

Auxiliary basis sets for density fitting–MP2 calculations: Nonrelativistic triple- ζ all-electron correlation consistent basis sets for the 3d elements Sc–Zn

J. Grant Hill and James A. Platts^{a)}

School of Chemistry, Cardiff University, Park Place, Cardiff CF10 3AT, United Kingdom

(Received 22 October 2007; accepted 27 November 2007; published online 24 January 2008)

Auxiliary basis sets for density fitting second-order Møller-Plesset perturbation theory (DF-MP2) have been optimized for use with the triple- ζ nonrelativistic all-electron correlation consistent orbital basis sets, cc-pVTZ-NR and aug-cc-pVTZ-NR, for the 3d elements Sc–Zn. The relative error in using these auxiliary basis sets is found to be around four orders of magnitude smaller than that from utilizing triple- ζ orbital basis sets rather than corresponding quadruple- ζ basis sets, in calculation of the correlation energy for a test set of 54 small to medium sized transition metal complexes. © 2008 American Institute of Physics. [DOI: 10.1063/1.2826348]

I. INTRODUCTION

The density fitting (DF) approximation of electron repulsion integrals has been shown to reduce the computer time necessary for second-order Møller-Plesset perturbation theory (MP2) calculations by around an order of magnitude.^{1–3} This is achieved by expanding each orbital product density in an optimized auxiliary density fitting basis set. Although density fitting, which is also referred to as the resolution-of-the-identity (RI) approximation by some authors, was first used in density functional theory,^{4,5} it has since been implemented as an approximation in Hartree-Fock theory^{6,7} along with several correlated *ab initio* methods such as explicitly correlated MP2 theories,^{8,9} approximate coupled-cluster singles-and-doubles model CC2 (in both its canonical¹⁰ and local¹¹ forms), and coupled-cluster methods.^{12,13}

The correlation consistent (cc) basis sets were originally introduced by Dunning for first row atoms,¹⁴ and have since been extended to cover *p*-block elements^{15–19} along with options available for core-valence electron correlation^{20,21} and the description of anions and noncovalent interactions.^{22,23} In their simplest, valence-only correlated form these basis sets are usually denoted as cc-pVnZ, where $n=D, T, Q, 5$, and so on. As these basis sets systematically converge towards the complete basis set (CBS) limit, they have seen extensive use in high accuracy correlated *ab initio* investigations of systems involving the main group elements. Given this popularity, it is perhaps unsurprising that auxiliary basis sets for DF-MP2 (and other methods such as DF-CC2 that can utilize the same auxiliary basis sets) have been optimized for the correlation consistent orbital basis sets to significantly decrease the cost of such calculations.^{24,25}

Only recently have all-electron correlation consistent basis sets been developed for the 3d transition metal elements Sc–Zn.²⁶ These sets range from triple- ζ to quintuple- ζ in quality, with augmented and core-valence variants, and have

been determined both with nonrelativistic (NR) and relativistic Douglas-Kroll-Hess calculations. In the present work, optimized cc-pVTZ-NR and aug-cc-pVTZ-NR auxiliary fitting basis sets for the 3d transition metal elements Sc–Zn are presented, and then verified by assessing the relative error introduced by the density fitting approximation over a test set of molecules.

II. BASIS SET CONSTRUCTION AND OPTIMIZATION

Optimization of auxiliary basis sets was performed using analytic gradients for the optimization of auxiliary basis sets implemented in the RICC2 module^{10,27} of the TURBOMOLE package.^{28,29} The cc-pVTZ-NR and aug-cc-pVTZ-NR orbital basis sets were obtained from the William R. Wiley Environmental Molecular Sciences Laboratory Basis Set Exchange.³⁰ The auxiliary basis sets were constructed with the guidelines from Ref. 24 in mind, namely, that the number of auxiliary functions should be less than four times the number of orbital basis functions, that the value of the functional

$$\delta_{\text{DF}} = \frac{1}{4} \sum_{abij} \frac{(\langle ab||ij \rangle_{\text{DF}} - \langle ab||ij \rangle)^2}{\varepsilon_a - \varepsilon_i + \varepsilon_b - \varepsilon_j}$$

divided by the canonical MP2 correlation energy is less than 10^{-6} , and that there are the same number of functions for every atom within a specific basis set. However, as transition metals are found in varying oxidation states in different chemical compounds, neutral atoms do not provide a realistic test of auxiliary basis set performance for common transition metal complexes. Hence, the criteria used to assess the auxiliary basis sets in the current investigation are that the mean error, the standard deviation, and the mean absolute error due to the DF approximation over a series of small and medium sized transition metal complexes should be insignificant when compared to the error due to orbital basis set incompleteness (highlighted in this case by the change in energy due to increasing the basis sets to quadruple- ζ quality). Once more, following the guidelines of Ref. 24, the density fitting

^{a)}Electronic mail: platts@cf.ac.uk.

TABLE I. Ratio of the number of GTOs in the auxiliary basis sets compared to the generally contracted orbital basis sets.

	Orbital basis	Auxiliary basis	Ratio of functions
cc-pVTZ-NR	[7s6p4d2f1g]	(13s11p9d7f5g3h2i)	2.5
aug-cc-pVTZ-NR	[8s7p5d3f2g]	(14s12p10d8f6g4h3i)	2.3

error should be at least two orders of magnitude smaller than the error in employing triple-rather than quadruple- ζ orbital basis sets.

The fact that transition metals are commonly found in more than one oxidation state across several chemical compounds means that optimizing an auxiliary basis set purely for the neutral atom is likely to lead to poor results when the auxiliary basis set is applied to test compounds. Following the procedure outlined for the optimization of TZVP and QZVPP auxiliary basis sets for the 3d transition metal elements,²⁵ the exponents of the auxiliary basis are divided into different categories. In the cases of Sc and Ni–Zn there are two categories: Tight and diffuse, while the elements Ti–Co have an additional intermediate category. The range of exponents corresponding to each of these categories was determined by inspecting how the orbital basis functions contribute to the atomic orbitals (AOs). Exponents contributing to the outermost valence AOs and those of greater angular momentum are considered diffuse, while those contributing to the core AOs are considered tight. These boundary guides were then permitted to change by a maximum of two functions when optimizing the auxiliary basis, in order to ensure that satisfactory convergence is achieved. Initially, the whole auxiliary basis set is optimized for the cation of the element at the highest oxidation state generally found in chemical compounds. The tight exponents in the resulting set are then frozen while the diffuse exponents are optimized for the neutral atom. For elements with an intermediate exponent category, those exponents are optimized for the cation typically corresponding to the most commonly found oxidation state of the element, again, with the optimized tight exponents frozen. A complete list of the exponent ranges for each atom is provided as supporting information. Again, following the procedure of Ref. 25, occupied orbitals from the 1s2s2p core shells were excluded from the calculation of the δ_{DF} functional and gradient.

Unlike previous auxiliary augmented correlation consistent basis sets, it did not prove possible to simply take the cc-pVTZ-NR auxiliary set and optimize an additional diffuse auxiliary function for each angular momentum quantum number as the resulting set was insufficiently accurate. Instead, the entire auxiliary basis set was reoptimized, following the above procedure, with an additional diffuse function added to each angular momentum function used in the non-augmented sets.

III. RESULTS AND DISCUSSION

The total number of Gaussian type orbitals (GTOs) included in the auxiliary basis sets for each angular momentum quantum number are shown in Table I, along with the num-

TABLE II. Relative percentage error in the MP2 correlation energy due to orbital basis set size, assessed by comparing with quadruple- ζ correlation consistent basis set for a test set of 54 molecules containing 3d transition metal elements. See text for further details.

	cc-pVTZ-NR	aug-cc-pVTZ-NR
$\bar{\Delta}^{\text{bas}}$	9.96	8.81
$\Delta_{\text{std}}^{\text{bas}}$	1.91	1.42
$\Delta_{\text{max}}^{\text{bas}}$	16.21	11.41

ber of GTOs in the generally contracted orbital basis. It can be seen that, just as with other cc auxiliary basis sets,^{24,25} it was necessary to include functions with an angular momentum quantum number = $(l_{\text{occ}} + l_{\text{bas}})$, where l_{occ} and l_{bas} are the highest occupied angular momentum for the atom and largest angular momentum included in the orbital basis set, respectively, to reach the desired accuracy. Although one of the desirable features of the cc basis sets is the ability to systematically approach the CBS limit, the majority of current quantum chemistry codes do not support auxiliary basis functions with an angular momentum quantum number of k or higher. A consequence of the auxiliary basis sets requiring such high angular momentum functions for a proper description of the orbital basis is that the current work has been restricted to triple- ζ basis sets, despite the latest versions of TURBOMOLE supporting higher angular momentum auxiliary functions.

In order to verify that the auxiliary basis sets accurately describe transition metal complexes, DF-MP2 and canonical MP2 calculations were carried out on a set of 54 small to medium sized 3d transition metal complexes to evaluate the error introduced by density fitting. These 54 molecules are a subset of the K–Kr set from Ref. 25, with the K–Ca and Ga–Kr molecules removed, and are summarized in the supporting information.

Just as the use of the density fitting approximation introduces an error, there is also an error due to orbital basis set incompleteness. The relative percentage error in the canonical MP2 correlation energy for the cc-pVTZ and aug-cc-pVTZ orbital basis sets when compared to the cc-pVQZ and aug-cc-pVQZ sets, respectively, has been calculated for the test set of molecules. While quadruple- ζ basis sets are not at the CBS limit, the relative errors that are highlighted by this approach are a useful comparison for the errors introduced by density fitting. Table II presents the relative orbital basis set error between triple- and quadruple- ζ basis sets [$\Delta_i^{\text{bas}} = (E^{\text{VQZ-corr}} - E^{\text{VTZ-corr}}) / E^{\text{VQZ-corr}} \times 100\%$] as the mean of this error over all 54 molecules ($\bar{\Delta}^{\text{bas}}$), the standard deviation of the error ($\Delta_{\text{std}}^{\text{bas}}$), and the largest error within the test set ($\Delta_{\text{max}}^{\text{bas}}$). All non-3d elements were described with the appropriate cc orbital basis sets, with all MP2 calculations carried out using the frozen core approximation.

The relative errors in the correlation energy from Table II seem to be rather large, especially when this error will only increase as the basis set size increased towards the CBS limit. However, the error is in line with that observed for the cc-pVTZ basis relative to a CBS extrapolation for molecules containing second row atoms,²⁵ and due to the current high

TABLE III. Relative percentage errors introduced via the density fitting approximation for a test set of 54 molecules containing 3*d* transition metal elements. See text for further details.

	cc-pVTZ-NR	aug-cc-pVTZ-NR
$\bar{\Delta}^{\text{DF}}$	0.0004	0.0006
$\Delta_{\text{std}}^{\text{DF}}$	0.0054	0.0020
$ \bar{\Delta}^{\text{DF}} $	0.0026	0.0017
$\Delta_{\text{max}}^{\text{DF}}$	0.0317	0.0057

computational cost, triple- ζ quality basis sets are likely to be a pragmatic choice for investigations involving medium sized transition metal complexes. Additionally, the small standard deviation suggests this may be a reasonably consistent error and, hence, less important in the case of relative energies. While this may be worthy of further investigation, it is beyond the scope of the current paper.

The error introduced by the density fitting approximation in the evaluation of the test set of molecules is assessed as the relative error in the correlation energy: $\Delta_i^{\text{DF}} = (E^{\text{corr}} - E^{\text{DF-corr}}) / E^{\text{corr}} \times 100\%$, where E^{corr} and $E^{\text{DF-corr}}$ are the canonical and density fitting MP2 correlation energies, respectively. Table III summarizes this error as the mean of the error over all 54 molecules ($\bar{\Delta}^{\text{DF}}$), the standard deviation of the error ($\Delta_{\text{std}}^{\text{DF}}$) and the largest error within the test set ($\Delta_{\text{max}}^{\text{DF}}$). Although larger basis sets will always lead to a lower energy and, therefore, an error of the same sign in the case of canonical MP2, this is not the case with density fitting and, hence, the mean absolute error ($|\bar{\Delta}^{\text{DF}}|$) is also reported. All non-3*d* elements were described with the appropriate cc auxiliary and orbital basis sets.

Table III indicates that both of the auxiliary basis sets currently presented are of a similar quality. Comparing Tables II and III shows that the mean error introduced using these auxiliary basis sets is around four orders of magnitude smaller than that from using triple- rather than quadruple- ζ orbitals basis sets. In the case of the maximum errors, the relative DF error rises to approximately three orders of mag-

nitude smaller than that of the orbital basis set. This confirms that using these auxiliary sets in the density fitting approximation produces an insignificant error in the MP2 correlation energy. Although these tests also include errors introduced from other auxiliary basis sets due to the inclusion of non-transition metal elements, they are more indicative of the size of errors likely to be introduced in real world applications than examining the effect on 3*d* atoms alone. Figure 1 plots the normalized Gaussian distributions of the relative errors due to density fitting over the test set of 54 molecules, along with distributions for the TZVPP and QZVPP auxiliary basis sets determined from data in Ref. 25. It should be noted that the statistics presented in Ref. 25 cover molecules that include atoms from K–Kr and not just the 3*d* elements Sc–Zn, which may have some bearing upon this comparison, but it does show that the fitting of the auxiliary basis sets presented in the current paper is of a similar quality to that of the QZVPP auxiliary and orbital basis set.

Although the computational benefits of using the DF approximation are well known, a simple illustration of its use in small to medium sized transition metal complexes is that a canonical single point MP2 calculation at the aug-cc-pVTZ level, excluding the initial SCF, on ferrocene took 4540 s while the same calculation using the DF approximation was completed in 140 s. In both cases, the calculations were carried out in serial on an AMD 2.4 GHz Opteron based machine.

IV. CONCLUSIONS

Triple- ζ nonrelativistic all-electron correlation consistent auxiliary basis sets in both their standard and augmented forms (cc-pVTZ-NR and aug-cc-pVTZ-NR, respectively) have been optimized for the 3*d* transition metals Sc–Zn. The accuracy of using these auxiliary basis sets in the density fitting approximation has been investigated in terms of the relative error within the MP2 correlation energy over a test set of 54 small to medium sized transition metal complexes. As the error introduced is such a negligible percentage of the

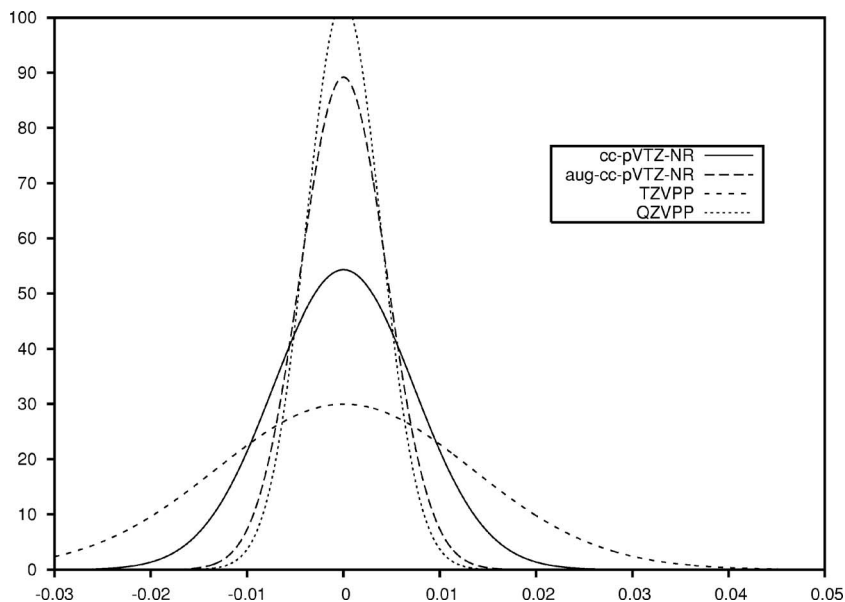


FIG. 1. Normalized Gaussian distributions of the relative percentage errors introduced via the density fitting approximation. The TZVPP and QZVPP distributions are determined from $|\bar{\Delta}^{\text{DF}}|$ and $\Delta_{\text{std}}^{\text{DF}}$ values taken from Ref. 25 and cover a test set of molecules containing the atoms K–Kr. The cc-pVTZ-NR and aug-cc-pVTZ-NR distributions are based on a 3*d* subset of these molecules.

total correlation energy, and is approximately four orders of magnitude smaller than the relative error in using a triple- rather than quadruple- ζ orbital basis, the use of these auxiliary basis sets in DF-MP2 calculations will provide an accurate approximation of canonical MP2 correlation energies.

All of the basis sets presented in the current work can be found in the supporting information, and will be made available via the Environmental Molecular Sciences Laboratory Basis Set Exchange³⁰ website (<http://gnode2.pnl.gov/bse/portal>).

ACKNOWLEDGMENTS

The EPSRC National Service for Computational Chemistry Software³¹ is acknowledged for the CPU time and access to the TURBOMOLE package it provided. The authors would also like to thank Dr. Kirk A. Peterson for invaluable discussions that allowed this work to be carried out.

¹F. Weigend, M. Häser, H. Patzelt, and R. Ahlrichs, *Chem. Phys. Lett.* **294**, 143 (1998).

²O. Vahtras, J. E. Almlöf, and M. W. Feyereisen, *Chem. Phys. Lett.* **213**, 514 (1993).

³H.-J. Werner, F. R. Manby, and P. J. Knowles, *J. Chem. Phys.* **118**, 8149 (2003).

⁴E. J. Baerends, D. E. Ellis, and P. Ros, *Chem. Phys.* **2**, 41 (1973).

⁵B. I. Dunlap, J. W. D. Connolly, and J. R. Sabin, *J. Chem. Phys.* **71**, 3396 (1979).

⁶F. Weigend, *Phys. Chem. Chem. Phys.* **4**, 4285 (2002).

⁷R. Polly, H.-J. Werner, F. R. Manby, and P. J. Knowles, *Mol. Phys.* **102**,

2311 (2004).

⁸A. J. May and F. R. Manby, *J. Chem. Phys.* **121**, 4479 (2004).

⁹H.-J. Werner and F. R. Manby, *J. Chem. Phys.* **124**, 054114 (2006).

¹⁰C. Hättig and F. Weigend, *J. Chem. Phys.* **113**, 5154 (2000).

¹¹D. Kats, T. Korona, and M. Schütz, *J. Chem. Phys.* **127**, 064107 (2007).

¹²A. P. Rendell and T. J. Lee, *J. Chem. Phys.* **101**, 400 (1994).

¹³M. Schütz and F. R. Manby, *Phys. Chem. Chem. Phys.* **5**, 3349 (2003).

¹⁴T. H. Dunning, *J. Chem. Phys.* **90**, 1007 (1989).

¹⁵D. E. Woon and T. H. Dunning, *J. Chem. Phys.* **98**, 1358 (1993).

¹⁶A. K. Wilson, K. A. Peterson, D. E. Woon, and T. H. Dunning, *J. Chem. Phys.* **110**, 7667 (1999).

¹⁷T. H. Dunning, K. A. Peterson, and A. K. Wilson, *J. Chem. Phys.* **114**, 9244 (2001).

¹⁸K. A. Peterson, *J. Chem. Phys.* **119**, 11099 (2003).

¹⁹K. A. Peterson, D. Figgen, E. Goll, H. Stoll, and M. Dolg, *J. Chem. Phys.* **119**, 11113 (2003).

²⁰D. E. Woon and T. H. Dunning, *J. Chem. Phys.* **103**, 4572 (1995).

²¹K. A. Peterson and T. H. Dunning, *J. Chem. Phys.* **117**, 10548 (2002).

²²R. A. Kendall, T. H. Dunning, and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).

²³D. E. Woon and T. H. Dunning, *J. Chem. Phys.* **100**, 2975 (1994).

²⁴F. Weigend, A. Köhn, and C. Hättig, *J. Chem. Phys.* **116**, 3175 (2002).

²⁵C. Hättig, *Phys. Chem. Chem. Phys.* **7**, 59 (2005).

²⁶N. B. Balabanov and K. A. Peterson, *J. Chem. Phys.* **123**, 064107 (2005).

²⁷C. Hättig, *J. Chem. Phys.* **118**, 7751 (2003).

²⁸R. Ahlrichs, M. Bär, M. Häser, H. Horn, and C. Kölmel, *Chem. Phys. Lett.* **162**, 165 (1989).

²⁹TURBOMOLE, version 5.8, see http://www.cosmologic.de/QuantumChemistry/main_turbomole.html

³⁰K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li, and T. L. Windus, *J. Chem. Inf. Model.* **47**, 1045 (2007).

³¹EPSRC National Service for Computational Chemistry Software, see <http://www.nscs.ac.uk>