Theor Chem Acc 97:331–340
30. Weigend F, Häser M, Patzelt H, Ahlrichs R (1998) Chem Phys Lett 294:143–152
31. Weigend F, Köhn A, Hättig C (2002) J Chem Phys 116:3175–3183
32. Whitten JL (1973) J Chem Phys 58:4496–4501
33. Wilson AK, Woon DE, Peterson KA, Dunning TH (1999) J Chem Phys 110:7667–7676
34. Woon DE, Dunning TH (1993) J Chem Phys 98:1358–1371