# ON THE PERFORMANCE OF BOND FUNCTIONS AND BASIS SET EXTRAPOLATION TECHNIQUES IN HIGH-ACCURACY CALCULATIONS OF INTERATOMIC POTENTIALS. A HELIUM DIMER STUDY

Małgorzata JEZIORSKA[a1,b], Robert BUKOWSKI[b1], Wojciech CENCEK[b2], Michał JASZUŃSKI[c], Bogumił JEZIORSKI[a2,b,]* and Krzysztof SZALEWICZ[b3]

[a] *Department of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland; e-mail:* [1] *mjezior@chem.uw.edu.pl,* [2] *jeziorsk@chem.uw.edu.pl*

[b] *Department of Physics and Astronomy, University of Delaware, Newark, Delaware 19716, U.S.A.; e-mail:* [1] *bukowski@udel.edu,* [2] *cencek@physics.udel.edu,* [3] *szalewic@udel.edu*

[c] *Institute of Organic Chemistry, Polish Academy of Sciences, Kasprzaka 44, 01 224 Warszawa, Poland; e-mail: michal@ewa.ichf.edu.pl*

*Dedicated to Professors Petr Čársky, Ivan Hubač and Miroslav Urban on the occasion of their 60th birthdays.*

Helium dimer interaction energies, $E_{int}$, obtained recently using the Gaussian geminal implementation of the coupled cluster doubles (CCD) and singles and doubles (CCSD) theory, were employed to evaluate the performance of conventional orbital calculations applying the correlation-consistent polarized valence $X$-tuple zeta (cc-pV$X$Z) bases, with $X$ ranging from 4 to 7, and very large sets of bond functions. We found that while the bond functions improve dramatically the convergence of the doubles and triples contribution to the interaction energy, these functions are inefficient or even counterproductive in predicting the effect of the single excitations and the small contribution beyond the CCSD(T) (CCSD model with noniterative account of triple excitations) level of electronic structure theory. We also found that bond functions are very effective in extrapolation techniques. Using simple two-point extrapolations based on the single-power laws $X^{-2}$ and $X^{-3}$ for the basis set truncation error, the Gaussian geminal CCSD result for $E_{int}$, equal to –9.150 ± 0.001 K at the equilibrium interatomic distance of $R$ = 5.6 bohr, could be reproduced with an error of 2–3 mK. Linear extrapolation of the functional dependence of the CCSD energy on the value of the second-order Møller–Plesset energy and the use of the known accurate value of the latter leads to an even smaller error. Using these extrapolation techniques with basis sets up to doubly augmented septuple-zeta quality and containing large sets of bond functions, we estimated the contribution of triple excitations within the CCSD(T) model to be –1.535 ± 0.002 K, with the error bars reflecting the spread of extrapolated results. The contribution beyond the CCSD(T) model, estimated from full configuration interaction (FCI) calculations with up to 255 orbitals, amounts to –0.323 ± 0.005 K. Combining the Gaussian geminal value of the CCSD energy with the orbital estimations of the CCSD(T) and FCI contributions, we found that $E_{int}$ = –11.008 ± 0.008 K. This value is consistent with recent high-

The selection of an optimal composition of orbitals and their exponents is important in all electronic structure calculations. In the case of correlation energy calculations for isolated molecules, this issue has been the subject of extensive investigations in the past and several libraries of optimized basis sets are available. An example are the correlation-consistent polarized valence *X*-tuple zeta (cc-pV*XZ*) bases of Dunning and collaborators[1–4]. These types of basis sets are also often used in calculations of interaction energies. However, it has been known for some time that basis sets optimized for the correlation energy (we will refer to them as CR bases) are not optimal for calculations of intermolecular interaction energies (see, *e.g.*, ref.[5]). CR-type basis sets are also known to perform poorly in calculations of molecular properties such as multipole moments and polarizabilities. The former issue is a consequence of the latter since interaction energies are closely related to static and dynamic electric properties of monomers. In fact, basis sets optimized for polarizability[6] or both polarizability and multipole moments[7] perform well in interaction energy calculations. The property-optimized basis sets differ from CR-type bases by the range of orbital exponents being shifted to smaller values, *i.e.*, the former bases are more diffuse than the latter. A simple method to make CR bases to acquire more diffuse character is to add functions with exponents extrapolated from the even-tempered progression of the smallest CR-optimized exponents. In this way the so-called "augmented", aug-cc-pV*XZ*, basis sets were created. More sets of diffuse functions can also be added, leading to doubly (d-aug-cc-pV*XZ*) or triply (t-aug-cc-pV*XZ*) augmented cc-pV*XZ* bases. The aug-cc-pV*XZ* bases have often been used in calculations of interaction energies in recent years. One should note, however, that bases such as aug-cc-pV*XZ* require more functions to reach a given accuracy of interaction energies compared to property-optimized bases or to the dispersion-optimized bases discussed below.

Basis sets can be directly optimized for the purpose of calculations of intermolecular interaction energies. This is due to the fact that the dispersion energy – the interaction energy component which is most difficult to

converge – possesses variational character[8]. Minimization of the leading dispersion energy term produces basis sets with exponents intermediate between those appearing in CR-type and in property-optimized bases[5]. We will denote such bases by the acronym "DS".

Even with the use of DS-type basis sets, the interaction energy is difficult to converge, and the bulk of the error remains in the dispersion component. This problem can be alleviated by the introduction of functions located on the intermolecular bond. Such functions were used already in the 1970s in calculations of dispersion energies using Gaussian-type geminals (GTG)[9,10] since these bases are usually applied with "floating" centers. The first calculations utilizing bond functions within orbital basis sets were probably those of Burton[11] and of Gutowski *et al.*[12] Bond functions were later popularized by the work of Tao *et al.*[13–16] who applied them to a number of dimers and developed working criteria for proper inclusion of such functions. Burcl *et al.*[17] and Williams *et al.*[18] have shown that it is the dispersion energy that requires basis functions located on the intermolecular bond since this is the region where important, dispersion-type electron correlation of electronic motion takes place. DS-type basis sets extended by bond functions appear to be optimal for the saturation of the dispersion part of the correlation energy[18]. Such basis sets are therefore optimal for interaction energy calculations in small (double- or triple-zeta quality) basis sets. At some point, however, the errors resulting from an inadequate description of the intramonomer correlation effects may become larger than the errors of the dispersion energy. CR-type exponents must then be included in the basis set to correctly describe the intramonomer correlation. The importance of CR-type and bond functions for the calculations of dispersion and intramolecular correlation components of the interaction energy was thoroughly discussed by Torheyden and Jansen[19].

With the development of computer capabilities, bases of quadruple-zeta quality and larger can be used in practical calculations of interaction energies. Although it has been known that as the basis set increases, the composition of the basis has to include more CR-type exponents[18], precise guidelines for this procedure have not been established. It is not clear whether the number of bond functions should increase with the size of the basis set or perhaps these functions become less important when large nuclei-centered bases are used. Little is known about the importance of bond functions at different levels of inclusion of the electron correlation. Furthermore, to our knowledge, there have been no attempts to use bases with bond functions in extrapolations of interaction energies to the complete basis set limit. The present paper is aimed at filling these gaps by investigat-

ing optimal strategies for accurately computing weak interaction energies using correlation-consistent bases of quadruple-zeta and higher quality. To gauge the effectiveness of various basis sets and extrapolations strategies, we shall use the helium dimer as a model. Interaction energies for this system have recently been obtained[20] using the Gaussian geminal implementation of CCSD theory[21] and these energies, believed to be accurate up to 1 mK, can be used as our reference data.

Although the present paper is devoted to methods of constructing optimal basis sets for interaction energy calculations, a brief explanation why the helium dimer requires calculations of such extreme accuracy is in order. Helium is a system of great interest in physics. Its superfluid phase has been the subject of extensive investigations for a long time. The gaseous helium is a benchmark medium in measurements of thermophysical properties of gases[22,23]. Superfluid helium nanodroplets are of increasing importance in high-resolution spectroscopy[24,25]. The experimental observations involving helium are predicted and analyzed by various theoretical methods, most of them utilizing the helium dimer interaction potential. Some of these methods require a highly accurate potential. For example, accuracy better than 0.01 K at the minimum is needed to create a new pressure standard[26,27]. Initially, most accurate helium potentials were constructed *via* empirical fits to experimental data. In 1995 van Mourik and van Lenthe[28] computed interaction energies that were employed to construct a potential which turned out to be more accurate than the empirical ones[22]. These authors used the supermolecular full configuration interaction (FCI) method and orbital basis sets optimized on interatomic interaction energy. The computed interaction energy at the internuclear distance $R$ = 5.6 bohr was –10.947 ± 0.01 K. Soon afterwards, papers by Klopper and Noga[29] and by Bukowski *et al.*[30] presented supermolecular calculations of largest components of the interaction energy for a few values of $R$ using explicitly correlated basis sets and adding the remaining components taken from orbital calculations. Both investigations arrived at the interaction energy at $R$ = 5.6 bohr equal to –11.00 K, significantly lower than the van Mourik and van Lenthe[28] value. A similar hybrid basis set approach was used by Williams *et al.*[31] within symmetry-adapted perturbation theory[32] (SAPT). These authors (see also Korona *et al.*[33] for a more detailed description of this potential) obtained the depth at the minimum of the potential equal to 11.06 ± 0.03 K. Since then, several other *ab initio* calculations have been published, most of them at high levels of theory, utilizing very large orbital or explicitly correlated basis sets, and sometimes extrapolating in basis set size. These calculations resulted in potentials with depths ranging from –10.95 K to –11.10 K [34–42],

sometimes with incompatible estimates of error bounds. Clearly, this situation is not satisfactory and one of the goals of the present work was to converge the interaction energy contributions beyond the CCSD level to an accuracy better than 0.01 K. Estimates of error bounds of *ab initio* calculations are almost always to some extent arbitrary. However, in the present work we attempted to converge the calculations so well that such estimates could be made in a very conservative way.

**THEORY**

In the research reported in this communication, the helium dimer interaction energy $E_{int}$ was computed using the conventional supermolecular approach[43]. We represented $E_{int}$ as the following sum

$$E_{int} = E_{CCSD} + E_T + \delta E_{FCI} , \tag{1}$$

where $E_{CCSD}$ is the interaction energy defined in the CCSD model[44], $E_T$ is the contribution of three-electron correlations (triple excitations)

$$E_T = E_{CCSD(T)} - E_{CCSD} , \tag{2}$$

with $E_{CCSD(T)}$ being the interaction energy of the CCSD(T) method (*i.e.*, the CCSD model with approximate, noniterative account of triple excitations) defined in ref.[45], and $\delta E_{FCI}$ is the remainder, $\delta E_{FCI} = E_{int} - E_{CCSD(T)}$, accounting for the quadruple excitations and for the three-electron correlation effects neglected by the CCSD(T) approach. Actually, the latter effects dominate $\delta E_{FCI}$, as discovered by Burda and co-workers[46]. The only practical way of computing $\delta E_{FCI}$ is to perform FCI and CCSD(T) calculations of $E_{int}$ in the same basis set and subtract the results. It may be noted here that the idea of using the converged CCSD amplitudes to compute the perturbation-theory-type correction for triple excitations, which underlies the CCSD(T) model, was proposed and applied for the first time by Urban and co-workers[47].

The energy partition of Eq. (*1*) is useful from a practical point of view since the identified interaction energy components $E_{CCSD}$, $E_T$, and $\delta E_{FCI}$ have very different computational requirements and can be optimally obtained using different methods and/or basis sets. For example, the largest contribution, $E_{CCSD}$, is the most basis set sensitive (as far as the absolute

value of the basis set error is concerned) and must be computed with bases which are impractical in calculating $\delta E_{FCI}$. Since the CCSD contribution depends only on one- and two-electron functions, it can be, unlike $E_T$ or $\delta E_{FCI}$, computed very accurately using Gaussian geminal methods[21].

From the point of view of Gaussian geminal calculations, it is very convenient to further partition $E_{CCSD}$ as

$$E_{CCSD} = E_{CCD} + E_S , \tag{3}$$

where $E_{CCD}$ is the interaction energy predicted by the coupled-pair many-electron theory (CPMET) of Čížek[48], referred nowadays to as the coupled cluster doubles (CCD) model, and $E_S$ is the single excitation contribution to the CCSD energy, defined essentially by Eq. (*3*). For a particularly comprehensible, first-quantized presentation of CCD and its properties, we refer the reader to the book by Čársky and Urban[49]. Note that the repulsive, self-consistent field (SCF) part of the interaction energy, denoted by $E_{SCF}$, is included in the CCD interaction energy. In the conventional, orbital implementation of the coupled cluster theory, the computational requirements of the CCD and CCSD methods are comparable. The explicitly correlated geminal CCSD theory[21] is, however, much more demanding computationally than the Gaussian geminal version of the CCD method[50]. Therefore, in very accurate Gaussian geminal applications it is convenient to compute $E_{CCD}$ and $E_S$ using different geminal basis sets, so the partitioning of Eq. (*3*) becomes useful.

Since the many-body Rayleigh–Schrödinger perturbation theory[49] based on the Møller–Plesset partition of the Hamiltonian is so often applied in practice, we also considered it in the present work. The interaction energy computed (in the supermolecular way) through the second order of this theory will be denoted by $E_{MP2}$ and referred to as the MP2 interaction energy. The pure correlation part of $E_{MP2}$ will be denoted by $E_{MP2}^{cr}$. Thus, $E_{MP2} = E_{SCF} + E_{MP2}^{cr}$.

It is well known[51–53] that, as a result of the interelectronic cusp singularity of the exact wave function, the orbital basis set expansions of the components of Eqs (*1*) and (*3*) converge frustratingly slowly and to improve accuracy, some kind of extrapolation technique is necessary. From this point of view, the family of correlation-consistent cc-pV*X*Z basis sets developed by Dunning and co-workers[1–4] appear to be especially useful since the energy or other properties computed using these bases can be expected to

converge in a systematic way to appropriate complete basis set limits when the cardinal number $X$ grows to infinity. There is an extensive literature on this subject and many different extrapolation schemes have been proposed and applied, also to study van der Waals interactions. This literature is reviewed and critically discussed in refs[54,55]. For a more recent work, see refs[41,56–61]. Three extrapolation formulas were investigated in the present work. The simplest of them consists of fitting the expression

$$E(X) = E^{(n)}(\infty) + A^{(n)} X^{-n} \tag{4}$$

to two consecutive interaction energies obtained using basis sets with cardinal number $X - 1$ and $X$. The number

$$E^{(n)}(\infty) = E(X-1) + \frac{E(X) - E(X-1)}{1 - (1 - 1/X)^n} \tag{5}$$

obtained by solving the system of linear equations for $A^{(n)}$ and $E^{(n)}(\infty)$, represents an estimate of the basis set limit for the energy $E$. The choice $n = 3$ can be rationalized by the rigorous results of refs[51–53] and was advocated by Helgaker and co-workers[62,63].

   We also considered a three-point extrapolation formula based on the expression

$$E(X) = E^{(nn')}(\infty) + A^{(n)} X^{-n} + A^{(n')} X^{-n'} \tag{6}$$

fitted to three consecutive interaction energies obtained using basis sets with cardinal numbers $X - 2$, $X - 1$, and $X$. The corresponding estimate $E^{(nn')}(\infty)$ of the complete basis set limit is obtained by solving the system of three linear equations for $A^{(n)}$, $A^{(n')}$, and $E^{(nn')}(\infty)$. When $n = 3$ and $n' = 5$, Eq. (6) can be justified by the fact that the MP2 or CCD energies are sums of singlet and triplet contributions which are expected to converge as $X^{-3}$ and $X^{-5}$, respectively[41,60].

   Since the MP2 energy (or energy of other theories based on the concept of pair functions) can be computed with high accuracy using Gaussian geminals[21,30,64] or the MP2-R12 method of Kutzelnigg and Klopper[65,66], one may try to use the accurate value of this energy in an extrapolation procedure employing sequences of cc-pV$X$Z bases. The method we applied can be

described as follows. Let us assume that a hard-to-compute approximate energy denoted by $E_H$ (for instance $E_H = E_T$ or $E_H = \delta E_{FCI}$) is an analytic function $f$ of an approximate energy $E$ (for instance $E = E_{MP2}$) which is much easier to accurately compute than $E_H$. For a given (potentially infinite) sequence of basis sets $\mathcal{B}_X$ (for instance $\mathcal{B}_X = $ d-aug-cc-pV$X$Z), the function $f(z)$ is defined by

$$f(E(\mathcal{B}_X)) = E_H(\mathcal{B}_X), \qquad X = 2, 3, 4, \dots \, . \tag{7}$$

We assume that the exact value of $E$, given by $E(\infty) \equiv \lim_{X \to \infty} E(\mathcal{B}_X)$, is known accurately. For a specified infinite sequence $\mathcal{B}_X$, the analytic function $f(z)$ is uniquely defined in some neighborhood of $z = E(\infty)$ since the sequence of numbers $E(\mathcal{B}_X)$, $X = 2, 3, 4, \dots$, has an accumulation point (equal to $E(\infty)$). If the function $f(z)$ were known, the exact value of $E_H$, denoted by $E_H(\infty)$, could be found from $E_H(\infty) = f(E(\infty))$. If we make the simplest possible approximation and assume that the function $f(z)$ is linear, $f(z) = az + b$, when $z$ is close to $E(\infty)$, and use the bases with cardinal numbers $X - 1$ and $X$ to fix the coefficients $a$ and $b$, we find

$$E_H(\infty) = E_H(X) + \frac{E(\infty) - E(X)}{E(X-1) - E(X)}[E_H(X-1) - E_H(X)] \, . \tag{8}$$

We will refer to this technique as $E_H$ *vs* $E$ extrapolation. Since very accurate Gaussian geminal values of the helium dimer interaction energy at the MP2, CCD, and CCSD level of theory are known[20], we can set $E = E_{MP2}$ and $E_H = E_{CCD}$ or $E = E_{MP2}$ and $E_H = E_{CCSD}$ and test the accuracy of Eq. (*8*) before applying it to unknown quantities like $E_T$ or $\delta E_{FCI}$. All three extrapolation schemes outlined above were applied both to the original (d)-aug-cc-pV$X$Z bases and to sequences obtained by including several sets of bond functions. Unless otherwise noted, the extrapolated MP2, CCD, and CCSD interaction energies include the Hartree–Fock part $E_{SCF}$.

### COMPUTATIONAL

The computations of $E_{CCD}$, $E_{CCSD}$, and $E_T$, as well as the FCI calculations, were performed for sequences of singly and doubly augmented cc-pV$X$Z basis sets for helium[1,3,4,39] with $X = Q, 5, 6,$ and $7$. For brevity, these basis sets will be denoted as a$X$Z for singly and d$X$Z for doubly augmented bases. We

also used sequences which were additionally augmented by bond functions located precisely at the middle between the interacting atoms. These sequences will be denoted by a$XZ$+b$m$ or d$XZ$+b$m$, where $m$ is the number of bond functions. We used four sets of bond functions: the standard set 3s3p2d1f1g recommended by Tao and Pan[14] and denoted here as b38, the 6s6p6s3f1g1h and 6s6p6s3f3g3h sets developed by Partridge and Bauschlicher[67], denoted by us as b95 and b135, respectively, and the set 3s3p2d2f1g1h1i, denoted as b69, obtained by adding one set of i functions with the exponent 0.5 to the set of 56 bond functions used to construct the interaction optimized Mc191 basis set of ref.[33] We also used the dimer-centered version of the basis Mc191 of ref.[33] to perform FCI calculations. It contains the same 56 bond functions as the Mc191 basis and the set of 9s7p6d4f1g functions from the Mc191 basis which are now centered on both helium atoms. This dimer-centered basis includes 250 functions and will be denoted here as DC250.

The MP2 and CCSD(T) energies were computed using the MOLPRO package[68]. The FCI calculations were performed with the LUCIA [69] program. Some of the CCSD(T) energies were obtained employing also the DALTON [70] program. The interaction energies were always calculated using the Boys and Bernardi counterpoise method[71] and converted to the kelvin units (K) with the conversion factor 1 hartree = 315774.65 K.

### RESULTS AND DISCUSSION

Results of our computations and the follow-up extrapolations are presented in Tables I–VI. We found that, except for the theoretically motivated case $n = 3$, $n' = 5$, the three-point extrapolations based on Eq. (*6*) work much worse than the simpler scheme based on the single-power convergence law. Although occasionally these three-point extrapolations can give very accurate results, such extrapolations can become unstable when $n' = n + 1$ and we cannot recommend their application. We also tried a three-point extrapolation based on the fractional power convergence law[72] $X^{-\alpha}$ with optimized $\alpha$, but we did not obtain satisfactory results. We found that when $n \geq 4$, the single-power law $X^{-n}$ significantly overestimates the convergence rate and leads to inaccurate results. Therefore, we limited ourselves here to the $X^{-2}$, $X^{-3}$, and $X^{-3} + X^{-5}$ extrapolations. We, of course, realize that the $X^{-2}$ convergence law does not have a theoretical foundation, but we have empirically found that it leads to very accurate results when bond functions are included in the basis set. We shall now discuss separately the results obtained for different contributions to the interaction energy. In this dis-

cussion, we will use the notation a$(X - 1, X)$Z+b$m$ for the extrapolations employing the energies obtained with a$(X - 1)$Z+b$m$ and a$X$Z+b$m$ bases. Similar definition applies to the notation d$(X - 1, X)$Z+b$m$. When $m = 0$, the suffix +b0 will be omitted.

TABLE I
Computed and extrapolated helium dimer CCD interaction energies (in K) at $R = 5.6$ bohr. The Gaussian geminal value of the CCD interaction energy is $-8.972 \pm 0.001$ K. The Hartree–Fock part of this energy is 9.220 K

| Basis | Size | $E_{\mathrm{CCD}}$[a] | $X^{-3}$ [b] | $X^{-3} + X^{-5}$ [c] | $E_{\mathrm{CCD}}$ ($E_{\mathrm{MP2}}$)[d] |
|---|---|---|---|---|---|
| aQZ | 92 | −7.674 | | | |
| a5Z | 160 | −8.199 | −8.749 | | −9.013 |
| a6Z | 254 | −8.591 | −9.130 | −9.347 | −9.036 |
| a7Z | 378 | −8.748 | −9.014 | −8.926 | −9.022 |
| a5Z + b69 | 229 | −8.912 | | | |
| a6Z + b69 | 323 | −8.935 | −8.968 | | −8.983 |
| a7Z + b69 | 447 | −8.948 | −8.969 | −8.969 | −8.978 |
| aQZ + b95 | 187 | −8.893 | | | |
| a5Z + b95 | 255 | −8.929 | −8.966 | | −8.980 |
| a6Z + b95 | 349 | −8.946 | −8.971 | −8.974 | −8.979 |
| a7Z + b95 | 473 | −8.955 | −8.969 | −8.968 | −8.977 |
| a5Z + b135 | 295 | −8.935 | | | |
| a6Z + b135 | 389 | −8.950 | −8.970 | | −8.977 |
| a7Z + b135 | 513 | −8.958 | −8.971 | −8.971 | −8.976 |
| d5Z | 210 | −8.710 | | | |
| d6Z | 326 | −8.873 | −9.096 | | −9.000 |
| d7Z | 476 | −8.911 | −8.975 | −8.884 | −8.972 |
| d5Z + b69 | 279 | −8.926 | | | |
| d6Z + b69 | 395 | −8.943 | −8.967 | | −8.979 |
| d7Z + b69 | 545 | −8.953 | −8.970 | −8.972 | −8.977 |

[a] Computed, nonextrapolated values, including the Hartree–Fock component of the interaction energy. [b] Results of the conventional $X^{-3}$ extrapolation, see Eq. (5). [c] Results of the $X^{-3} + X^{-5}$ extrapolation, see Eq. (6). [d] Results of CCD *versus* MP2 extrapolation, see Eq. (8). The assumed accurate (Gaussian geminal) value of the MP2 interaction energy is −6.791 K.

*CCD and MP2 Interaction Energies*

The results presented in Table I show that the bond functions are very effective in reproducing the accurate value of $E_{CCD}$. For instance, the a5Z+b69 basis gives a better result than the d7Z basis containing twice as many functions. With simple $X^{-3}$ extrapolations, all sequences except for a$X$Z give very accurate results when $X = 7$ bases are used. The errors are of the order of only few mK. The extrapolations using bond functions are more effective than those using bases centered only at atomic nuclei. While the a(67)Z extrapolation overshoots by 42 mK, the extrapolations from sequences a$X$Z+b$m$, $m = 69$, 95, and 135 give stable results with errors of 1–3 mK. The d(67)Z extrapolation is accurate to 3 mK but the corresponding d(56)Z and d(45)Z extrapolations are much less accurate. We have verified that extrapolations of pure correlation energies give practically the same errors.

Somewhat disappointingly, the three-point $X^{-3} + X^{-5}$ extrapolation of Eq. (*6*) is not an improvement over the simple $X^{-3}$ extrapolation. In fact, for the sequences without bond functions, the extrapolated results are less accurate than the computed ones. This observation is inconsistent with the results of Klopper[41] who argued that the singlet and triplet pair contributions to the interaction energy (25 and 75% of $E_{CCD}$, respectively, at large $R$) should converge as $X^{-3}$ and $X^{-5}$ to their basis set limits.

In the last column of Table I we give the results of the extrapolations of $E_{CCD}$ as a function of $E_{MP2}$, obtained using Eq. (*8*). In this case the d(67)Z extrapolation gives the best (within 1 mK) and a(56)Z the worst (64 mK off the mark) result (note, however, that the a(56)Z value is more accurate than the results of purely orbital extrapolations employing aQZ, a5Z, and a6Z data as an input). The a(67)Z+b$m$ and d(67)Z+b69 extrapolations overshoot by 4–6 mK, which is a little disappointing since the computed (without extrapolation) values of $E_{CCD}$ are in error by only 15–25 mK. Overall, the CCD *vs* MP2 extrapolations perform little worse than $X^{-3}$ extrapolations, but the closeness of the predictions by these two quite different techniques provides an additional argument for the reliability of either method.

An important observation based on the results of Table I is that the singly-augmented bases with $X \leq 7$ are not able to reproduce the interaction energy of the helium dimer with errors smaller than about 50 mK, even with extrapolations. The double augmentation brings significant improvement[41], however, not as dramatic as that provided by bond functions.

The computed and extrapolated MP2 energies are shown in Table II. We present here the values of the pure correlation energies to eliminate any dependence on the basis set inaccuracy of the SCF contributions to the inter-

action energy. In contrast to other components, extrapolations of $E_{MP2}^{cr}$ led to slightly more accurate results than extrapolations of $E_{MP2}$. One can see that the $X^{-3}$ extrapolations of the MP2 energies are significantly less accurate than those of the CCD or CCSD energy. Perhaps the reason for this behavior may be that for the same basis set the computed MP2 interaction energies are significantly less accurate than the energies of the CCD or CCSD method. Apparently the $r_{12}$ singularity of the MP2 pair functions is

TABLE II
Convergence of the computed and extrapolated MP2 correlation energy contributions $E_{MP2}^{cr} = E_{MP2} - E_{SCF}$ to the interaction energy in the helium dimer (in K) at $R$ = 5.6 bohr. The reference (Gaussian geminal) value of $E_{MP2}^{cr}$ is $-16.011 \pm 0.001$ K

| Basis | Size | $E_{MP2}^{cr}$ $^a$ | $X^{-3}$ $^b$ | $X^{-3} + X^{-5}$ $^c$ | $X^{-2}$ $^b$ |
|---|---|---|---|---|---|
| aQZ | 92 | −14.753 | | | |
| a5Z | 160 | −15.256 | −15.784 | | −16.151 |
| a6Z | 254 | −15.607 | −16.088 | −16.262 | −16.403 |
| a7Z | 378 | −15.752 | −16.000 | −15.933 | −16.156 |
| a5Z + b69 | 229 | −15.935 | | | |
| a6Z + b69 | 323 | −15.961 | −15.998 | | −16.021 |
| a7Z + b69 | 447 | −15.975 | −15.999 | −16.000 | −16.013 |
| aQZ + b95 | 187 | −15.906 | | | |
| a5Z + b95 | 255 | −15.947 | −15.991 | −16.006 | −16.022 |
| a6Z + b95 | 349 | −15.968 | −15.997 | −16.001 | −16.016 |
| a7Z + b95 | 473 | −15.980 | −16.000 | −16.002 | −16.012 |
| a5Z + b135 | 295 | −15.957 | | | |
| a6Z + b135 | 389 | −15.975 | −16.000 | | −16.016 |
| a7Z + b135 | 513 | −15.985 | −16.003 | −16.005 | −16.014 |
| dQZ | 124 | −15.336 | | | |
| d5Z | 210 | −15.707 | −16.097 | −16.106 | −16.367 |
| d6Z | 326 | −15.874 | −16.102 | −16.105 | −16.251 |
| d7Z | 476 | −15.926 | −16.015 | −15.949 | −16.071 |
| d5Z + b69 | 279 | −15.947 | | | |
| d6Z + b69 | 395 | −15.968 | −15.997 | | −16.016 |
| d7Z + b69 | 545 | −15.980 | −16.000 | −16.003 | −16.013 |

$^a$ Computed, nonextrapolated values. $^b$ Results of extrapolations defined by Eq. (*5*). $^c$ Result of the $X^{-3} + X^{-5}$ extrapolation, see Eq. (*6*).

somewhat more serious than that appearing at the level of an infinite-order coupled-pair theory. Inclusion of the $X^{-5}$ term in the extrapolation formula does not help. To our surprise, however, we found that the $X^{-2}$ extrapolation works remarkably well for the MP2 energy, but only when the bond functions are included. We verified that this holds also when the interac-

TABLE III
Computed and extrapolated single excitation contributions $E_S = E_{CCSD} - E_{CCD}$ to the interaction energy of the helium dimer (in K) at $R = 5.6$ bohr. The Gaussian geminal value of $E_S$ is $-0.178 \pm 0.001$ K

| Basis | Size | $E_S{}^a$ | $X^{-3\ b}$ | $E_S\ (E_{CCD})^c$ |
|---|---|---|---|---|
| aQZ | 92 | −0.192 | | |
| a5Z | 160 | −0.179 | −0.166 | −0.161 |
| a6Z | 254 | −0.179 | −0.178 | −0.178 |
| a7Z | 378 | −0.179 | −0.180 | −0.179 |
| a5Z + b69 | 229 | −0.198 | | |
| a6Z + b69 | 323 | −0.188 | −0.175 | −0.174 |
| a7Z + b69 | 447 | −0.184 | −0.177 | −0.175 |
| aQZ + b95 | 187 | −0.225 | | |
| a5Z + b95 | 255 | −0.199 | −0.171 | −0.166 |
| a6Z + b95 | 349 | −0.189 | −0.175 | −0.174 |
| a7Z + b95 | 473 | −0.184 | −0.177 | −0.175 |
| a5Z + b135 | 295 | −0.199 | | |
| a6Z + b135 | 389 | −0.189 | −0.175 | −0.174 |
| a7Z + b135 | 513 | −0.184 | −0.177 | −0.176 |
| dQZ | 124 | −0.214 | | |
| d5Z | 210 | −0.194 | −0.173 | −0.180 |
| d6Z | 326 | −0.187 | −0.177 | −0.182 |
| d7Z | 476 | −0.183 | −0.177 | −0.177 |
| d5Z + b69 | 279 | −0.198 | | |
| d6Z + b69 | 395 | −0.189 | −0.175 | −0.172 |
| d7Z + b69 | 545 | −0.184 | −0.177 | −0.176 |

[a] Computed, nonextrapolated values. [b] Result of the conventional $X^{-3}$ extrapolation, see Eq. (5). [c] Results of $E_S$ *versus* CCD extrapolation, see Eq. (8). The assumed accurate (Gaussian geminal) value of the CCD energy is −8.972 K.

tion energies rather than the pure correlation energies are extrapolated. The $X^{-2}$ extrapolation was also applied to $E_{CCD}$ (and $E_S$), but was found to be less effective (especially for $E_S$) than the $X^{-3}$ extrapolation.

### Singles Contribution

Since for the same Gaussian geminal basis set a computation of the CCSD energy is much (an order of magnitude) more expensive than the computation of the CCD contributions, it is of interest to see if the relatively small $E_S$ contribution can be accurately computed using pure orbital bases. The results gathered in Table III show that this might be possible but only if an extrapolation is performed. The computed results are clearly not converged, even with bases containing more than 300 functions (except, fortuitously, for the a$XZ$ sequence). The basis set error ranges from 5 to 11 mK (3–6%) with respect to the Gaussian geminal result. It is interesting that the addition of bond functions makes the convergence only worse.

Despite the different computed results, all basis sets sequences except for a$XZ$ give –177 mK as a result of the $X^{-3}$ extrapolation from $X$ = 6 and $X$ = 7 results, in very good agreement with the Gaussian geminal result of –178(1) mK. The a(56)Z and a(67)Z extrapolations give –178 and –179.5 mK, respectively, also in good agreement with the geminal result.

Since the CCSD theory and, consequently, the singles contribution are in a closer relation to the CCD than to the MP2 theory, we decided to perform the extrapolation of $E_S$ as a function of $E_{CCD}$ rather than as a function of $E_{MP2}$. In this case, the sequences a$XZ$ and d$XZ$ give the most accurate results (within 1 mK), whereas the sequences with bond functions underestimate the absolute value of the singles effect by 2–3 mK. We also found that the extrapolations of $E_S$ as a function of $E_{MP2}$ based on the a$XZ$ and d$XZ$ sequences give the same results as the extrapolations of $E_S$ *versus* the CCD energy. When the bond functions are used, the latter extrapolations are slightly more accurate.

### CCSD Interaction Energy

Similarly to the calculations of the CCD energy, the inclusion of bond functions dramatically improves the convergence of both singly and doubly augmented sequences. This is clearly seen in Table IV and in Fig. 1, where, for comparison, the complete basis set limit, obtained using Gaussian geminals, is drawn as a horizontal line at $E_{CCSD}$ = –9.150 K. Since the $E_S$ contributions are too negative, the errors of the computed CCSD interac-

tion energies are somewhat smaller than the corresponding errors of CCD energies.

In the case of the $X^{-3}$ extrapolation, the error of CCSD extrapolated values equals to the sum of the errors of $E_{CCD}$ and $E_S$ extrapolations, so the fourth column in Table IV does not really carry a new information. It may be noticed, however, that although the d(67)Z extrapolation is slightly more accurate than the extrapolations employing bases with bond functions, the d(56)Z extrapolation gives a much worse result than the corresponding extrapolations involving bond functions.

When the three-point $X^{-3}$ + $X^{-5}$ extrapolation is used, the bases containing bond functions are effective (similarly to the CCD case), while the d$X$Z bases give poor results. The $X^{-2}$ extrapolation of the CCSD interaction energies gives very accurate results when the bond functions are included. This holds also when the pure correlation rather than the total energies are extrapolated.

The last column of Table IV shows that when the CCSD energy is extrapolated as a function of the MP2 energy, all bases except a$X$Z give excellent results (the errors are within 1 mK). It is interesting that such extrapolations of the CCSD energy are more accurate than the separate extrapola-
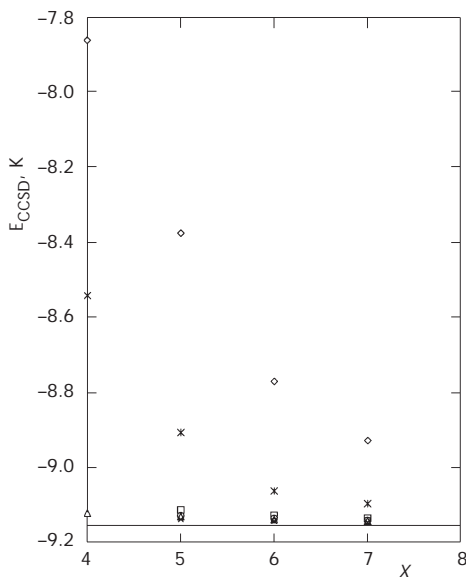


FIG. 1

Convergence of the computed CCSD energies to the Gaussian geminal result (the thick horizontal line): ◇ a$X$Z, □ a$X$Z+b69, △ a$X$Z+b95, × a$X$Z+b135, ✳ d$X$Z, ○ d$X$Z+b69, ---- GTG

tions of the CCD energy and of the singles contributions (there is no additivity of errors in this case). The remarkable linearity of the functional dependence of $E_{CCSD}$ on $E_{MP2}$ is illustrated in Fig. 2. Actually, if we did not know the accurate value of the MP2 energy, it could be read from this figure with quite a good accuracy (together with a very accurate value of

TABLE IV
Computed and extrapolated helium dimer CCSD interaction energies (in K) at $R$ = 5.6 bohr. The reference Gaussian geminal value of the CCSD interaction energy is –9.150 ± 0.001 K. The Hartree–Fock part of this energy is 9.220 K

| Basis | Size | $E_{CCSD}{}^a$ | $X^{-3}{}^b$ | $X^{-3} + X^{-5}{}^c$ | $X^{-3}{}^b$ | $E_{CCSD} (E_{MP2})^d$ |
|---|---|---|---|---|---|---|
| aQZ | 92 | –7.866 | | | | |
| a5Z | 160 | –8.378 | –8.915 | | –9.288 | –9.173 |
| a6Z | 254 | –8.770 | –9.308 | –9.532 | –9.660 | –9.213 |
| a7Z | 378 | –8.927 | –9.194 | –9.107 | –9.361 | –9.201 |
| a5Z + b69 | 229 | –9.109 | | | | |
| a6Z + b69 | 323 | –9.124 | –9.144 | | –9.157 | –9.153 |
| a7Z + b69 | 447 | –9.132 | –9.145 | –9.146 | –9.154 | –9.152 |
| aQZ + b95 | 187 | –9.118 | | | | |
| a5Z + b95 | 255 | –9.127 | –9.137 | –9.153 | –9.143 | –9.140 |
| a6Z + b95 | 349 | –9.135 | –9.146 | –9.151 | –9.153 | –9.149 |
| a7Z + b95 | 473 | –9.139 | –9.146 | –9.146 | –9.150 | –9.150 |
| a5Z + b135 | 295 | –9.134 | | | | |
| a6Z + b135 | 389 | –9.139 | –9.145 | | –9.150 | –9.147 |
| a7Z + b135 | 513 | –9.142 | –9.147 | –9.149 | –9.151 | –9.150 |
| d5Z | 210 | –8.904 | | | | |
| d6Z | 326 | –9.059 | –9.272 | –9.266 | –9.412 | –9.181 |
| d7Z | 476 | –9.094 | –9.152 | –9.061 | –9.189 | –9.149 |
| d5Z + b69 | 279 | –9.125 | | | | |
| d6Z + b69 | 395 | –9.132 | –9.142 | | –9.149 | –9.147 |
| d7Z + b69 | 545 | –9.138 | –9.147 | –9.150 | –9.152 | –9.151 |

$^a$ Computed, nonextrapolated values, including the Hartree–Fock component of the interaction energy. $^b$ Results of extrapolations defined by Eq. (5). $^c$ Result of the $X^{-3} + X^{-5}$ extrapolation, see Eq. (6). $^d$ Results of CCSD *versus* MP2 extrapolation, see Eq. (8). The assumed accurate (Gaussian geminal) value of the MP2 interaction energy is –6.791 K.

the CCSD energy) by finding the abscissa (and the ordinate) of the point where different straight lines intersect. This is a consequence of the fact that different basis sets $\mathcal{B}_X$ lead generally to different functions $f(z)$ defined by Eq. (7), but all these different functions must intersect at the same point. Figure 2 shows also the advantage of using bases containing bond functions. These bases define the $f(z)$ functions with different slopes, which cross at the accurate values of $E$ and $E_H$. In Fig. 2 we also marked (by a black diamond) the point corresponding to the CCSD-R12 and MP2-R12 interaction energies obtained by Klopper and Noga[29]. It can be seen that the CCSD-R12 value, equal to –9.14 K, is only 10 mK above the limit, but is substantially less accurate than the energies obtained from the pure orbital extrapolations described here. The MP2-R12 value of ref.[29], equal to –6.74 K, is about 50 mK off the complete basis set limit obtained using Gaussian geminals[20]. It may be noted here that the (exponentially) extrapolated values of the MP2 and CCSD interaction energies obtained by van Mourik and Dunning[35] using dQZ, d5Z, and d6Z bases and equal to –6.775 and –9.145 K, respectively, are more accurate than the results of the R12 calculations of ref.[29]
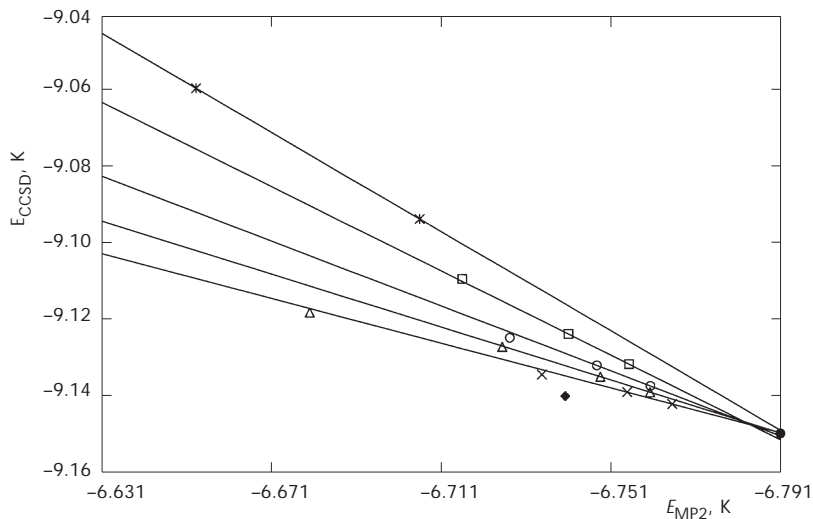


FIG. 2

Extrapolation of the CCSD energy as a function of the MP2 energy, see Eq. (8). The straight lines are drawn through the energies obtained with $X = 6$ and $X = 7$. The result obtained with the d5Z basis is out of range and is not shown but it lies exactly on the straight line determined by the d6Z and d7Z results. The black circle and black diamond correspond to the results obtained in ref.[20] and ref.[29], respectively. □ a$XZ$+b69, △ a$XZ$+b95, × a$XZ$+b135, ✻ d$XZ$, ○ d$XZ$+b69, ● GTG, ◆ CCSD-R12

### Triples Contribution

The computed values of the triple-excitation contribution $E_T$ are shown in the third column of Table V and plotted in Fig. 3. No explicitly correlated calculations at the CCSD(T) level are feasible at the moment and the value of $E_T$ is not known with accuracy of the order of 1 mK. To put the computed values of $E_T$ in some perspective, we show in Fig. 3 two horizontal lines (at –1.533 and –1.537 K) representing the upper and lower bound, respectively, of our estimation of $E_T$ made on the basis of extrapolations described in this subsection.

Table V and Fig. 3 clearly show the importance of bond functions. For instance, the basis aQZ+b95 gives a better value of $E_T$ than the basis a7Z containing more than twice as many functions. It is not known how the $E_T$ contribution should be extrapolated and we cannot compare extrapolated results with an accurate one. We can judge the extrapolations only by their internal consistency.

In Table V and Fig. 4 we show the results of the extrapolation of $E_T$ as a function of $E_{CCSD}$ (which are expected to be more accurate than extrapolations *versus* the MP2 energy). As the $X^{-2}$ decay of the basis set error for $E_T$
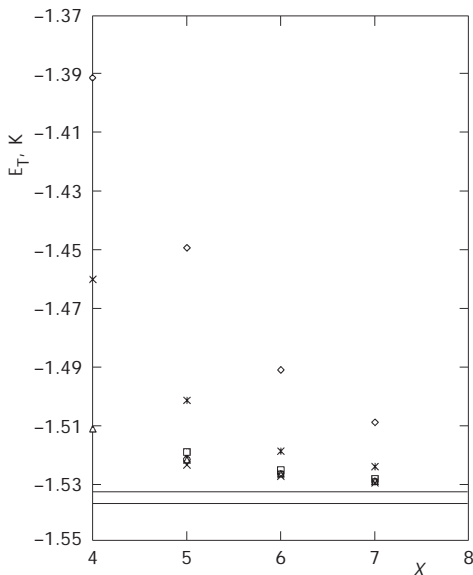


FIG. 3
Convergence of the computed triples contribution, $E_T$. The two horizontal lines denote the upper and lower error bounds of the complete basis set limit estimated from the extrapolations described in the text. ◇ a*XZ*, □ a*XZ*+b69, △ a*XZ*+b95, × a*XZ*+b135, ∗ d*XZ*, ○ d*XZ*+b69

cannot be excluded on theoretical grounds and since the $X^{-2}$ extrapolation gives excellent results for the CCSD and the MP2 energies (when bond functions are used), we present in Table V the $X^{-2}$ extrapolations together with the results of the more conventional $X^{-3}$ extrapolation.

TABLE V
Computed and extrapolated triple-excitation contributions $E_T = E_{CCSD(T)} - E_{CCSD}$ to the helium dimer interaction energies (in K) at $R = 5.6$ bohr

| Basis | Size | $E_T{}^a$ | $X^{-3\ b}$ | $X^{-2\ b}$ | $E_T\ (E_{CCSD})^c$ |
|---|---|---|---|---|---|
| aQZ | 92 | –1.392 | | | |
| a5Z | 160 | –1.449 | –1.510 | –1.553 | –1.537 |
| a6Z | 254 | –1.491 | –1.548 | –1.585 | –1.531 |
| a7Z | 378 | –1.509 | –1.539 | –1.558 | –1.534 |
| a5Z + b69 | 229 | –1.519 | | | |
| a6Z + b69 | 323 | –1.525 | –1.533 | –1.539 | –1.536 |
| a7Z + b69 | 447 | –1.528 | –1.533 | –1.537 | –1.535 |
| aQZ + b95 | 187 | –1.511 | | | |
| a5Z + b95 | 255 | –1.522 | –1.533 | –1.540 | –1.548 |
| a6Z + b95 | 349 | –1.526 | –1.533 | –1.537 | –1.535 |
| a7Z + b95 | 473 | –1.529 | –1.533 | –1.536 | –1.536 |
| a5Z + b135 | 295 | –1.523 | | | |
| a6Z + b135 | 389 | –1.527 | –1.533 | –1.536 | –1.536 |
| a7Z + b135 | 513 | –1.530 | –1.534 | –1.536 | –1.535 |
| dQZ | 124 | –1.460 | | | |
| d5Z | 210 | –1.501 | –1.545 | –1.575 | –1.529 |
| d6Z | 326 | –1.519 | –1.543 | –1.558 | –1.529 |
| d7Z | 476 | –1.524 | –1.533 | –1.539 | –1.533 |
| d5Z + b69 | 279 | –1.522 | | | |
| d6Z + b69 | 395 | –1.526 | –1.533 | –1.537 | –1.538 |
| d7Z + b69 | 545 | –1.529 | –1.534 | –1.537 | –1.535 |
| d7Z + b135 | 611 | –1.530 | | | |

$^a$ Computed, nonextrapolated values. $^b$ Results of extrapolations defined by Eq. (*5*). $^c$ Results of $E_T$ *versus* $E_{CCSD}$ extrapolation, see Eq. (*8*). The assumed accurate (Gaussian geminal) value of $E_{CCSD}$ is –9.150 K.

All $X^{-3}$ extrapolations based on $X = 7$ and $X = 6$ energies, except for a(67)Z, give –1.533 or –1.534 K as a result, in very good agreement with the results of $E_T$ *vs* $E_{CCSD}$ extrapolations which range between –1.533 and –1.536 K when $X = 7$ and $X = 6$ bases are used. As can be expected, the results of $X^{-2}$ extrapolations are somewhat larger in absolute value and give –1.536 or –1.537 K if the bond functions are used. The pure a$X$Z and d$X$Z bases appear to give unreasonably large extrapolated values (similarly as for the MP2 and CCSD energies when extrapolated assuming the $X^{-2}$ convergence law). In view of the above results, we believe that –1.535 ± 0.002 K is at the moment the most reliable estimate of $E_T$. Combining this value with the CCSD energy from GTG calculations[20], we find that the CCSD(T) value of the interaction energy at $R = 5.6$ bohr is –10.685 ± 0.003 K. This value differs somewhat from the value of –10.67 K obtained using the CCSD(T)-R12 approach[29] and even more from a later, unpublished CCSD(T)-R12 result by Klopper[73] equal to –10.659 K (in ref.[41] Klopper quotes the value –10.66 K as the CCSD(T)-R12/B result). The exponential extrapolation of van Mourik and Dunning[35] leads in this case to the value of –10.673 K. Partridge and Bauschlicher[67] report the value of –10.71 K obtained using the a5Z+b38 basis and the interatomic separation of $R = 5.628$ bohr. We re-
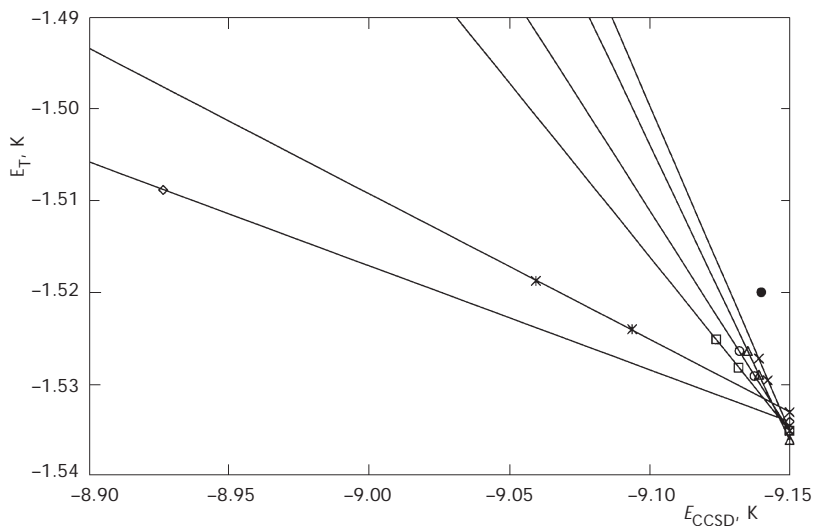


FIG. 4
Extrapolation of the triples contribution, $E_T$, as a function of the CCSD energy, see Eq. (*8*). The straight lines are drawn through the energies obtained with $X = 6$ and $X = 7$. The result obtained with the a6Z basis is out of range. The black circle corresponds to the result obtained in ref.[29] ◇ a$X$Z, □ a$X$Z+b69, △ a$X$Z+b95, × a$X$Z+b135, ∗ d$X$Z, ○ d$X$Z+b69, ● CCSD(T)-R12

computed this value and obtained –10.624 K, in agreement with the value of –10.616 K obtained by us at $R$ = 5.6 bohr using the same a5Z+b38 basis.

### FCI Contribution

The results of the FCI calculations are presented in Table VI. Although the FCI contribution varies relatively little with the size of the basis set, a clear convergence pattern has not been established with the pentuple and sixtuple-zeta basis sets used by us and it is very hard to estimate the value of $\delta E_{FCI}$ with an error smaller than 5 mK. The inclusion of bond functions does not help and appears to be even counterproductive. These functions improve, of course, dramatically the accuracy of the total FCI energy (because the CCD component is improved) but the accuracy of $\delta E_{FCI}$ itself does

TABLE VI
Computed and extrapolated values of the FCI contribution $\delta E_{FCI} = E_{FCI} - E_{CCSD(T)}$ to the helium dimer interaction energy (in K) at $R$ = 5.6 bohr

| Basis | Size | $\delta E_{FCI}$ [a] | $X^{-3}$ [b] | $\delta E_{FCI}\ (E_{CCD})$ [c] |
|---|---|---|---|---|
| aQZ | 92 | –0.321 | | |
| a5Z | 160 | –0.319 | –0.317 | –0.317 |
| a6Z | 254 | –0.321 | –0.324 | –0.323 |
| aTZ + b95 | 141 | –0.337 | | |
| aQZ + b95 | 187 | –0.331 | –0.327 | –0.316 |
| a5Z + b95 | 255 | –0.325 | –0.319 | –0.318 |
| dTZ | 64 | –0.333 | | |
| dQZ | 124 | –0.334 | –0.334 | –0.334 |
| d5Z | 210 | –0.328 | –0.321 | –0.323 |
| dTZ + b38 | 102 | –0.338 | | |
| dQZ + b38 | 162 | –0.333 | –0.329 | –0.319 |
| d5Z + b38 | 248 | –0.326 | –0.319 | –0.318 |
| DC147 [d] | 147 | –0.308 | | |
| DC250 [e] | 250 | –0.326 | | |

[a] Computed, nonextrapolated values. [b] Result of the conventional $X^{-3}$ extrapolation, see Eq. (5). [c] Results of $\delta E_{FCI}$ *versus* CCD extrapolation, see Eq. (8). The assumed accurate (Gaussian geminal) value of the CCD energy is –8.972 K. [d] Interaction optimized Dc147 basis set of ref.[33] [e] Dimer-centered version of the Mc191 basis set of ref.[33], see the text.

not become better. For instance, an addition of a small b38 set of bond functions to the basis dTZ changes the total FCI interaction energy from –9.255 to –10.921 K, *i.e.*, reduces its error from 1.75 to only 0.09 K. Judging from calculations with much larger bases, we believe, however, that the value of $\delta E_{\text{FCI}}$ obtained with basis dTZ+b38 is less accurate than the value obtained with the dTZ basis. Also the extrapolations are not particularly helpful because the convergence trends are not yet well set in for the small

TABLE VII
Recent (since 1995) theoretical values (in K) of the helium–helium interaction energy $E_{\text{int}}$ at $R$ = 5.6 bohr. $N$, $K$, and $L$ denote the sizes of the employed orbital, Gaussian geminal, and four-electron explicitly correlated Gaussian (ECG) bases, respectively

| Reference | $E_{\text{int}}$ | Method |
|---|---|---|
| van Mourik and van Lenthe[28] | –10.947 ± 0.01 | FCI ($N$ = 155) with bond function |
| Klopper and Noga[29] | –11.00 | CCSD(T)-R12 ($N$ = 338) + FCI ($N$ = 159) |
| Bukowski *et al.*[30] | –11.00 | MP3 ($K$ = 490) + CCSD(T) ($N$ = 135) + FCI ($N$ = 71) |
| Williams *et al.*[31] | –11.06 ± 0.03 | SAPT ($K$ = 190) + SAPT ($N$ = 191) + FCI ($N$ = 147) |
| Huiszoon and Caffarel[74] | –11.05 ± 0.38 | perturbational quantum Monte Carlo |
| Komasa and Rychlewski[75] | –10.94 | variational upper bound, $L$ = 1200 |
| Komasa and Rychlewski[34] | –10.947 | variational upper bound, $L$ = 1200 |
| van Mourik and Dunning[35] | –11.00 ± 0.03 | CCSD(T) ($N$ = 326) + CCSDT ($N$ = 124) + FCI ($N$ = 64) |
| Gdanitz[36] | –11.05 ± 0.10 | $r_{12}$-MR-ACPF ($N$ = 272) + extrapolation |
| van de Bovenkamp and van Duijneveldt[38] | –10.99 ± 0.02 | MRCI ($N$ = 301) + CCSD(T)-R12(338) |
| Komasa[37] | –10.978 | variational upper bound, $L$ = 2400 |
| Gdanitz[39] | –11.100 | MR-ACPF ($N$ = 326) + extrapolation |
| Gdanitz[40] | –10.980 ± 0.004 | $r_{12}$-MR-ACPF ($N$ = 420) |
| Klopper[41] | –10.99 ± 0.02 | CCSD(T) ($N$ = 476) + FCI ($N$ = 92) + extrapolation |
| Anderson[42] | –10.98 ± 0.02 | quantum Monte Carlo |
| Present work | –11.008 ± 0.008 | CCSD ($K$ = 800) + CCSD(T) ($N$ = 611) + FCI ($N$ = 255) + extrapolation |

values of $X$ considered by us. In Table VI we present only the $X^{-3}$ extrapolations and the results of extrapolations of $\delta E_{FCI}$ as a function of the CCD energy (the extrapolations *versus* MP2 would give the same results). The values from extrapolations with the largest bases used by us range from –0.318 to –0.325 K. The values obtained with the bond functions are always smaller in absolute value. The computed values obtained with the largest bases in each sequence considered by us range from –0.321 (a6Z basis) up to –0.328 K (d5Z). We (conservatively) estimate that the true FCI contribution is –0.323 ± 0.005 K. The proposed error bounds encompass all extrapolated and all calculated values obtained with the largest bases in each sequence. Using also the computed value of the CCSD contribution (–9.150 ± 0.001 K) and the extrapolated value of $E_T$ equal to –1.535 ± 0.002 K, we find that the total interaction energy at $R$ = 5.6 bohr is –11.008 ± 0.008 K, the error of 8 mK being dominated by the inaccuracy of the FCI contribution. To be on the safe side, we added the errors linearly.

Since $R$ = 5.6 bohr corresponds, to a very good approximation, to the distance of the minimum of the potential, the predicted value of the well depth $D_e$ is 11.008 ± 0.008 K. This value is consistent with recent high-level orbital computations[35,41] but it has substantially tighter error bounds, see the compilation of Table VII for the list of recent values of $D_e$ for the helium dimer. Our result is also consistent with the rigorous variational lower bound (10.978 K) for $D_e$ obtained recently by Komasa[37]. The present prediction differs somewhat from the value of 10.98 ± 0.02 K obtained recently by Anderson[42] from the "exact" quantum Monte Carlo calculations.

The only value published with error bounds below 0.01 K is the result of large scale multiconfigurational averaged coupled pair functional ($r_{12}$-MR-ACPF) calculations of Gdanitz[40]. His result, equal to 10.980 ± 0.004 K, is in a significant disagreement with our value. The value obtained in the present work lies also outside the error bounds of the 1996 SAPT value[31], equal to 11.06 ± 0.03 K. The reason for this discrepancy will be discussed in ref.[76]

### CONCLUSIONS

The results of our investigations can be summarized as follows:

– Even when very large correlation-consistent basis sets are used, the bond functions remain very effective in improving the computed values of the CCSD and CCSD(T) interaction energies.

– Only the double excitation part (CCD) of the CCSD energy is improved by bond functions. The single excitation contribution becomes usually less accurate when bond functions are added.

– Bond functions are not effective or even counterproductive in computations of the difference between the FCI and CCSD(T) energies.

– Bond functions are very effective when extrapolation techniques are used to reach the complete basis set limit of the CCSD and CCSD(T) energies.

– The $X^{-3}$ extrapolation appears to give the most reliable results overall.

– When the bond functions are used, the $X^{-2}$ extrapolation gives exceptionally good results for the MP2 and CCSD interaction energies.

– The $E_H$ *vs E* extrapolations of Eq. (*8*) give predictions of similar accuracy as $X^{-3}$ extrapolations. This consistency demonstrates the reliability of either approach.

– Using the extrapolated values of the contributions beyond the CCSD level and the CCSD energy obtained using Gaussian geminals, we found that the well depth of the helium pair potential amounts to $11.008 \pm 0.008$ K, the error of 0.008 K being dominated by the inaccuracy of the FCI contribution. We believe that this value of $D_e$ is more accurate than the results of published theoretical determinations.

## REFERENCES

1. Dunning T. H.: *J. Chem. Phys*. **1989**, *90*, 1007.
2. Kendall R. A., Dunning T. H., Harrison R. J.: *J. Chem. Phys*. **1992**, *96, 6796*.
3. Woon D. E., Dunning T. H.: *J. Chem. Phys*. **1994**, *100*, 2975.
4. van Mourik T., Wilson A. K., Dunning T. H.: *Mol. Phys*. **1999**, *96*, 529.
5. Szalewicz K., Cole S. J., Kolos W., Bartlett R. J.: *J. Chem. Phys*. **1988**, *89*, 3662.
6. Sadlej A. J.: *Collect. Czech. Chem. Commun*. **1988**, *53*, 1995.
7. van Duijneveldt-van de Rijdt J. G. C. M., van Duijneveldt F. B.: *J. Mol. Struct. (THEOCHEM)* **1982**, *89*, 185.
8. Jeziorski B., van Hemert M.: *Mol. Phys*. **1976**, *31*, 713.
9. Chalasinski G., Jeziorski B., Andzelm J., Szalewicz K.: *Mol. Phys*. **1977**, *33*, 971.
10. Szalewicz K., Jeziorski B.: *Mol. Phys*. **1979**, *38*, 191.
11. Burton P. G.: *J. Chem. Phys*. **1977**, *67*, 4696.
12. Gutowski M., Verbeek J., van Lenthe J. H., Chalasinski G.: *Chem. Phys*. **1987**, *111*, 271.
13. a) Tao F.-M., Pan Y.-K.: *J. Chem. Phys*. **1991**, *95*, 3582; b) Tao F.-M., Pan Y.-K.: *J. Chem. Phys*. **1991**, *95*, 9811; c) Tao F.-M., Pan Y.-K.: *J. Chem. Phys*. **1991**, *96*, 5815.
14. Tao F. M., Pan Y. K.: *J. Chem. Phys*. **1992**, *97, 4989*.

15. a) Tao F.-M.: *J. Chem. Phys.* **1993**, *98*, 3049; b) Tao F.-M., Klemperer W.: *J. Chem. Phys.* **1993**, *99*, 5976; c) Chang H.-C., Tao F.-M., Klemperer W., Healey C., Hutson J. M.: *J. Chem. Phys.* **1993**, *99*, 9337.
16. a) Tao F.-M.: *J. Chem. Phys.* **1994**, *100*, 3645; b) Tao F.-M., Klemperer W.: *J. Chem. Phys.* **1994**, *101*, 1129.
17. Burcl R., Chalasinski G., Bukowski R., Szczesniak M. M.: *J. Chem. Phys.* **1995**, *103*, 1498.
18. Williams H. L., Mas E. M., Szalewicz K., Jeziorski B.: *J. Chem. Phys.* **1995**, *103*, 7374.
19. Torheyden M., Jansen G.: *Theor. Chem. Acc.* **2000**, *104*, 370.
20. Cencek W., Bukowski R., Jaszunski M., Jeziorska M., Jeziorski B., Szalewicz K.: Unpublished results.
21. Bukowski R., Jeziorski B., Szalewicz K.: *J. Chem. Phys.* **1999**, *110*, 4165.
22. Aziz R. A., Janzen A. R., Moldover M. R.: *Phys. Rev. Lett.* **1995**, *74*, 1586.
23. Hurly J. J., Moldover M. R.: *J. Res. Natl. Inst. Stand. Technol.* **2000**, *105*, 667.
24. Toennies P., Vilesov A. F.: *Annu. Rev. Phys. Chem.* **1998**, *49*, 1.
25. Callegari C., Lehmann K. K., Schmied R., Scoles G.: *J. Chem. Phys.* **2001**, *115*, 10090.
26. Moldover M. R.: *J. Res. Natl. Inst. Stand. Technol.* **1998**, *103*, 167.
27. Buckley T. J., Hamelin J., Moldover M. R.: *Rev. Sci. Instrum.* **2000**, *71*, 2914.
28. van Mourik T., van Lenthe J. H.: *J. Chem. Phys.* **1995**, *102*, 7479.
29. Klopper W., Noga J.: *J. Chem. Phys.* **1995**, *103*, 6127.
30. Bukowski R., Jeziorski B., Szalewicz K.: *J. Chem. Phys.* **1996**, *101*, 3306.
31. Williams H. L., Korona T., Bukowski R., Jeziorski B., Szalewicz K.: *Chem. Phys. Lett.* **1996**, *262*, 431.
32. Jeziorski B., Moszynski R., Szalewicz K.: *Chem. Rev. (Washington, D. C.)* **1994**, *94*, 1887.
33. Korona T., Williams H. L., Bukowski R., Jeziorski B., Szalewicz K.: *J. Chem. Phys.* **1997**, *106*, 5109.
34. Komasa J., Rychlewski J.: *Mol. Phys.* **1997**, *91*, 909.
35. van Mourik T., Dunning T. H.: *J. Chem. Phys.* **1999**, *111*, 9246.
36. Gdanitz R. J.: *Mol. Phys.* **1999**, *96*, 1423.
37. Komasa J.: *J. Chem. Phys.* **1999**, *110*, 7909.
38. van de Bovenkamp J., van Duijneveldt F. B.: *J. Chem. Phys.* **1999**, *110*, 11141.
39. Gdanitz R. J.: *J. Chem. Phys.* **2000**, *113*, 5145.
40. Gdanitz R. J.: *Mol. Phys.* **2001**, *99*, 923.
41. Klopper W.: *J. Chem. Phys.* **2001**, *115*, 761.
42. Anderson J. B.: *J. Chem. Phys.* **2001**, *115*, 4546.
43. Chalasinski G., Szczesniak M. M.: *Chem. Rev. (Washington, D. C.)* **1994**, *94*, 1723.
44. Purvis G. G., Bartlett R. J.: *J. Chem. Phys.* **1982**, *76*, 1910.
45. Raghavachari K., Trucks G. W., Pople J. A., Head-Gordon M.: *Chem. Phys. Lett.* **1989**, *157*, 479.
46. Burda J. V., Zahradník R., Hobza P., Urban M.: *Mol. Phys.* **1996**, *89*, 425.
47. Urban M., Noga J., Cole S. J., Bartlett R. J.: *J. Chem. Phys.* **1985**, *83*, 4041.
48. Čížek J.: *J. Chem. Phys.* **1966**, *45*, 4256.
49. Čársky P., Urban M.: *Ab Initio Calculations. Methods and Applications in Chemistry*, p. 93. Springer-Verlag, Berlin 1980.
50. Jeziorski B., Monkhorst H. J., Szalewicz K., Zabolitzky J. G.: *J. Chem. Phys.* **1984**, *81*, 368.
51. Hill R. N.: *J. Chem. Phys.* **1985**, *83*, 1173.
52. Kutzelnigg W.: *Theor. Chim. Acta* **1985**, *68*, 445.
53. Kutzelnigg W., Morgan III J. D.: *J. Chem. Phys.* **1992**, *96*, 4484.

54. Klopper W., Bak K. L., Jørgensen P., Olsen J., Helgaker T.: *J. Phys. B: At., Mol. Opt. Phys.* **1999**, *32*, R103.

55. Helgaker T., Jørgensen P., Olsen J.: *Molecular Electronic-Structure Theory*, p. 322. J. Wiley, Chichester 2000.

56. Lee J. S., Park S. Y.: *J. Chem. Phys.* **2000**, *112*, 10746.

57. Varandas A. J. C.: *J. Chem. Phys.* **2000**, *113*, 8880.

58. Butler P. R., Ellis A. M.: *Mol. Phys.* **2001**, *99*, 525.

59. Lee J. S.: *Chem. Phys. Lett.* **2001**, *339*, 133.

60. Klopper W.: *Mol. Phys.* **2001**, *99*, 481.

61. Neogrády P., Medved M., Černušák I., Urban M.: *Mol. Phys.* **2002**, *100*, 541.

62. Helgaker T., Klopper W., Koch H., Noga J.: *J. Chem. Phys.* **1997**, *106*, 9639.

63. Halkier A., Helgaker T., Jørgensen P., Klopper W., Koch H., Olsen J., Wilson A. K.: *Chem. Phys. Lett.* **1998**, *286*, 243.

64. Szalewicz K., Jeziorski B., Monkhorst H. J., Zabolitzky J. G.: *J. Chem. Phys.* **1983**, *78*, 1420.

65. Kutzelnigg W., Klopper W.: *J. Chem. Phys.* **1991**, *94*, 1985.

66. Klopper W., Kutzelnigg W.: *J. Chem. Phys.* **1991**, *94*, 2020.

67. Partridge H., Bauschlicher C. W.: *Mol. Phys.* **1999**, *96*, 705.

68. Werner H.-J., Knowles P. J.: *MOLPRO, A Package of ab initio Programs*, with contributions from R. D. Amos, A. Bernhardsson, A. Berning, P. Celani, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, T. Korona, R. Lindh, A. W. Lloyd, S. J. McNicholas, F. R. Manby, W. Meyer, M. E. Mura, A. Nicklass, P. Palmieri, R. Pitzer, G. Rauhut, M. Schütz, H. Stoll, A. J. Stone, R. Tarroni and T. Thorsteinsson.

69. Olsen J.: *LUCIA, A Full CI, General Active Space Program*, with contributions from H. Larsen and M. Fulscher.

70. Helgaker T., Jensen H. J. Aa., Jørgensen P., Olsen J., Ruud K., Aagren H., Auer A. A., Bak K. L., Bakken V., Christiansen O., Coriani S., Dahle P., Dalskov E. K., Enevoldsen T., Fernandez B., Hättig C., Hald K., Halkier A., Heiberg H., Hettema H., Jonsson D., Kirpekar S., Kobayashi R., Koch H., Mikkelsen K. V., Norman P., Packer M. J., Pedersen T. B., Ruden T. A., Sanchez A., Saue T., Sauer S. P. A., Schimmelpfennig B., Sylvester-Hvid K. O., Taylor P. R., Vahtras O.: *DALTON, An ab initio Electronic Structure Program*, Release 1.2. 2001. See http://www.kjemi.uio.no/software/dalton/dalton.html.

71. Boys S. F., Bernardi F.: *Mol. Phys.* **1970**, *19*, 553.

72. Martin J. M. L.: *Chem. Phys. Lett.* **1996**, *259*, 669.

73. Klopper W.: Unpublished results cited in ref.[35]

74. Huiszoon C., Caffarel M.: *J. Chem. Phys.* **1996**, *104*, 4621.

75. Komasa J., Rychlewski J.: *Chem. Phys. Lett.* **1996**, *249*, 253.

76. Jeziorska M., Bukowski R., Cencek W., Jaszunski M., Jeziorski B., Szalewicz K.: Unpublished results.