Behavior of the Sapporo- n ZP-2012 Basis Set Family

Rebecca Weber, Benjamin Hovda, George Schoendorff, Angela K. Wilson*

*Department of Chemistry and Center for Advanced Scientific Computing and Modeling (CASCAM),
University of North Texas, Denton, Texas 76203-5017*

ABSTRACT

The behavior of Sapporo- n ZP-2012 ($n=D,T,Q$) basis sets for $3d$ atoms is investigated. It is shown that total atomic energies obtained using Sapporo sets of increasing quality converge to a limit. Several schemes are used to determine these limits and comparisons are made to numerical Hartree-Fock energies and to limits obtained with correlation consistent sets (cc-pV n Z, $n=T,Q,5$). While differences in energy relative to numerical HF are larger for Sapporo- n ZP-2012 sets (16-40 mE_h) than for correlation consistent basis sets (~0.005 mE_h), the systematic nature of the Sapporo- n ZP-2012 sets may provide a reasonable alternative when the correlation consistent basis sets are not available.

Keywords:

Basis set, Sapporo, extrapolation, complete basis set limit, convergence behavior

* akwilson@unt.edu, phone 940-565-4372, fax 940-565-4318

INTRODUCTION

The solution of the Schrödinger equation in electronic structure calculations requires a number of approximations. Commonly, a finite basis set is used, resulting in an error in the calculated energy in comparison to that of an infinite (*theoretical*), or complete, basis set. This error can be referred to as the basis set incompleteness error. Hierarchies of basis sets have emerged, however, which result in systematic convergence in properties such as bond distances or energies with respect to increasing basis set size to a limit, typically referred to as the complete basis set (CBS) limit. Extrapolation of these properties can minimize the basis set incompleteness error. In principle, when this error is reduced to zero, then the complete basis set (CBS) limit is obtained. This limit is the point at which no further improvement in the basis set will improve upon or change the property prediction, and the remaining error in the approximations can be attributed to the choice of method utilized.¹ These characteristics make the CBS limit a useful reference for approaches such as *ab initio* composite methods, as well as provide insight about the performance of computational methods.

In principle, the CBS limit can be obtained from calculations performed using only a double- ζ and triple- ζ basis set (though larger basis may be needed) and this can eliminate the need for larger, more computationally expensive (*i.e.*, computer time, memory, and disk space requirements) calculations. Systematic behavior from hierarchical basis sets has been well established²⁻⁶ for the correlation consistent basis sets,⁷⁻¹¹ with recent advances designed for density functional theory (DFT) calculations.^{12,13}

While there are many families of basis sets for which smooth convergence to the CBS limit does not occur for properties, there are a number of sets for which this unique characteristic has not yet been well studied. Particularly for areas of the periodic table for which cc-pVnZ basis sets, for example, have not yet fully emerged, considering the possibility of convergent behavior for properties with other basis sets is important. One hierarchical basis set family is the Sapporo-*n*ZP-2012 family.^{14,15} The Sapporo sets are compact segmented basis sets, designed to recover the correlation energy of a system efficiently. These basis sets were based on the original NOSeC-*n*ZP sets ($n = D, T, Q$),¹⁶ which were designed to represent the space spanned by atomic natural orbitals (ANOs) as determined by configuration interaction (CI) calculations. The Sapporo sets improve upon the NOSeC-*n*ZP basis sets by including important core-valence correlation through correlation of the subvalence electrons [*i.e.* the n and $(n-1)$ shell of electrons, n being the primary quantum number].

The *n*ZP sets are based upon the primitive sets of Huzinaga, *et al.*¹⁷ which were extended by adding higher angular momentum functions and used to determine a set of accurate natural orbitals (NOs). The *n*ZP sets were created by truncating the NOs derived from non-relativistic CI

calculations.* For the transition metal atoms, the ns and $(n-1)s$, p , and d electrons were included within the correlation space. The sizes of the n ZP sets were determined using cc-pVnZ sets as a guide, e.g. the size of the double- ζ set for scandium (Sc) is $[6s4p3d1f]$, beginning from a minimal basis of $[4s3p1d]$. To determine the core correlating functions, the HF sets of Tatewaki and Koga were used.¹⁸⁻²¹ The n and $(n-1)$ shells of orbitals were decontracted and the valence correlating functions from the NOSeC- n ZP sets were added. If the numbers of the decontracted s and p primitives included were found to be lacking, then additional primitives were added. The functions were optimized to reproduce the ANO's, and were optimized for computational efficiency. Because of small energetic contributions, the authors removed one s -type, p -type, and d -type function each from the quadruple- ζ sets for fourth and fifth row atoms, thereby reducing the size of the basis set. In CI calculations, the Sapporo sets have been shown to recover more than 90% of the atomic correlation energy compared to using ANOs.¹⁴ The Sapporo sets have fewer contracted basis functions than the corresponding correlation consistent sets. For example, for titanium, the Sapporo-TZP-2012 basis set results in 69 contracted basis functions, while the corresponding correlation consistent triple- ζ set (cc-pwCVTZ) results in 102 contracted functions.

While the Sapporo sets were designed to provide a hierarchy of efficient basis sets, limited, if any, attention has been given to the possible convergence of properties with respect to increasing basis set size for these sets. In part, this may be due to construction, particularly as the cc-pVnZ sets are available for so much of the periodic table, and are explicitly designed to address a systematic accounting for the correlation energy. In considering transition metal and heavier p -block systems, the correlation consistent family of sets includes all-electron and effective core potential (ECP)-based sets for heavy p -block²² and ECP-based sets for transition metals and all-electron sets for some of the transition metals.^{10,23-25}

For the heaviest of elements, there are fewer basis set options than for lighter elements. A basis set option that is available for the heaviest of elements is the Dyall correlation consistent family of basis sets designed for four-component relativistic calculations.²⁶ While these sets can be uncontracted for use in calculations, another approach is to utilize an ECP-based approach for basis set development, as has been recently utilized by Peterson for uranium and thorium.²⁷ Despite these and likely continuing developments, the Sapporo- n ZP-2012 sets already exist across the periodic table and provide a series of sets worth considering for possible convergence of properties towards the CBS limit. Having routes to achieve CBS limits for the heaviest of elements will enable a greater understanding of methodological needs for molecules including such elements.

Though our interest is in the heaviest of elements, the behavior of the sets for lighter elements can be considered, as there are other useful sources for comparison, including the cc-

* For the relativistic variant of the basis sets, the 3rd order Douglas-Kroll Hamiltonian was used for relativistic CI calculations.

pVnZ sets. As well, numerical Hartree-Fock (HF) calculations can be utilized as an excellent reference, particularly as numerical HF is a basis set independent method. Prior work has demonstrated that the numerical HF energy can be considered the HF limit, given a sufficiently large grid size.²⁸⁻³¹ In fact, numerical HF energies have been utilized as the HF limit in studies of the convergence behavior of different basis set families.³⁰⁻³⁴ The efficacy of various extrapolation schemes for estimating the HF limit may be tested by comparing the extrapolated energies to the numerical HF energy.

In this study, the convergence behavior of the total atomic energies calculated at the nonrelativistic HF level of theory using the Sapporo-*n*ZP-2012 basis sets is investigated by comparing computed energies to nonrelativistic numerical HF energies. Convergence of the correlation energy has also been considered.

Extrapolation Schemes

For cc-pVnZ, many schemes have been proposed for extrapolation of energies to the CBS limit, including utilizing an exponential form, an inverse power form, or a combination of the two to describe the asymptotic behavior of the total energy.^{5,6,35-39} Previous work has demonstrated that an exponential function better describes the convergence behavior for HF energies obtained with the cc-pVnZ sets than the power form, which is generally utilized for correlation energies.^{31,33,34,40} It should be noted that the extrapolation schemes utilized in this study have originally been identified for use with the correlation consistent basis sets. Therefore, it is important to consider different functional forms for the extrapolations in this work, as what is effective in extrapolating correlation consistent energies may not work well with the Sapporo sets. As such, three exponential schemes and three inverse power schemes have been considered in this study as possible routes to extrapolation.

The exponential extrapolation scheme originally proposed by Feller (Eqn. 1) is an approach that has been shown to perform well with the correlation consistent basis sets.^{33,34,36}

$$E_x = E_\infty + A \exp(-Bn) \quad (1)$$

In Eqn. 1, E_x is the energy at each level of basis set, where $n = D$ (2), T (3), and Q (4); E_∞ is the limit of the HF energy; A and B are fitting parameters; and n is the quality of basis set. Eqn. 1 is generally used as a three-point extrapolation. A value of 1.63 for B has been shown to provide the lowest mean absolute deviation (MAD) for main group diatomics in comparison to numerical HF when using correlation consistent basis sets, enabling a reduction in parameters from a three point to a two point extrapolation scheme.³³ A subsequent study demonstrated that, while optimizing the parameter B for transition metal diatomics does reduce the MAD from numerical HF energies, overall it is a negligible error reduction. Therefore, Eqn. 1 with a B value of 1.63 is suggested for use in the extrapolation of HF energies for both main group and transition metals

and will be tested with the Sapporo-2012 basis sets as well.³⁴ The second extrapolation formula is an exponential form proposed by Jensen³¹ and modified by Karton and Martin³⁹ (Eqn. 2) and is utilized within the W4 composite method.⁴¹

$$E_x = E_\infty + A(l_{\max} + 1)\exp(-9\sqrt{l_{\max}}) \quad (2)$$

In Eqn. 2, E_x is the energy at each level of basis set; E_∞ is the energy at the CBS limit; $l_{\max}$ is the maximum angular momentum included within the basis set; and A is a fitting parameter. Another extrapolation scheme is the widely used three-point mixed Gaussian/exponential scheme of Peterson and co-workers (Eqn. 3).⁵

$$E_x = E_\infty + A\exp(-(n-1)) + B\exp(-(n-1)^2) \quad (3)$$

In Eqn. 3, E_x and E_∞ are the energies at the basis set level and limit, respectively; A and B are fitting parameters; and n is the ζ -level of the basis set. Previous studies on the cc-pVnZ family have demonstrated that of the exponential extrapolation schemes listed here, Eqn. 1 with a B value of 1.63 results in the lowest mean absolute deviation (MAD) for energies in comparison to numerical HF.^{33,34,42} A variant of the Peterson extrapolation is proposed in which the $(n-1)$ term is replaced by the $(C\sqrt{l_{\max}})$ term from the Karton and Martin extrapolation (Eqn. 4), in which C was optimized to 1.544.

$$E_x = E_\infty + A\exp(-1.544\sqrt{l_{\max}}) + B\exp(-(1.544\sqrt{l_{\max}})^2) \quad (4)$$

Additionally, a combination exponential scheme is also proposed that combines aspects of the Feller and Karton and Martin schemes (Eqn. 5).

$$E_x = E_\infty + A\exp(-1.63n) + B\exp(-4\sqrt{n}) \quad (5)$$

Based on work by Schwartz,³⁵ Carroll *et al.*,⁴³ and Hill,⁴⁴ inverse cubic (Eqn. 6) and inverse quartic (Eqn. 7) schemes have shown to be successful for extrapolations to the CBS limit for the cc-pVnZ basis sets.

$$E_x = E_\infty + A(l_{\max})^{-3} \quad (6)$$

$$E_x = E_\infty + A(l_{\max} + 1/2)^{-4} \quad (7)$$

Another scheme dependent on a series of inverse power functions was determined by Wilson and Dunning to be very effective in extrapolating MP2 energies (Eqn. 8).⁶

$$E_x = E_\infty + \frac{A}{(l_{\max} + 1)^4} + \frac{B}{(l_{\max} + 1)^5} \quad (8)$$

In Eqns. 6-8, E_x and E_∞ are the energies at the basis set level and limit, respectively; A and B are fitting parameters; and l_{max} is the maximum angular momentum in the basis set.

While extrapolation to the CBS limit is possible in principle, the fact that there are numerous extrapolation schemes that produce varying results suggests that even the extrapolation methods are approximate.^{34,38,45} All of the extrapolation schemes minimize the basis set incompleteness error such that they recover the “CBS” limit. In a previous study, several different extrapolation schemes, including some that were considered in this study (i.e. F, P, S3, S4), were used to estimate the HF limit of 3d transition metal-containing systems utilizing the correlation consistent basis sets.³⁴ Using the cc-pVTZ and cc-pVQZ energies and the F scheme with varying B values, a range of 300-800 μE_h from the numerical Hartree-Fock limit was obtained. Inclusion of diffuse functions within the basis sets reduced this error to 200-400 μE_h . Utilizing the cc-pVQZ and cc-pV5Z sets within the extrapolation, the range of the results was much reduced to about 20 μE_h , yet the error from the numerical HF was still on the order of almost 200 μE_h . This demonstrates that the incompleteness error present from the basis sets is not completely eliminated, yet the deviation from the numerical HF limit is so small that the CBS limit is effectively obtained.

Previous studies with cc-pVnZ sets have advocated separately extrapolating the HF and correlation energies.^{25,41,46} In this work, extrapolations of the total energies and the separate energies obtained with the Sapporo-2012 sets are performed.

While certainly there are many other equation forms that could be considered for the calculation of CBS limits, the equations that have been examined in this work may enable the use of convergent behavior of properties determined with the Sapporo- n ZP-2012 basis sets to a limit and how this limit compares with the numerical HF limit for HF calculations to be ascertained.

COMPUTATIONAL METHODS

All calculations were performed using the NWChem software package.⁴⁷ The $4s^1 3d^n$ and $4s^2 3d^{n-1}$ states of the 3d transition metal atoms were calculated using restricted open-shell Hartree-Fock (ROHF) with the Sapporo- n ZP-2012 ($n = D, T, Q$) basis sets. The ROHF energies were compared to ROHF energies obtained using the cc-pVnZ ($n = T, Q, 5$) basis sets (previously published by Balabanov and Peterson)¹⁰ and nonrelativistic numerical Hartree-Fock energies that are accurate to at least 1 μE_h .^{28,29} Although a double- ζ energy can provide a useful anchor in extrapolations,^{48,49} the cc-pVDZ basis sets for Sc – Zn were not developed in a correlation consistent fashion[†] and so were not considered in this study. If required, orbitals were

[†] Beginning with the cc-pVTZ set, the last atomic natural orbital (ANO) contraction was removed from the s , p , and d shells, as well as the g function. The two f functions in the cc-pVTZ set were contracted to one function. Otherwise, the sets are the same. Ref. 23

rotated to the desired (i.e. lowest energy) occupation and the Maximum Overlap Method (MOM)⁵⁰ was utilized to maintain occupation of the selected electronic state. The electronic states of the atoms are the same states utilized by Balabanov and Peterson during the creation of the 3*d* correlation consistent sets. The final energies were obtained with no symmetry restrictions to ensure a positive-definite solution on the orbital potential energy surface. Relativistic effects were not included in the atomic calculations as the numerical HF energies did not include relativistic effects. The atomic energies were calculated for each level of basis set and then extrapolated using the schemes described below. The following abbreviations are used herein for the extrapolation schemes: F indicates the Feller extrapolation (Eqn. 1); KM indicates the Karton and Martin extrapolation (Eqn. 2); P indicates the Peterson extrapolation (Eqn. 3); P+KM is the combination of the Peterson and KM schemes (Eqn. 4) and F+KM is the combination of the Feller and KM schemes (Eqn. 5); S3 and S4 denote the inverse cubic and quartic schemes of Schwartz (Eqns. 6 and 7, respectively); WD denotes the Wilson and Dunning inverse power series (Eqn. 8). The quality of the basis set included within the extrapolation appears in parentheses after the abbreviation. For example, F(DTQ) corresponds to a three-point extrapolation utilizing the Feller extrapolation with the energies from the double- ζ , triple- ζ , and quadruple- ζ basis sets.

Additionally, CCSD(T) single point calculations were performed on the ground states of each atom using the Sapporo-*n*ZP-2012 (*n* = D, T, Q) and cc-pwCV*n*Z (*n* = T, Q, 5) basis sets. The valence and subvalence electrons were included within the correlation space when the core-valence basis sets are used. The same extrapolation schemes were utilized for the correlation energies as were used for HF energies.

RESULTS

Shown in Table 1 are deviations between ROHF/cc-pV*n*Z and numerical HF energies, in mE_h, of the ground and excited states as the size of the cc-pV*n*Z set is increased from triple- ζ to quintuple- ζ . The absolute energies at each basis set level were obtained from Balabanov and Peterson (Ref. 10) and are reproduced in the Supplemental Information. At the quadruple- ζ level, the mean absolute deviation (MAD) from numerical HF in the calculated atomic energies is less than 0.1 mE_h. At the quintuple- ζ level, the MAD is less than 0.02 mE_h, and so can be considered converged at this point. The MADs resulting from exponential schemes are smaller than those obtained using an inverse power scheme, as is expected with the correlation consistent basis sets. (The results from all extrapolation schemes can be found in Table S2 of the Supplemental Information.) All of the exponential extrapolations utilized have an accuracy of 10⁻³ mE_h in comparison to numerical HF energies, with the KM extrapolation resulting in the lowest MAD.

To help assess if the Sapporo-2012 HF set is able to describe the occupied orbitals adequately, the radial distribution functions were computed for the Zn 1*s*, 2*s*, 2*p*, 4*s*, and 3*d* orbitals and compared to those obtained with the cc-pVTZ set in Figure 1. The curves obtained

using each basis set family are almost identical for each orbital type. This indicates that the Sapporo-2012 HF set is able to describe the spatial extent of the orbitals on par with the correlation consistent HF set. However, the orbital energies are different. For example, the Sapporo-TZP-2012 energy of the Zn 1s orbital is -353.289 E_h, while the cc-pVTZ Zn 1s orbital energy is -353.304 E_h. The larger number of functions in cc-pVTZ allows for greater variational flexibility in the wave functions and the additional tight functions, allows a better description of the core orbitals and the nodal structure of the valence orbitals, which results in a lower energy.

As is expected with energies that converge to a limit, the largest improvement in energy recovered with the Sapporo-*n*ZP-2012 sets is upon increasing the basis set from double- ζ to triple- ζ . Typically the MAD from numerical HF can be used as a good metric of the basis set quality for the HF energy, as for the correlation consistent sets. Yet even visual inspection of the convergent behavior of the Sapporo-2012 sets, as reflected by the deviations in Table 2, indicates that the asymptotic limit of these sets is higher in energy than the numerical HF energy. Instead, to determine the quality of the functional fit to the data, a metric by which to judge the “best” extrapolation scheme for the Sapporo sets is the root-mean-squared deviation (RMSD) of the functional fit to the computed energies at each ζ level.

When a two-point extrapolation was performed, the two highest quality basis sets were used (e.g. the triple- ζ and quadruple- ζ sets). The double- ζ and triple- ζ energies were also tested in the two-point extrapolations, but this routinely led to higher extrapolated energies than an extrapolation using triple- ζ and quadruple- ζ energies (see Table S4 in Supplemental Information) and so will not be discussed further. As the extrapolation schemes used in this project were originally intended for use with the correlation consistent basis sets, additional functional forms as well as reoptimization of parameters were considered to determine if a better fit could be obtained. Reoptimization of the parameters generally leads to an extrapolated energy that is clearly lower than the asymptotic limit, so the extrapolation schemes are used as proposed originally, i.e. the unmodified extrapolation techniques. (More information on additional functional forms considered is provided in the Supplemental Information.)

The RMSDs for each extrapolation scheme are shown in Table 3. P, WD, P+KM, and F+KM have the lowest RMSDs, indicating the functional forms fit the calculated data well. Yet visual inspection of the graphs of the calculated and extrapolated energies of Cr and Ni in Figure 2 indicates that, while the WD and the F+KM schemes fit the calculated data well, the convergence behavior of the Sapporo energies is not always well described. The P and the P+KM schemes both fit the calculated data well and also have the correct asymptotic behavior at the limit.

The extrapolated energies obtained with the Sapporo sets are between 16 and 40 mE_h (10 – 25 kcal mol⁻¹) above the numerical HF limit for the ground state energies. The excited state energies converge to limits that are 16 to 80 mE_h (10 – 50 kcal mol⁻¹) above the numerical HF limit. This is in comparison to the range of 0.014 mE_h for the ground state energy limits and

0.017 mE_h for the excited state energy limits obtained using cc-pVnZ sets. The energies obtained with the Sapporo basis sets clearly do not converge to the numerical HF energies, yet they do converge to a limit. Because of this, the incompleteness of extrapolated energies should largely cancel out when relative energies are computed. If extrapolation of HF energies obtained with the Sapporo-2012 family of basis sets is needed, then use of either the Peterson three-point extrapolation or the P+KM extrapolation is suggested.

The lowering of the calculated energy as the basis set quality is increased is small when using the Sapporo-2012 sets, with a decrease of approximately 2 mE_h from double- ζ to quadruple- ζ . In comparison, the energies from the correlation consistent sets decrease by roughly an order of magnitude with each increase in basis set quality. This small change can be attributed to the small changes in the Sapporo basis sets as the quality is increased. For example, when comparing the Sapporo-TZP and Sapporo-QZP sets for Mn, the HF set is essentially unchanged. Most of the differences in the basis sets come from the correlating functions.

The correlation energy was calculated using CCSD(T) to see if the energy recovered by the Sapporo-2012 sets was comparable to that recovered by the cc-pVnZ sets. Overall, there is a much larger decrease in energy as the Sapporo-2012 basis set quality is increased compared with the cc-pVnZ sets; the decrease in energy is generally over 200 mE_h. Most of the decrease in the total energy results from the amount of correlation energy recovered. The RMSDs for each of the extrapolation schemes are listed in Table 4. Extrapolation schemes have been applied to both the total energies and the correlation energies. As the F and KM schemes were originally formulated for HF energies, it is not surprising that the RMSDs for these schemes are much higher than the others, indicating that the schemes do not fit the calculated data very well. The S3 and S4 schemes also are not appropriate for use with the Sapporo-2012 sets, which is reflected by the large RMSDs. The P, WD, P+KM, and F+KM schemes again result in the lowest RMSDs, both when extrapolating total energies and correlation energy alone. The fact that the 3-parameter schemes fit the calculated data well for HF energies, correlation energies, and total energies is due to the flexibility given by the additional parameter. Each of these four schemes has a very low RMSD and the shapes of the curves have the correct shape as well, although the Peterson extrapolation converges too quickly, as shown in Figure 3. Because of this and the clearly incorrect behavior demonstrated by the F+KM and WD schemes when extrapolating HF energies, use of the P+KM scheme is suggested when extrapolating the energies from the Sapporo-*n*ZP-2012 basis set family.

CONCLUSIONS

The convergence behavior of the Sapporo-*n*ZP-2012 family of basis sets was investigated using different functional forms for the extrapolation. Focus was on the 3*d* elements (Sc – Zn), as they can be more challenging to calculate than main-group elements due to the increased number

of valence electrons, the near degeneracy of the $3d$ and $4s$ orbitals, and the often partially occupied $3d$ shell. As well, both Sapporo and correlation consistent sets are widely available for comparison.

As the Sapporo- n ZP-2012 basis sets have been built in a hierarchical fashion, with increasing angular momentum, an increase in the amount of energy recovered is observed when improving the quality of the Sapporo basis set. As such, extrapolations of the energies may be possible to minimize the basis set incompleteness error intrinsic of any finite basis set. This work demonstrates that the energies from the Sapporo basis sets can be extrapolated to a limit, though do not converge to the numerical Hartree-Fock limit in contrast to the correlation consistent basis sets, which nearly reach the numerical HF limit. The limits from the Sapporo sets are 16-40 mE_h above the numerical HF energy, whereas the cc-pVnZ basis set limits deviate by ~0.005 mE_h from the numerical HF energy. As the correlation consistent basis sets include a larger number of functions than the corresponding Sapporo basis set, there is a greater variational flexibility within the wave function, allowing for a lower calculated energy when utilizing the correlation consistent sets. The extrapolated energies obtained with the Sapporo basis sets, however, do minimize the basis set incompleteness error and the remaining error associated with the incompleteness of the basis set may cancel out when computing relative energies.

Extrapolation schemes utilizing three parameters, such as the Wilson and Dunning inverse power scheme or the Peterson mixed Gaussian/exponential scheme, fit the calculated energies well as reflected by the low RMSD. Two new extrapolation formulas have been proposed: a mixed Gaussian/exponential scheme based on the Peterson and Karton and Martin formulas, and a combination of the Feller and Karton and Martin exponential formulas. Of the functional forms considered in this study, the combination of the Peterson and Karton and Martin schemes (P+KM) is recommended for extrapolation of energies for the Sapporo- n ZP-2012 basis set family. The (P+KM) extrapolation form yielded total atomic energies that improved upon the quadruple- ζ level of theory (the highest quality set available within the Sapporo family) by 100 to 250 mE_h, demonstrating the benefit of extrapolating the energies obtained by using the Sapporo- n ZP-2012 basis sets.

Overall, the Sapporo- n ZP-2012 basis sets may provide a reasonable alternative when other hierarchical basis sets, such as the correlation consistent basis sets, that demonstrate convergent behavior for energetic properties with respect to increasing basis set size may not be available.

ACKNOWLEDGEMENTS

This material is based on work supported by the National Science Foundation under grant numbers CHE-1213874 and CHE-1362479. RW acknowledges support from the National Science Foundation GFRP under grant number DGE-1144248. Computing resources were provided by the Academic Computing Services at the University of North Texas. Support from

the United States Department of Energy for the Center for Advanced Scientific Computing and Modeling (CASCaM) is acknowledged.

References

1. Petersson, G. A.; Bennett, A.; Tensfeldt, T. G.; Al-Laham, M. A.; Shirley, W. A.; Mantzaris, J. J. *Chem. Phys.* **1988**, *89*, 2193.
2. Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.
3. Peterson, K. A.; Kendall, R. A.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, *99*, 1930.
4. Peterson, K. A.; Kendall, R. A.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, *99*, 9790.
5. Peterson, K. A.; Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1994**, *100*, 7410.
6. Wilson, A. K.; Dunning, T. H., Jr. *J. Chem. Phys.* **1997**, *106*, 8718.
7. Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.
8. Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, *98*, 1358.
9. Wilson, A. K.; Woon, D. E.; Peterson, K. A.; Dunning, T. H., Jr. *J. Chem. Phys.* **1999**, *110*, 7667.
10. Balabanov, N. B.; Peterson, K. A. *J. Chem. Phys.* **2005**, *123*, 064107.
11. Prascher, B. P.; Woon, D. E.; Peterson, K. A.; Dunning, T. H., Jr.; Wilson, A. K. *Theor. Chem. Acc.* **2011**, *128*, 69.
12. Prascher, B. P.; Wilson, A. K. *Mol. Phys.* **2007**, *105*, 2899.
13. Prascher, B. P.; Wilson, B. R.; Wilson, A. K. *J. Chem. Phys.* **2007**, *127*, 124110.
14. Noro, T.; Sekiya, M.; Koga, T. *Theor. Chem. Acc.* **2012**, *131*, 1124.
15. Sekiya, M.; Noro, T.; Koga, T.; Shimazaki, T. *Theor. Chem. Acc.* **2012**, *131*, 1247.
16. Noro, T.; Sekiya, M.; Koga, T. *Theor. Chem. Acc.* **1997**, *98*, 25.
17. Huzinaga, S.; Miguel, B. *Chem. Phys. Lett.* **1990**, *175*, 289.
18. Tatewaki, H.; Koga, T. *J. Chem. Phys.* **1996**, *104*, 8493.
19. Tatewaki, H.; Koga, T.; Takashima, H. *Theor. Chem. Acc.* **1997**, *96*, 243.
20. Koga, T.; Tatewaki, H.; Matsuyama, H.; Satoh, Y. *Theor. Chem. Acc.* **1999**, *102*, 105.
21. Koga, T.; Yamamoto, S.; Shimazaki, T.; Tatewaki, H. *Theor. Chem. Acc.* **2002**, *108*, 41.
22. Mahler, A.; Wilson, A. K. *J. Chem. Phys.* **2015**, *142*, 084102.
23. Balabanov, N. B.; Peterson, K. A. *J. Chem. Phys.* **2006**, *125*, 074110.
24. Peterson, K. A.; Figgen, D.; Dolg, M.; Stoll, H. *J. Chem. Phys.* **2007**, *126*, 124101.
25. Peterson, K. A.; Puzzarini, C. *Theor. Chem. Acc.* **2005**, *114*, 283.
26. Figri, K., Jr.; Dyall, K. G. In *Basis sets for relativistic calculations*; Schwerdtfeger, P., Ed.; Theoretical and Computational Chemistry; Elsevier Science B. V.: 2002; Vol. 11, pp 259-290.
27. Peterson, K. A. *J. Chem. Phys.* **2015**, *142*, 074105.
28. Fischer, C. F. *The Hartree-Fock Method for Atoms: A Numerical Approach*; John Wiley & Sons, Inc.: New York, USA, 1977; .
29. Tatewaki, H.; Koga, T. *Chem. Phys. Lett.* **1994**, *228*, 562.
30. Jensen, F. *Theor. Chem. Acc.* **2005**, *113*, 187.
31. Jensen, F. *Theor. Chem. Acc.* **2005**, *113*, 267.
32. Feller, D.; Boyle, C. M.; Davidson, E. R. *J. Chem. Phys.* **1987**, *86*, 3424.
33. Halkier, A.; Helgaker, T.; Jørgenson, P.; Klopper, W.; Olsen, J. *Chem. Phys. Lett.* **1999**, *302*, 437.
34. Williams, T. G.; DeYonker, N. J.; Wilson, A. K. *J. Chem. Phys.* **2008**, *128*, 044101.
35. Schwartz, C. *Phys. Rev.* **1962**, *126*, 1015.
36. Feller, D. *J. Chem. Phys.* **1992**, *96*, 6104.
37. Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *J. Chem. Phys.* **1997**, *106*, 9639.
38. Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243.
39. Karton, A.; Martin, J. M. L. *Theor. Chem. Acc.* **2006**, *115*, 330.
40. Jensen, F. *Theor. Chem. Acc.* **2000**, *104*, 484.
41. Karton, A.; Rabinovich, E.; Martin, J. M. L.; Ruscic, B. *J. Chem. Phys.* **2006**, *125*, 144108-1.
42. Feller, D. *J. Chem. Phys.* **1993**, *98*, 7059.
43. Carroll, D. P.; Silverstone, H. J.; Metzger, R. M. *J. Chem. Phys.* **1979**, *71*, 4142.
44. Hill, R. N. *J. Chem. Phys.* **1985**, *83*, 1173.

- 1
2
3
4 45. de Lara-Castells, M. P.; Krems, R. V.; Buchachenko, A. A.; Delgado-Barrio, G.; Villarreal, P. *J.*
5 *Chem. Phys.* **2001**, *115*, 10438.
6 46. DeYonker, N. J.; Wilson, B. R.; Pierpont, A. W.; Cundari, T. R.; Wilson, A. K. *Mol. Phys.* **2009**, *107*,
7 1107.
8 47. Valiev, M.; Bylaska, E. J.; Govind, N.; Kowalski, K.; Straatsma, T. P.; van Dam, H. J. J.; Wang, D.;
9 Nieplocha, J.; Apra, E.; Windus, T. L.; de Jong, W. A. *Comput. Phys. Commun.* **2010**, 1477.
10 48. Dunning, T. H., Jr.; Peterson, K. A.; Wilson, A. K. *J. Chem. Phys.* **2001**, *114*, 9244.
11 49. Wilson, A. K.; Dunning, T. H., Jr. *J. Chem. Phys.* **2003**, *119*, 11712.
12 50. Gilbert, A. T. B.; Besley, N. A.; Gill, P. M. W. *J. Phys. Chem. A* **2008**, *112*, 13164.
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Table 1 – Deviation from the numerical Hartree-Fock (NHF) energies for the ground and excited states of the 3d transition metals (mE_h) calculated at the ROHF/cc-pVnZ ($n = \text{T, Q, 5}$) level of theory using C_1 symmetry. The mean absolute deviation (MAD) from NHF is also reported in mE_h . F indicates the Feller extrapolation; KM indicates the Karton and Martin extrapolation; P indicates the Peterson extrapolation; the letters in parenthesis indicate which basis set levels were included within the extrapolation. The atomic energies are from Ref. 10; the extrapolations have been performed for this work.

Atom	Electronic State		TZ	QZ	5Z	F(TQ5)	F(Q5)	KM(Q5)	P(TQ5)
Sc	[Ar]4s ² 3d	² D	0.268	0.034	0.004	0.000	-0.003	-0.002	-0.008
	[Ar]4s3d ²	⁴ F	0.359	0.047	0.005	-0.002	-0.005	-0.004	-0.013
Ti	[Ar]4s ² 3d ²	³ F	0.386	0.039	0.005	0.001	-0.003	-0.002	-0.006
	[Ar]4s3d ³	⁵ F	0.469	0.051	0.006	0.001	-0.005	-0.003	-0.010
V	[Ar]4s ² 3d ³	⁴ F	0.504	0.045	0.007	0.004	-0.002	-0.001	-0.003
	[Ar]4s3d ⁴	⁶ D	0.582	0.055	0.007	0.002	-0.005	-0.003	-0.007
Cr	[Ar]4s3d ⁵	⁷ S	0.698	0.060	0.012	0.008	0.000	0.002	0.002
	[Ar]4s ² 3d ⁴	⁵ D	0.749	0.069	0.009	0.003	-0.006	-0.003	-0.008
Mn	[Ar]4s ² 3d ⁵	⁶ S	0.880	0.061	0.008	0.004	-0.005	-0.003	0.001
	[Ar]4s3d ⁶	⁶ D	0.970	0.076	0.013	0.008	-0.002	0.000	0.002
Fe	[Ar]4s ² 3d ⁶	⁵ D	1.108	0.076	0.016	0.012	0.001	0.004	0.011
	[Ar]4s3d ⁷	⁵ F	1.209	0.091	0.016	0.011	-0.002	0.001	0.004
Co	[Ar]4s ² 3d ⁷	⁴ F	1.353	0.079	0.012	0.008	-0.004	-0.002	0.011
	[Ar]4s3d ⁸	⁴ F	1.466	0.098	0.012	0.006	-0.009	-0.006	0.001
Ni	[Ar]4s ² 3d ⁸	³ F	1.658	0.098	0.020	0.016	0.001	0.004	0.022
	[Ar]4s3d ⁹	³ D	1.737	0.122	0.021	0.014	-0.004	0.000	0.009
Cu	[Ar]4s3d ¹⁰	² S	1.814	0.126	0.015	0.007	-0.012	-0.008	-0.001
	[Ar]4s ² 3d ⁹	² D	1.414	0.088	0.012	0.007	-0.007	-0.004	0.007
Zn	[Ar]4s ² 3d ¹⁰	¹ S	1.465	0.095	0.016	0.011	-0.003	0.000	-0.010
	[Ar]4s3d ¹⁰ 4p ¹	³ P	1.669	0.152	0.045	0.037	0.019	0.023	0.025
MAD			1.038	0.078	0.013	0.008	0.005	0.004	0.008

Table 2 – Deviation from the numerical Hartree-Fock energies for the ground and excited states of the $3d$ transition metals (mE_h) calculated at the ROHF/Sapporo-2012- n ZP ($n = D, T, Q$) level of theory using C_1 symmetry. The mean absolute deviation (MAD) from numerical HF is also reported.

Atom	Electronic State		DZ	TZ	QZ
Sc	[Ar]4s ² 3d	² D	18.97	18.03	17.41
	[Ar]4s3d ²	⁴ F	19.37	18.30	17.60
Ti	[Ar]4s ² 3d ²	³ F	21.09	19.54	19.25
	[Ar]4s3d ³	⁵ F	21.42	20.35	19.80
V	[Ar]4s ² 3d ³	⁴ F	23.22	21.82	21.64
	[Ar]4s3d ⁴	⁶ D	23.23	21.77	21.60
Cr	[Ar]4s3d ⁵	⁷ S	26.81	25.72	25.61
	[Ar]4s ² 3d ⁴	⁵ D	25.49	23.46	23.09
Mn	[Ar]4s ² 3d ⁵	⁶ S	30.23	29.08	28.62
	[Ar]4s3d ⁶	⁶ D	25.13	23.48	23.19
Fe	[Ar]4s ² 3d ⁶	⁵ D	28.88	27.14	26.55
	[Ar]4s3d ⁷	⁵ F	28.82	27.64	27.30
Co	[Ar]4s ² 3d ⁷	⁴ F	32.63	31.31	30.71
	[Ar]4s3d ⁸	⁴ F	32.35	30.94	30.56
Ni	[Ar]4s ² 3d ⁸	³ F	36.32	34.92	34.26
	[Ar]4s3d ⁹	³ D	35.55	33.62	33.04
Cu	[Ar]4s3d ¹⁰	² S	41.50	40.25	39.73
	[Ar]4s ² 3d ⁹	² D	40.56	38.39	37.55
Zn	[Ar]4s ² 3d ¹⁰	¹ S	42.46	41.36	40.61
	[Ar]4s3d ¹⁰ 4p ¹	³ P	115.36	81.67	80.81
MAD			33.47	30.44	29.95
MAD w/o Zn excited state			29.16	27.75	27.27

Figure 1 – Radial distribution functions for the Zn 1s, 2s, and 2p orbitals (top graph) and 4s and 3d orbitals (bottom graph), using the cc-pVTZ and Sapporo-TZP-2012 basis sets. The distributions obtained from each basis set family are almost completely identical, making it appear as though there is only one line.

Table 3 – The root-mean-squared deviation (RMSD) is reported for each extrapolation scheme when used to extrapolate the ground state energies for the $3d$ transition metals. The energies are calculated at the ROHF/Sapporo- n ZP-2012 ($n = D, T, Q$) level of theory. The following abbreviations are used: F = Feller, KM = Karton and Martin, P = Peterson, WD = Wilson and Dunning, S3 = Schwartz 3, S4 = Schwartz 4, PS3 = average of P and S3, P+KM = combination of P and KM, F+KM = combination of F and KM.

Scheme	Functional Form	RMSD
Feller (B = 1.63)	Exponential	1.96×10^{-4}
Feller (B = 1.0577)	Exponential	1.28×10^{-4}
KM	Exponential	2.65×10^{-4}
Peterson	Mixed Gaussian/exponential	1.59×10^{-9}
WD	Inverse power	6.14×10^{-13}
S3	Inverse power	1.27×10^{-4}
S4	Inverse power	1.34×10^{-4}
P+KM	Mixed Gaussian/exponential	1.75×10^{-12}
F+KM	Sum of exponentials	9.25×10^{-13}

Figure 2 – Total atomic energy (E_h) of Cr (top graph) and Ni (bottom graph), calculated at the ROHF/Sapporo- n ZP-2012 ($n = D, T, Q$) level of theory. The squares indicate the double- ζ , triple- ζ , and quadruple- ζ energies. The following extrapolation schemes are used: P+KM is a combination of aspects of the Peterson form and the Karton and Martin form; F+KM is a combination of the Feller and Karton and Martin extrapolations; P is the Peterson extrapolation; WD is the Wilson and Dunning extrapolation.

Table 4 – The root-mean-squared deviation (RMSD) is reported for each extrapolation scheme when used to extrapolate the ground state energies for the 3*d* transition metals. The energies are calculated at the CCSD(T,FC1)/Sapporo-*n*ZP-2012 (*n* = D, T, Q) level of theory. The following abbreviations are used: F = Feller, KM = Karton and Martin, P = Peterson, WD = Wilson and Dunning, S3 = Schwartz 3, S4 = Schwartz 4, PS3 = average of P and S3, P+KM = combination of P and KM, F+KM = combination of F and KM.

Scheme	Functional Form	Energy	RMSD
Feller (B = 1.63)	Exponential	Total	3.18×10^{-2}
		Correlation	3.17×10^{-2}
Feller (B = 1.0577)	Exponential	Total	1.68×10^{-2}
		Correlation	1.67×10^{-2}
KM	Exponential	Total	6.93×10^{-2}
		Correlation	4.15×10^{-2}
Peterson	Mixed Gaussian/exponential	Total	1.12×10^{-12}
		Correlation	7.04×10^{-16}
WD	Inverse power	Total	4.21×10^{-13}
		Correlation	2.19×10^{-16}
S3	Inverse power	Total	1.60×10^{-2}
		Correlation	1.60×10^{-2}
S4	Inverse power	Total	1.94×10^{-2}
		Correlation	1.93×10^{-2}
P+KM	Mixed Gaussian/exponential	Total	8.04×10^{-13}
		Correlation	4.72×10^{-16}
F+KM	Sum of exponentials	Total	7.96×10^{-13}
		Correlation	8.87×10^{-16}

Figure 3 - Correlation energy (E_h) of Ni (top graph) and total atomic energy of Ni (bottom graph), calculated at the CCSD(T,FC1)/Sapporo- n ZP-2012 ($n = D, T, Q$) level of theory. The squares indicate the double- ζ , triple- ζ , and quadruple- ζ energies. The following extrapolation schemes are used: P+KM is a combination of aspects of the Peterson form and the Karton and Martin form; F+KM is a combination of the Feller and Karton and Martin extrapolations; P is the Peterson extrapolation; WD is the Wilson and Dunning extrapolation.