

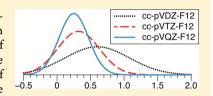
Auxiliary Basis Sets for Density Fitting in Explicitly Correlated Calculations: The Atoms H-Ar

Stella Kritikou and J. Grant Hill*

Department of Chemistry, University of Sheffield, Sheffield S3 7HF, United Kingdom

Supporting Information

ABSTRACT: Auxiliary basis sets specifically matched to the correlation consistent ccpVnZ-F12 and cc-pCVnZ-F12 orbital basis sets for the elements H-Ar have been optimized at the density-fitted second-order Møller-Plesset perturbation theory level of theory for use in explicitly correlated (F12) methods, which utilize density fitting for the evaluation of two-electron integrals. Calculations of the correlation energy for a test set of small to medium sized molecules indicate that the density fitting error when using these auxiliary sets is 2 to 3 orders of magnitude smaller than the F12 orbital basis set



incompleteness error. The error introduced by the use of these fitting sets within the resolution-of-the-identity approximation of the many-electron integrals arising in F12 theory has also been assessed and is demonstrated to be negligible and well-controlled. General guidelines are proposed for the optimization of density fitting auxiliary basis sets for use with F12 methods for other elements.

1. INTRODUCTION

One of the biggest drawbacks of standard correlated electronic structure methods is the slow convergence with respect to oneparticle basis sets. 1-3 Since the introduction of the correlation consistent (cc) family of basis sets,⁴ this error has been addressed by a number of strategies that extrapolate an estimated complete basis set (CBS) limit for a given correlation method. 5-7 Nevertheless, large basis sets of quadruple- ζ or higher quality must be used in order to achieve high accuracy. An efficient route around the basis set problem is the use of explicitly correlated techniques, which include terms that depend explicitly on the interelectronic distance in the wave function. The roots of this method stretch back to the work of Hylleraas on the helium atom, 8,9 but the resulting manyelectron integrals prevented practical applications on larger systems. The introduction of the resolution-of-the-identity (RI) technique, 10 which uses an auxiliary basis set (ABS) to robustly approximate the many-electron integrals, 11 led to the development of the R12 method and a resurgence of interest in explicitly correlated wave functions. The rapid progress in this area has been the subject of several reviews, and interested readers are referred to those publications for further details. 12-16 However, it is important to note that the introduction of a nonlinear correlation factor (the F12 methods) has resulted in explicit correlation becoming almost routine for high-accuracy investigations of small molecular systems. 17-19 F12 has been implemented for a number of correlated wave function methods, including second-order Møller-Plesset perturbation theory (MP2), coupled cluster with single and double excitations (CCSD), complete active space second-order perturbation theory (CASPT2),²⁰ and internally contracted multireference configuration interaction (icMRCI).^{21,22} While conventional correlation consistent basis sets can be used in F12 methods, it has been shown that the use

of Gaussian basis sets optimized specifically for use with explicitly correlated methods produces an additional increase in accuracy. These sets, denoted cc-pVnZ-F12, 23,24 typically produce results of a quality equivalent to that of aug-cc-pV(n)+1)Z (both in F12 calculations). F12-specific basis sets have also been optimized for the treatment of core-valence correlation effects (cc-pCVnZ-F12)²⁵ and paired with pseudopotentials (cc-pVnZ-PP-F12)²⁶ for the post-d main group elements (Ga-Rn).

As alluded to above, a major bottleneck in explicitly correlated methods is the evaluation of numerous three- and four-electron integrals, as well as multiple many-index twoelectron integrals. The former is usually ameliorated by the RI approximation, and in many implementations, an auxiliary basis that is complementary to the orbital basis, leading to the complementary auxiliary basis set (CABS) approach, ²⁷ is used. ABSs optimized for this purpose are suffixed OptRI and are designed to be compact in order to increase computational efficiency.^{28,29} The two-electron integrals are often computed using the density fitting (DF) approximation, 30-32 which has experienced widespread adoption in the quantum chemistry community in the form of the DF-MP2 method.^{33,34} It is noted that this approximation of two-electron integrals is also referred to as RI by some groups. In order to reduce confusion, DF will herein refer to the approximation of two-electron integrals, and RI is reserved for the approximation of many-electron integrals. DF requires an additional ABS, suffixed MP2Fit, optimized for this purpose. In an F12 context, additional two-electron integrals are required in the computation of the V and B matrices (see ref 15 for full details) and several implementations use Manby's generalized robust DF formula for this

Received: August 25, 2015 Published: October 12, 2015 purpose.^{35,36} It has been demonstrated that ABSs optimized for conventional DF-MP2 produce errors of a similar magnitude when used for this robust fitting in DF-MP2-R12,³⁵ suggesting that optimization at the DF-MP2 level combined with careful analysis of the effect on the **V** and **B** matrices should result in MP2Fit sets particularly suitable for use in F12 calculations.

The optimization of ABSs for use in DF-MP2 commonly follows one of two routes: either automatic generation on-thefly using the Cholesky decomposition (CD) and related methods of Aquilante and co-workers^{37–39} or individual optimization matched to a specific orbital basis set (OBS). The latter usually follows the procedure of Weigend et al., later refined by Hättig, ^{34,40,41} which is the approach adopted in the present work. A number of guidelines for the design and optimization of MP2Fit ABSs emerged from these investigations, namely, that the number of functions in the auxiliary basis should be no greater than four times the number of functions in the orbital basis and that the absolute error in the molecular MP2 correlation energy due to density fitting should be less than 100 $\mu E_{\rm h}$ per atom.^{34,40} It has become apparent that accurate density fitting requires functions in the ABS with an angular momentum of at least $l_{\rm occ} + l_{\rm OBS}$, where $l_{\rm occ}$ and $l_{\rm OBS}$ are the highest angular momentum symmetry occupied in the neutral ground-state atom and the largest angular momentum symmetry in the orbital basis, respectively. This rises to $l_{\rm occ}+l_{\rm OBS}+1$ for the hydrogen atom. ^{34,40,41} An additional rule of thumb has also been established: the error due to density fitting in molecules should be at least 2 orders of magnitude smaller than the error due to basis set incompleteness (BSIE).40,41

The design guidelines for MP2Fit sets mentioned above have proven to be successful and led to the development of auxiliary sets matched to the correlation consistent ^{24,40-45} and Karlsruhe segmented contracted (def2) families of basis sets ⁴⁶⁻⁴⁹ for most elements of the periodic table. However, these guidelines have emerged from investigations at the conventional MP2 level, and there is no reason to assume that they will be appropriate in the specific context of explicitly correlated methods. As well as the additional two-electron integrals required in the aforementioned computation of the V and B matrices, BSIE is drastically reduced in explicitly correlated methods, meaning that the fitting accuracy must be increased in order for any density fitting errors to remain negligible. Recently published MP2Fit sets matched to the cc-pVnZ-PP-F12 basis sets for the post-d main group elements Ga-Rn required functions with an angular momentum of $l_{\rm occ} + l_{\rm OBS} + 1$ in order to reach sufficient accuracy. This, in turn, meant that the number of functions in the auxiliary basis reached a maximum of 5.6 times the number of functions in the orbital basis.²⁶ An analysis of density fitting in F12 methods by Hättig et al. concluded that auxiliary sets for this purpose should include functions up to $\max(l_{\text{OBS}}, l_{\text{CABS}}) + l_{\text{occ}}$, where l_{CABS} is the maximum angular momentum of the CABS auxiliary basis set. 15 The same study also noted that MP2Fit-type sets for use with F12 methods should contain tighter functions than those in the analogous conventional case in order to accurately fit the products of two occupied orbitals in the integrals for the V and B matrices.

In the current investigation, new MP2Fit auxiliary sets specifically matched to the cc-pVnZ-F12 and cc-pCVnZ-F12 orbital sets for the elements H—Ar have been developed, and new guidelines for the design of MP2Fit sets to be used in F12 calculations are proposed. The resulting sets are validated using

a number of criteria both for atoms and a test set of 104 small to medium sized molecules.

2. BASIS SET CONSTRUCTION AND OPTIMIZATION

MP2Fit auxiliary sets specifically matched to the correlation consistent cc-pVnZ-F12 and cc-pCVnZ-F12 (where n=D, T, Q) basis sets for the elements He–Ar²³-25 have been optimized using the TURBOMOLE program. Although the resulting sets are intended for use in F12 calculations, they have been optimized at the conventional DF-MP2 level. This is partly due to a lack of access to an efficient "density fitting free" MP2-F12 code, but it has been demonstrated elsewhere that this approach can lead to accurate and efficient fitting sets. The optimization followed the approach of Hättig, ⁴¹ where the following functional, $\delta_{\rm DF}$, is minimized for neutral ground-state atoms using the analytical ABS gradients available in RICC2 module of TURBOMOLE ^{41,51,52}

$$\delta_{\rm DF} = \frac{1}{4} \sum_{aibj} \frac{(\langle ab|lij \rangle_{\rm DF} - \langle ab|lij \rangle)^2}{\epsilon_a - \epsilon_i + \epsilon_b - \epsilon_j} \tag{1}$$

where $\langle ab|lij \rangle = (ai|bj) - (aj|bi)$, with i and j denoting occupied orbitals, a and b denoting virtual orbitals, and ϵ_x denoting the HF orbital energies.³⁴

Technical difficulties were encountered in attempting to optimize an MP2Fit auxiliary basis for the H2 molecule in TURBOMOLE; hence, the exponents were optimized using the MOLPRO^{53,54} package. This proceeded by minimizing the density fitting error in the MP2 correlation energy, $\Delta DF = |E_{MP2}^{corr}|$ E_{DF-MP2}l, using either the Broyden-Fletcher-Goldfarb-Shanno (BFGS) conjugate gradient method or Nelder-Mead simplex algorithm for a hydrogen-hydrogen bond length of 1.40 a_0 . 55,86 In all cases, the new sets were designed so that the number of basis functions within a particular ABS will remain the same for all elements belonging to a single row of the periodic table. For the cc-pVnZ-F12 sets, the standard definition of the frozen core approximation (the 1s electrons for Li-Ne and 1s2s2p electrons for Na-Ar) was used, with those electrons excluded from the correlation treatment and the evaluation of δ_{DF} . The exceptions were Li, where all electrons were correlated during the ABS optimizations, and Na, where only the 1s electrons were frozen. The standard frozen core definitions were used for Li and Na during the subsequent testing of the resulting ABSs. The cc-pCVnZ-F12 series of basis sets for the atoms Li-Ar are designed for core-valence correlation, with only the 1s electrons of Na-Ar excluded from the correlation treatment. MP2Fit auxiliary sets matched to other conventional core-valence correlation consistent basis sets have typically fixed the exponents of the analogous ccpVnZ/MP2Fit set while augmenting it with a number of additional tight functions. 41 In the present investigation, initial testing demonstrated that reaching the desired level of accuracy with such an approach led to auxiliary sets with a very large number of functions; thus, the decision was made to completely reoptimize all exponents for cc-pCVnZ-F12/MP2Fit.

In the density fitting approximation of four-index electron repulsion integrals (ERIs), three index ERIs of the type (ailP) are evaluated, where P denotes auxiliary basis functions. The follows from symmetry considerations that only P with angular momentum equal to or less than $l_{\rm corr}+l_{\rm OBS}$ will possess a gradient when evaluating $\delta_{\rm DF}$ for atoms. However, initial molecular tests revealed that P with angular momentum equal to $l_{\rm occ}+l_{\rm OBS}+1$ was important for accurate fitting. Such

Table 1. Composition of the Valence-Only cc-pVnZ-F12 Auxiliary MP2Fit and Orbital Basis Sets for the Elements H-Ar^a

element	basis set designation	OBS	MP2Fit	ratio
H/He	cc-pVDZ-F12	[3s2p]/[4s2p]	(6s4p2d1f)	3.9/3.5
	cc-pVTZ-F12	[4s3p1d]/[5s3p1d]	(7s5p3d2f1g)	3.3/3.2
	cc-pVQZ-F12	[5s4p2d1f]/[6s4p2d1f]	(8s6p4d3f2g1h)	2.8/2.7
Li–Ne	cc-pVDZ-F12	[5s5p2d]	(9s8p7d3f2g)	3.6
	cc-pVTZ-F12	[6s6p3d2f]	(11s10p9d5f4g1h)	3.2
	cc-pVQZ-F12	[7s7p4d3f2g]	(13s12p11d7f6g3h1i)	2.9
Na-Ar	cc-pVDZ-F12	[6s6p3d]	(12s10p9d5f2g)	3.6
	cc-pVTZ-F12	[7s7p4d2f]	(14s12p11d7f4g1h)	3.2
	cc-pVQZ-F12	[8s8p5d3f2g]	(16s14p13d9f6g3h1i)	3.0

^aThe ratio of functions (ABS/OBS) assumes spherical angular momenta.

Table 2. Composition of the Core-Valence cc-pCVnZ-F12 Auxiliary MP2Fit and Orbital Basis Sets for the Elements Li-Ar

element	basis set designation	OBS	MP2Fit	ratio
Li–Ne	cc-pCVDZ-F12	[6s6p2d]	(11s10p8d3f2g)	3.6
	cc-pCVTZ-F12	[7s7p4d2f]	(13s12p10d6f4g1h)	3.0
	cc-pCVQZ-F12	[8s8p5d3f2g]	(16s15p13d9f7g3h1i)	3.1
Na-Ar	cc-pCVDZ-F12	[7s7p4d]	(14s12p10d6f2g)	3.6
	cc-pCVTZ-F12	[8s8p5d3f]	(16s14p12d8f5g1h)	2.8
	cc-pCVQZ-F12	[9s9p6d4f2g]	(18s16p14d10f7gh1i)	2.8
aThe ratio of functions (ARC/ORC) assumes spherical angular	mamanta		

^aThe ratio of functions (ABS/OBS) assumes spherical angular momenta.

Table 3. Errors in the MP2 Density Fitting for cc-pVnZ-F12 (He-Ar) and cc-pCVnZ-F12 (Li-Ar), as Errors in Correlation Energy (Δ DF, per Correlated Electron) and Two-Electron Integrals ($\delta_{\rm DF}/|E_{\rm MP2}^{\rm corr}|$, per Atom) for the Auxiliary Basis Sets Developed in This Work

		$\Delta { m DF} \left({\mu E_{ m h}} ight)$			$\delta_{ m DF}/ E_{ m MP2}^{ m corr} $		
basis	MUE	σ	MAX	MUE	σ	MAX	
cc-pVDZ-F12	0.20	0.21	0.73	2.09×10^{-9}	2.41×10^{-9}	8.37×10^{-9}	
cc-pVTZ-F12	0.09	0.12	0.47	9.73×10^{-10}	1.05×10^{-9}	3.17×10^{-9}	
cc-pVQZ-F12	0.05	0.07	0.31	7.54×10^{-10}	5.73×10^{-10}	1.80×10^{-9}	
cc-pCVDZ-F12	0.25	0.20	0.70	3.11×10^{-9}	2.44×10^{-9}	8.42×10^{-9}	
cc-pCVTZ-F12	0.19	0.14	0.49	1.76×10^{-9}	1.10×10^{-9}	3.92×10^{-9}	
cc-pCVQZ-F12	0.03	0.03	0.11	9.72×10^{-10}	8.12×10^{-10}	2.52×10^{-9}	

functions without a gradient were interpolated from the previous angular momentum symmetry as

$$\alpha_l = c \cdot \sqrt{\alpha_{(l-1),x} \alpha_{(l-1),y}} \tag{2}$$

where c is a scalar coefficient (taking the value 1, 2, or 0.5). The exact choice of c and which two exponents (x,y) of the previous angular momentum symmetry used in each interpolation was somewhat guided by molecular tests, but in the majority of cases, c = 1 and the two most diffuse exponents were selected.

The compositions of the final MP2Fit sets developed in this work are shown in Table 1 for cc-pVnZ-F12 and Table 2 for ccpCVnZ-F12. The compositions of the corresponding orbital sets are also shown for reference and to provide context to the size of the auxiliary sets. As with previous MP2Fit ABSs matched to correlation consistent basis sets, all of the ABSs developed in this work are uncontracted. Initial testing indicated that a general contraction scheme leads to relatively large contraction errors. Indeed, reaching the target accuracy (see below) including general contraction required a large number of ABS primitives and thus offered no computational advantage. The maximum angular momentum for every MP2Fit set is equal to $l_{\rm occ} + l_{\rm OBS} + 1$, which is consistent with the equivalent fitting sets matched to the cc-pVnZ-PP-F12 sets for the post-d main group elements. 26 The requirement for the additional shell of higher angular momentum functions

relative to the $l_{\rm occ}$ + $l_{\rm OBS}$ rule of thumb was clearly evident from initial molecular testing, a finding that was also noted by Tanaka et al. when optimizing MP2Fit sets matched to 6-31G(d,p) and 6-311G(d,p) orbital sets.⁵⁸ The maximum angular momentum symmetry for each MP2Fit set developed in this work can also be compared with the suggestion of Hättig et al. 15 that functions of at least $l_{\text{MP2Fit}} = \max(l_{\text{OBS}}, l_{\text{CABS}}) + l_{\text{occ}}$ should be included for density fitting in explicitly correlated methods. By assuming that the CABS basis is the appropriate OptRI set of Yousaf and Peterson, ²⁸ the MP2Fit sets developed in this work meet or exceed this suggestion in all cases except the double- ζ sets for Li–Ar, where g functions in the OptRI basis indicate that it may be necessary to include h functions in the MP2Fit (which, in the present case, includes only up to g functions). Initial testing of cc-pVDZ-F12/MP2Fit sets that included interpolated h functions demonstrated a negligible change both in molecular correlation energies and the functional evaluating the accuracy of the RI (see below). Pragmatically, the inclusion of h functions for cc-pVDZ-F12/ MP2Fit appears to increase computation time for very little benefit; hence, the decision was made to exclude them.

The exponents of the MP2Fit sets developed in this work are presented in the Supporting Information. A comparison of the exponents with those of auxiliary basis sets matched to other correlation consistent basis sets reveals that they are generally

Table 4. RI Errors Due to Density Fitting (Per Correlated Electron, Relative to a Large Even-Tempered Reference Set) at the MP2-F12 Level for Valence-Only Calculations on the Atoms He, Be-Ne, and Mg-Ar Using the MP2Fit Auxiliary Basis Sets Developed in This Work^a

				$\delta_{ m RI} \; (E_{ m h})$	
orbital basis	MP2Fit	$I_{ m MP2Fit}^{ m MAX}$	MUE	σ	MAX
cc-pVDZ-F12	cc-pVDZ-F12	4	4.49×10^{-10}	3.37×10^{-10}	1.19×10^{-9}
	aug-cc-pVTZ	4	2.38×10^{-9}	2.71×10^{-9}	9.78×10^{-9}
	aug-cc-pwCVDZ	3	7.40×10^{-8}	6.47×10^{-8}	2.16×10^{-7}
	aug-cc-pwCVTZ	4	7.12×10^{-10}	3.95×10^{-10}	1.24×10^{-9}
cc-pVTZ-F12	cc-pVTZ-F12	5	1.23×10^{-10}	1.55×10^{-10}	5.77×10^{-10}
	aug-cc-pVQZ	5	5.49×10^{-9}	1.49×10^{-8}	5.92×10^{-8}
	aug-cc-pwCVTZ	4	2.78×10^{-7}	6.58×10^{-7}	2.56×10^{-6}
	aug-cc-pwCVQZ	5	4.08×10^{-9}	1.53×10^{-8}	5.92×10^{-8}
cc-pVQZ-F12	cc-pVQZ-F12	6	5.55×10^{-11}	4.76×10^{-11}	1.26×10^{-10}
•	aug-cc-pV5Z	6	1.16×10^{-9}	8.75×10^{-10}	3.10×10^{-9}
	aug-cc-pwCVQZ	5	8.78×10^{-8}	2.03×10^{-7}	6.81×10^{-7}
	aug-cc-pwCV5Z	6	3.27×10^{-10}	7.72×10^{-10}	3.10×10^{-9}

^aSeveral other MP2Fit sets are included for comparison purposes; see text for further details.

Table 5. RI Errors Due to Density Fitting (Per Correlated Electron, Relative to a Large Even-Tempered Reference Set) at the MP2-F12 Level for Core—Valence Calculations on the Atoms Li—Ar Using the MP2Fit Auxiliary Basis Sets Developed in This Work^a

			$\delta_{ m RI} \; (E_{ m h})$	
orbital basis	MP2Fit	MUE	σ	MAX
cc-pCVDZ-F12	cc-pCVDZ-F12	4.59×10^{-9}	7.04×10^{-9}	2.30×10^{-8}
	aug-cc-pwCVDZ	1.28×10^{-6}	2.04×10^{-6}	7.18×10^{-6}
	aug-cc-pwCVTZ	1.26×10^{-8}	1.94×10^{-8}	5.19×10^{-8}
cc-pCVTZ-F12	cc-pCVTZ-F12	1.88×10^{-10}	2.06×10^{-10}	7.15×10^{-10}
	aug-cc-pwCVTZ	1.69×10^{-5}	2.24×10^{-5}	6.02×10^{-5}
	aug-cc-pwCVQZ	2.68×10^{-9}	6.04×10^{-9}	1.80×10^{-8}
cc-pCVQZ-F12	cc-pCVQZ-F12	8.06×10^{-11}	1.27×10^{-10}	5.24×10^{-10}
	aug-cc-pwCVQZ	7.41×10^{-8}	1.45×10^{-7}	5.59×10^{-7}
	aug-cc-pwCV5Z	1.63×10^{-10}	2.55×10^{-10}	1.06×10^{-9}

[&]quot;Several other MP2Fit sets are included for comparison purposes; see text for further details.

somewhat tighter, for example, the tightest s exponent in the cc-pVQZ-F12/MP2Fit set for Ne is roughly 1235, compared to 812 for aug-cc-pV5Z/MP2Fit. The ratio between the number of functions in the MP2Fit auxiliary sets developed in this work and the given orbital set is displayed in Tables 1 and 2. It can be seen that this ratio ranges between 2.8 and 3.9, which is less than the guideline of 4.0 established for density fitting in conventional MP2. This ratio is somewhat lower than the equivalent required for accurate fitting of the cc-pVnZ-PP-F12 sets for the post-d main group elements, ²⁶ perhaps as a result of the requirement for the latter sets to correlate an outer core of d electrons in addition to the valence s and p.

3. RESULTS AND DISCUSSION

The performance of the new MP2Fit auxiliary sets has been initially validated at the atomic level by inspection of both the error in the conventional MP2 correlation energy (Δ DF) and the quantity $\delta_{\rm DF}/|E_{\rm MP2}^{\rm corr}|$, which represents the error in the two-electron integrals at the conventional MP2 level. This data is presented in Table 3 as the mean unsigned error (MUE), the standard deviation of the error (σ), and the maximum absolute error (MAX), with hydrogen excluded since it can be treated only at a molecular level with MP2. It can be seen that the error in the correlation energy decreases as the basis set cardinal number is increased. The mean unsigned Δ DF error is always

equal to or less than 0.2 $\mu E_{\rm h}$ (per correlated electron) for the valence-only sets and is less than 0.3 $\mu E_{\rm h}$ for core—valence. The maximum error is always less than 1 $\mu E_{\rm h}$ per correlated electron, which is smaller than standard criteria for atomic density fitting of 50 $\mu E_{\rm h}$. The error in the integrals represented by $\delta_{\rm DF}/|E_{\rm MP2}^{\rm corr}|$ also decreases with basis set cardinal number and is several orders of magnitude smaller than the 1 \times 10⁻⁶ guideline upper limit for density fitting in conventional methods. The combination of $\Delta {\rm DF}$ and $\delta_{\rm DF}/|E_{\rm MP2}^{\rm corr}|$ indicates that density fitting with the sets developed in this work introduces negligible errors for conventional MP2 at the atomic level.

As outlined above, density fitting is also used in F12 methods for additional two-electron integrals that appear within the intermediate V and B matrices of the RI; hence, it is important to ensure that the choice of density fitting ABS does not introduce any significant errors at this stage. A functional that has been previously used to evaluate the accuracy of the RI is known as $\delta_{\rm RI}$

$$\delta_{\text{RI}} = \sum_{ij} \frac{(V_{ij,ij} - V_{ij,ij}^{\text{ref}})^2}{|V_{ij,ij}^{\text{ref}}|} + \frac{(B_{ij,ij} - B_{ij,ij}^{\text{ref}})^2}{|B_{ij,ij}^{\text{ref}}|}$$
(3)

where the superscript ref indicates that the diagonal elements of the V and B matrices have been computed using a large reference set.²⁸ δ_{RI} has units of energy and is always positive. Although the form of the δ_{RI} functional above appears to differ slightly from that published for the optimization of OptRI auxiliary sets (it was previously shown without taking the absolute values of the diagonal elements of the V and B matrices in the denominator), $^{28,29}_{\text{Pl}}$ previous investigations did calculate δ_{RI} as in eq 3 above. 59 If the CABS auxiliary basis is held fixed while the MP2Fit basis is varied relative to a large reference set, then the effect of density fitting on the RI can be elucidated. A locally modified version of the MOLPRO program was used to calculate $\delta_{\rm RI}$ with the large even-tempered sets of Hill et al.^{24,60} used as the reference density fitting ABS. These uncontracted sets have a 21s18p14d12f10g8h6i composition for H and He, 28s26p22d22f20g18h15i for Li-Ne, and 29s27p23d20f18g17h15i for Na-Ar. Full technical details of the MP2-F12 calculations, including fitting sets, ^{28,61} used in the calculation of δ_{RI} are included in the Supporting Information. The effect of density fitting on the RI is presented in terms of δ_{RI} in Tables 4 and 5 for the cc-pVnZ-F12 and cc-pCVnZ-F12 basis sets, respectively. In addition to the MP2Fit auxiliary sets developed in this work, results are also presented for a number of other MP2Fit sets that have previously been recommended for this purpose. 15,25,26,60 The atoms H, Li, and Na are not included in the error statistics of Table 4 because they possess only one valence electron, and while it would be possible to correlate core electrons for Li and Na in order to produce values for the RI errors, the cc-pVnZ-F12 basis sets are not designed for this purpose, and reporting such values would not be indicative of the performance that may be expected in general applications.

Examining Tables 4 and 5 and focusing momentarily on the MP2Fit sets developed in this work, it can be seen that as the basis set cardinal number increases the δ_{RI} MUE, σ , and MAX decrease. If a cc-pVnZ-F12 set is compared with the cc-pCVnZ-F12 set with the same cardinal number, then it is evident that the errors per correlated electron are roughly equivalent, indicating a similar level of accuracy in the RI approximation. When considering Table 4, it should be noted that the aug-ccpVnZ auxiliary set was necessarily used for He when aug-ccpwCVnZ was used for all other elements, leading to MAX errors that do not appear to change between the aug-cc-pVnZ and aug-cc-pwCVnZ fitting sets. Overall, the errors in $\delta_{\rm RI}$ due to the use of the MP2Fit sets developed in this work appear to be well-controlled, with a mean average error of less than 0.5 n E_h for the valence sets and less than 5.0 n E_h for core—valence. The maximum errors are also small, always less than 23 n E_h .

Table 4 shows the maximum angular momentum functions included in each MP2Fit basis as $I_{\text{MP2Fit}}^{\text{MAX}}$ and also compares the results from MP2Fit sets developed in this work with those that have commonly been used for density fitting in F12 methods in the past. It can be seen that including functions with an angular momentum of $l_{\text{occ}} + l_{\text{OBS}} + 1$ reduces the errors in the RI by between 1 and 2 orders of magnitude, whereas the addition of tight functions (compare aug-cc-pVnZ/MP2Fit with aug-ccpwCVnZ/MP2Fit) also significantly reduces the average error. Taking into account Tables 4 and 5, it is evident that the sets developed in this work introduce significantly smaller errors than those used previously, with the cc-p(C)VnZ-F12 fitting sets performing better than aug-cc-pwCV(n+1)Z (which possesses a significantly greater number of auxiliary basis functions; see Table SI in the Supporting Information). Although it is not particularly surprising that fitting sets

specifically matched to a given orbital set outperform those designed for a different purpose, the data indicates that the sets developed in this work should ensure an increase in both accuracy and efficiency.

In addition to the atomic validation of the new MP2Fit sets presented above, it is vital to ensure that the errors due to density fitting are also negligible for molecules. A standard approach from the literature is to compare the density fitting error at the MP2 level with the BSIE at the same level of theory and basis set, with a rule of thumb being that ΔDF should be at least 2 orders of magnitude smaller than the BSIE. ^{40,41} In order to modify this approach for the reduced BSIE expected from explicitly correlated calculations, Tables 6 and 7 compare ΔDF

Table 6. cc-pVnZ-F12 Valence-Only Correlation Energy Errors ($\mu E_{\rm h}$, per Correlated Electron) Due to MP2-F12 Explicitly Correlated Basis Set Incompleteness Error (BSIE) and Conventional MP2 Density Fitting (Δ DF) Using the Auxiliary Sets Developed in This Work For a Test Set of Molecules Containing Elements H–Ar

orbital basis	error type	MUE	σ	MAX
cc-pVDZ-F12	BSIE	590.37	292.79	1189.00
	$\Delta \mathrm{DF}$	0.60	0.48	2.01
cc-pVTZ-F12	BSIE	173.32	87.92	367.64
	$\Delta \mathrm{DF}$	0.32	0.31	1.95
cc-pVQZ-F12	BSIE	49.53	25.12	105.05
	$\Delta \mathrm{DF}$	0.25	0.22	1.18

Table 7. cc-pCVnZ-F12 Core—Valence Correlation Energy Errors ($\mu E_{\rm h}$, per Correlated Electron) Due to MP2-F12 Explicitly Correlated Basis Set Incompleteness Error (BSIE) and Conventional MP2 Density Fitting (Δ DF) Using the Auxiliary Sets Developed in This Work For a Test Set of Molecules Containing Elements H—Ar

orbital basis	error type	MUE	σ	MAX
cc-pCVDZ-F12	BSIE	1271.51	792.45	3971.28
	$\Delta \mathrm{DF}$	0.50	0.39	1.78
cc-pCVTZ-F12	BSIE	276.36	178.32	1200.66
	$\Delta \mathrm{DF}$	0.32	0.25	1.00
cc-pCVQZ-F12	BSIE	78.97	50.95	343.07
	$\Delta \mathrm{DF}$	0.17	0.13	0.65

at the conventional MP2 level with BSIE from MP2-F12 calculations with the same basis sets. The BSIE was obtained by estimating the CBS limit using a Schwenke-type extrapolation of

$$E_{\text{CBS}}^{\text{corr}} = (E_{\text{QZ}}^{\text{corr}} - E_{\text{TZ}}^{\text{corr}})F + E_{\text{TZ}}^{\text{corr}}$$
(4)

where QZ and TZ refer to the zeta level of the basis sets and F takes the value 1.4148. Separate estimates of the CBS limit were produced for valence-only (cc-pVnZ-F12) and corevalence (cc-pCVnZ-F12) correlation. The value of F used has not been well-tested for core-valence correlation, but it should prove to be sufficient for the purposes of estimating BSIE rather than establishing benchmark data on the CBS limits. The correlation energies used to determine the BSIE were evaluated at the (R)MP2-F12 level 2,63 with the 3C(FIX) Ansatz 1 in the MOLPRO program. Full technical details of the calculations are provided in the Supporting Information. The test set of molecules used to evaluate the performance of the new MP2Fit sets is a subgroup of those from the work of Weigend (with the additions of Ne, Ar, and O₂) and consists of 104 small to

Journal of Chemical Theory and Computation

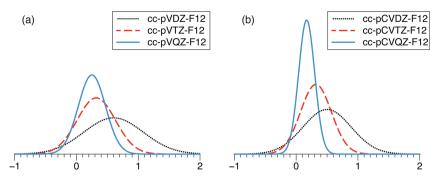


Figure 1. Gaussian distributions of the Δ DF error (μ E_b, per correlated electron) in MP2 correlation energy for a test set of molecules that include elements H–Ar, using (a) valence cc-pVnZ-F12 basis sets and (b) core–valence cc-pCVnZ-F12 basis sets.

medium sized molecules (56 molecules containing only first-row elements, and 48 molecules containing both first- and second-row elements). 65

Tables 6 and 7 immediately show that, even though MP2-F12 greatly accelerates convergence with respect to basis set, the basis set incompleteness error is still significantly larger than that due to density fitting. In terms of ΔDF , all of the error statistics presented are 2 to 3 orders of magnitude smaller than the error in the orbital basis; hence, the density fitting errors due to the auxiliary sets developed in this work can be considered to be negligible. As was observed at the atomic level, as the basis set cardinal number is increased all of the error statistics related to ΔDF are reduced. Comparing Table 6 with Table 7 indicates that while, as expected, the BSIE increases when core–electron correlation is included, ΔDF actually decreases, perhaps indicating particularly accurate density fitting for core-valence in this case. A comparison of ΔDF from Tables 6 and 7 with those reported for the cc-pVnZ-PP-F12 sets for the post-d main group elements shows that the fitting errors are of the same magnitude.26

The accuracy of the density fitting for the molecular test set is also presented visually as normalized Gaussian distributions in Figure 1, where the tight grouping of the Gaussians reflects the high accuracy at all zeta levels. For both the valence-only and core–valence correlation cases, an increase in accuracy as the basis set cardinal number is also increased can be clearly seen. The ΔDF error per correlated electron for each molecule in the test set is presented in bar chart form in Figures SI–SIV in the Supporting Information. Inspection of the heights of the bars indicates that the ΔDF error is reasonably consistent across all of the molecules, and it does not appear that the fitting accuracy is noticeably worse for any given element.

To provide a rough indication of the performance of the newly developed auxiliary basis sets in terms of computational time, MP2-F12 calculations were carried out on the glycine molecule using the cc-pVnZ-F12 and cc-pCVnZ-F12 basis sets. In valence-only calculations, relative to the aug-cc-pwCV(n+1)Z MP2Fit auxiliary basis set, the auxiliary basis sets developed in this work produced a decrease in the wall-time taken to evaluate the F12 integrals of up to 17.3% (cc-pVDZ-F12), with cc-pVTZ-F12 and cc-pVQZ-F12 savings of 13.2 and 13.6%, respectively. In calculations with core-valence electron correlation using the larger cc-pCVnZ-F12 basis sets, the savings were 6.3, 6.5, and 7.0% (with increasing basis set cardinal number), again relative to the aug-cc-pwCV(n+1)Z MP2Fit auxiliary basis set. All calculations were performed direct and without symmetry on a single core of a 2.6 GHz Intel Xeon E5-2650 CPU.

Combining the results of the present investigation with those from the optimization of auxiliary basis sets for cc-pVnZ-PP-F12 allows for a set of guidelines to be proposed for the optimization of MP2Fit auxiliary sets for use in explicitly correlated calculations. These guidelines are very much in the spirit of those that emerged for conventional MP2 density fitting,³⁴ but they are revised and expanded in order to reflect the requirements of density fitting in explicitly correlated methods:

- (1) The number of auxiliary basis functions should not be greater than $6 \times N_{\rm OBS}$ (number of orbital basis set functions).
- (2) The ΔDF error in atoms should be less than 1.0 $\mu E_{\rm h}$ per correlated electron, with an optimization criteria of $\delta_{\rm DF}/|E_{\rm MP2}^{\rm corr}| \leq 10^{-8}$.
- (3) To ensure that there are negligible errors due to density fitting within the RI approximation of many-electron integrals, $\delta_{\rm RI} \leq 5 \times 10^{-8} \, E_{\rm h}$ (per correlated electron).
- (4) In molecular tests, the ΔDF error per correlated electron at the conventional MP2 level should be at least 2 orders of magnitude smaller than the BSIE at the explicitly correlated MP2-F12 level.

We note that the above guidelines may need to be revisited for suitability should basis sets be optimized specifically for use in explicitly correlated methods for transition metals or the lanthanides and actinides.

4. CONCLUSIONS

New MP2Fit auxiliary basis sets specifically matched to the ccpVnZ-F12 and cc-pCVnZ-F12 families of correlation consistent basis sets for the atoms H-Ar have been optimized for use in the density fitting approximation of two-electron integrals in explicitly correlated methods. The number of functions in the auxiliary sets is always less than four times the number of functions in the orbital sets, and the error both in the atomic MP2 correlation energy and in the atomic electron repulsion integrals decreases as the basis set cardinal number is increased. The error introduced by using the auxiliary basis sets to evaluate the two-electron terms within the F12 resolution-ofthe-identity of many-electron integrals has also been demonstrated to be negligible. At the molecular level, the density fitting error using the new auxiliary sets has been determined with the (DF-)MP2 method and compared to the orbital basis set incompleteness error (at the MP2-F12 level) for a test set of 104 small to medium sized molecules. The mean unsigned error, standard deviation, and maximum error of the density fitting is always 2 to 3 orders of magnitude smaller than the BSIE, indicating the accuracy and suitability of the auxiliary sets

for F12 methods such as MP2-F12 and approximate CCSD-(T)-F12. Again, an increase in accuracy of the density fitting is observed as the basis set cardinal number is increased. Comparison with auxiliary basis sets previously developed for use with the cc-pVnZ-PP-F12 sets for post-d main group elements has allowed for a set of guidelines to be formulated for the optimization of MP2Fit auxiliary basis sets, which are likely to be useful as new F12-specific basis sets are developed for other elements of the periodic table.

All of the ABSs optimized in this work can be found in the Supporting Information and will be made available for electronic download from the Basis Set Exchange Web site (https://bse.pnl.gov/bse/portal; accessed October 9, 2015). 66,67

ASSOCIATED CONTENT

S Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b00816.

Auxiliary basis sets in MOLPRO format, density fitting errors in MP2 correlation energies for all molecules in the test set, Gaussian distributions of the density fitting error with different MP2Fit auxiliary sets, a comparison of the number of functions in different MP2Fit auxiliary sets, and full technical details of explicitly correlated calculations (PDF)

Auxiliary basis sets in machine-readable MOLPRO format (TXT)

AUTHOR INFORMATION

Corresponding Author

*Phone: +44 (0)114 2229392. Fax: +44 (0)114 2229346 E-mail: grant.hill@sheffield.ac.uk.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors acknowledge the EPSRC UK National Service for Computational Chemistry Software (NSCCS) at Imperial College London for providing access to the TURBOMOLE package. The authors also thank Prof. Kirk A. Peterson (Washington State) for initial technical assistance in the calculation of $\delta_{\rm RI}$.

REFERENCES

- (1) Helgaker, T.; Klopper, W.; Tew, D. P. Mol. Phys. 2008, 106, 2107-2143.
- (2) Jensen, F. WIREs Comput. Mol. Sci. 2013, 3, 273-295.
- (3) Hill, J. G. Int. J. Quantum Chem. 2013, 113, 21-34.
- (4) Dunning, T. H., Jr. J. Chem. Phys. 1989, 90, 1007-1023.
- (5) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. J. Chem. Phys. 1997, 106, 9639-9646.
- (6) Schwenke, D. J. Chem. Phys. 2005, 122, 014107.
- (7) Feller, D.; Peterson, K. A.; Hill, J. G. J. Chem. Phys. 2011, 135, 044102.
- (8) Hylleraas, E. A. Z. Phys. A 1928, 48, 469-494.
- (9) Hylleraas, E. A. Z. Phys. A 1929, 54, 347–366.
- (10) Kutzelnigg, W. Theor. Chim. Acta 1985, 68, 445-469.
- (11) Klopper, W.; Samson, C. C. M. J. Chem. Phys. 2002, 116, 6397—6410.
- (12) Klopper, W.; Manby, F. R.; Ten-no, S.; Valeev, E. F. Int. Rev. Phys. Chem. **2006**, 25, 427–468.

- (13) Tew, D. P.; Hättig, C.; Bachorz, R.; Klopper, W. In Recent Progress in Coupled Cluster Methods: Theory and Applications; Čársky, P., Paldus, J., Pittner, J., Eds.; Springer: Berlin, 2010; pp 535–572.
- (14) Werner, H.-J; Adler, T. B.; Knizia, G.; Manby, F. R. In Recent Progress in Coupled Cluster Methods: Theory and Applications; Čársky, P., Paldus, J., Pittner, J., Eds.; Springer: Berlin, 2010; pp 573–619.
- (15) Hättig, C.; Klopper, W.; Köhn, A.; Tew, D. P. Chem. Rev. 2012, 112, 4-74.
- (16) Kong, L.; Bischoff, F. A.; Valeev, E. F. Chem. Rev. 2012, 112, 75–107.
- (17) Ten-no, S. Chem. Phys. Lett. 2004, 398, 56-61.
- (18) Tew, D. P.; Klopper, W. J. Chem. Phys. 2005, 123, 074101.
- (19) May, A. J.; Valeev, E. F.; Polly, R.; Manby, F. R. Phys. Chem. Chem. Phys. **2005**, 7, 2710–2713.
- (20) Shiozaki, T.; Werner, H.-J. J. Chem. Phys. 2010, 133, 141103.
- (21) Shiozaki, T.; Knizia, G.; Werner, H.-J. J. Chem. Phys. 2011, 134, 034113.
- (22) Shiozaki, T.; Werner, H.-J. Mol. Phys. 2013, 111, 607-630.
- (23) Peterson, K. A.; Adler, T.; Werner, H.-J. *J. Chem. Phys.* **2008**, 128, 084102.
- (24) Hill, J. G.; Peterson, K. A. Phys. Chem. Chem. Phys. 2010, 12, 10460–10468.
- (25) Hill, J. G.; Mazumder, S.; Peterson, K. A. J. Chem. Phys. 2010, 132, 054108.
- (26) Hill, J. G.; Peterson, K. A. J. Chem. Phys. 2014, 141, 094106.
- (27) Valeev, E. F. Chem. Phys. Lett. 2004, 395, 190-195.
- (28) Yousaf, K. E.; Peterson, K. A. J. Chem. Phys. 2008, 129, 184108.
- (29) Yousaf, K. E.; Peterson, K. A. Chem. Phys. Lett. 2009, 476, 303–307.
- (30) Whitten, J. L. J. Chem. Phys. 1973, 58, 4496-4501.
- (31) Vahtras, O.; Almlöf, J. E.; Feyereisen, M. W. Chem. Phys. Lett. 1993, 213, 514-518.
- (32) Kendall, R. A.; Frúchtl, H. A. Theor. Chem. Acc. 1997, 97, 158-
- (33) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. Chem. Phys. Lett. **1993**, 208, 359–363.
- (34) Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. Chem. Phys. Lett. 1998, 294, 143-152.
- (35) Manby, F. R. J. Chem. Phys. 2003, 119, 4607-4613.
- (36) Ten-No, S.; Manby, F. R. J. Chem. Phys. 2003, 119, 5358-5363.
- (37) Aquilante, F.; Lindh, R.; Pedersen, T. B. J. Chem. Phys. 2007, 127, 114107.
- (38) Pedersen, T. B.; Aquilante, F.; Lindh, R. Theor. Chem. Acc. 2009, 124, 1-10.
- (39) Aquilante, F.; Boman, L.; Boström, J.; Koch, H.; Lindh, R.; Sánchez de Merás, A.; Pedersen, T. B. In *Linear-Scaling Techniques in Computational Chemistry and Physics*; Zalešny, R., Papadopoulos, M. G., Mezey, P. G., Leszczynski, J., Eds.; Springer: Berlin, 2011; Chapter 13, pp 301–343.
- (40) Weigend, F.; Köhn, A.; Hättig, C. J. Chem. Phys. 2002, 116, 3175-3183.
- (41) Hättig, C. Phys. Chem. Chem. Phys. 2005, 7, 59-66.
- (42) Hill, J. G.; Platts, J. A. J. Chem. Phys. 2008, 128, 044104.
- (43) Hill, J. G.; Platts, J. A. J. Chem. Theory Comput. 2009, 5, 500-505.
- (44) Hill, J. G. J. Chem. Phys. 2011, 135, 044105.
- (45) Hättig, C.; Schmitz, G.; Koßmann, J. Phys. Chem. Chem. Phys. 2012, 14, 6549–6555.
- (46) Weigend, F.; Ahlrichs, R. Phys. Chem. Chem. Phys. 2005, 7, 3297-3305.
- (47) Weigend, F. Phys. Chem. Chem. Phys. 2006, 8, 1057-1065.
- (48) Hellweg, A.; Hättig, C.; Höfener, S.; Klopper, W. Theor. Chem. Acc. 2007, 117, 587–597.
- (49) Hellweg, A.; Rappoport, D. Phys. Chem. Chem. Phys. 2015, 17, 1010-1017.
- (50) TURBOMOLE, V6.5; TURBOMOLE GmbH: Karlsruhe, Germany, 2013. http://www.turbomole.com.
- (51) Hättig, C.; Weigend, F. J. Chem. Phys. 2000, 113, 5154-5161.
- (52) Hättig, C. J. Chem. Phys. 2003, 118, 7751-7761.

- (53) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. WIREs Comput. Mol. Sci. **2012**, 2, 242–253.
- (54) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Shamasundar, K. R.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Köppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; O'Neill, D. P.; Palmieri, P.; Peng, D.; Pflüger, K.; Pitzer, R.; Reiher, M.; Shiozaki, T.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M. MOLPRO, version 2012.1, a package of ab initio programs. https://www.molpro.net.
- (55) Fletcher, R. *Practical Methods of Optimization*, 2nd ed.; Wiley: Chichester, 1987; pp 49-57.
- (56) Nelder, J. A.; Mead, R. Comput. J. 1965, 7, 308-313.
- (57) Weigend, F.; Häser, M. Theor. Chem. Acc. 1997, 97, 331–340.
- (58) Tanaka, M.; Katouda, M.; Nagase, S. J. Comput. Chem. **2013**, 34, 2568–2575.
- (59) Peterson, K. A. Private communication.
- (60) Hill, J. G.; Peterson, K. A.; Knizia, G.; Werner, H.-J. *J. Chem. Phys.* **2009**, *131*, 194105.
- (61) Weigend, F. Phys. Chem. Chem. Phys. 2002, 4, 4285-4291.
- (62) Werner, H.-J.; Adler, T. B.; Manby, F. R. J. Chem. Phys. 2007, 126, 164102.
- (63) Knizia, G.; Werner, H.-J. J. Chem. Phys. 2008, 128, 154103.
- (64) Weigend, F. J. Comput. Chem. 2008, 29, 167-175.
- (65) The molecular test set contains the following compounds: B₂H₆, B₃N₃H₆, B₄H₄, Be₂F₄, Be₂H₄, Be₄, BeC₂H₆, BeF₂O₂H₄, BeH₂, BF₃, BH₃CO, BH₃, BH₃NH₃, C₂H₂, C₂H₃N, C₂H₄, C₂H₆, C₄H₄, C₆H₆, CF₄, CH₂O₂, CH₂O, CH₃N, CH₃OH, CH₄, CO₂, CO, F₂, H₂CO₃, H₂H₂O₂, H₂O, HCN, He₂, HF, HNC, HNO₂, HNO₃, HNO, Li₂, Li₂O, Li₄H₄, Li₈, LiBH₄, LiF, LiH, N₂, N₂H₂, N₂H₄, N₄, NF₃, NH₃, NH₄F, Ne, OF₂, O₂, Al₂O₃, Al₂S₃, AlCl₃, AlF₃, AlH₃, AlN, Ar, BeS, Cl₂, ClF₃, ClF, CS₂, H₂SO₄, H₃PO₄, HCl, HCP, HSH, HSSH, Li₄Cl₄, LiCl, LiSLi, Mg₄, MgCl₂, MgF₂, MgF, MgH₂, Na₂O, Na₂S, Na₃N, Na₃P, NaCl, NaF, NaH, P₂, PF₃, PF₅, PH₃, PLi₃, S₂, S₅, SF₂, SF₄, SF₆, SiCl₄, SiF₄, SiH₄, SiO₂, and SiS₂.
- (66) Feller, D. J. Comput. Chem. 1996, 17, 1571-1586.
- (67) Schuchardt, K. L.; Didier, B. T.; Elsethagen, T.; Sun, L.; Gurumoorthi, V.; Chase, J.; Li, J.; Windus, T. L. J. Chem. Inf. Model. **2007**, 47, 1045–1052.