Accuracy of the Boys and Bernardi function counterpoise method

Maciej Gutowski
Department of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland and Theoretical Chemistry Group, University of Utrecht, Padualaan 14, 3584 CH Utrecht, The Netherlands and Department of Chemistry, University of Utah, Salt Lake City, Utah 84112

Jeanne G. C. M. van Duijneveldt-van de Rijdt, Joop H. van Lenthe, and Frans B. van Duijneveldt
Theoretical Chemistry Group, University of Utrecht, Padualaan 14, 3584 CH Utrecht, The Netherlands

(Received 19 June 1992; accepted 4 December 1992)

The performance of the Boys and Bernardi function counterpoise (FCP) method in eliminating the basis set superposition error (BSSE) is studied for He₂, at R = 5.6 a.u., within the supermolecular coupled electron pair approximation (CEPA-1) method. A series of one-electron Gaussian basis sets is designed that allows a systematic approach to the basis set limit value of the interaction energy. Every basis set contains a part suitable to reproduce the atomic correlation energy and a second part optimized for the dispersion interaction in He₂. BSSE-free correlated first-order interaction energies [E⁽¹⁾], calculated using perturbation theory, are reported for each of these sets. Extrapolation to the basis set limit yields a new value of 33.60 ± 0.02 µH for E⁽¹⁾ at R = 5.6 a.u. Extending previous work, the supermolecular CEPA-1 interaction energies for each set are then compared to the total of E⁽¹⁾ and the BSSE-free Møller–Plesset second-order dispersion energy reported previously. While for some basis sets the uncorrected ΔE values deviate up to 43 K from the perturbation estimate, the FCP-corrected results always agree within 0.4 K. A virtuals-only counterpoise procedure is considered as well, but fails badly. The remaining discrepancies in the FCP results are ascribed to a failure of the Møller–Plesset approach to precisely model the dispersion energy at the CEPA level. This problem is removed in a further, more stringent test where supermolecular E⁽¹⁾ results, in which only the intra-atomic correlation (at the CEPA-1 level) is taken into account, are directly compared to the BSSE-free E⁽¹⁾ values. In this test the FCP-corrected supermolecular results agree, for the larger sets, to within 0.001 K with the results expected on the basis of E⁽¹⁾. These findings demonstrate, for the first time, that at least in He₂ the FCP recipe yields interaction energies that correspond precisely (to machine precision) to the basis set and correlation method at hand.

I. INTRODUCTION

The successful application of highly correlated ab initio techniques in the field of van der Waals forces has been seriously hampered by the problem of basis set superposition error (BSSE). For a thorough discussion of BSSE related problems up to 1987 see Ref. 1. BSSE is a consequence of the incompleteness of the one-electron basis used in calculations for the monomers and the dimer. In a finite basis set calculation for the dimer the individual monomer takes advantage of the basis set of the whole dimer rather than of just "its own" basis set centered at this monomer. The related lowering of monomer energies is called BSSE, although strictly speaking the lowering is an improvement rather than an error. The magnitude of BSSE proved to be large enough to prevent a quantitative characterization of the van der Waals potential energy surfaces.

Uncertainty about the way to handle BSSE has induced the development of techniques, that aim to eliminate the BSSE by imposing restrictions on the dimer calculation. Examples are the "chemical Hamiltonian approach" and the "constraint dimer approach." Both methods, unfortunately, become ill-defined if the monomer basis-sets approach completeness. Other methods try to avoid BSSE by keeping the (nonorthogonal) basis sets used by the monomers in a dimer calculation separate throughout the calculation in a valence bond like manner. These methods do have some attractive features, which need to be explored further. Alternatively, some authors have advocated the use of large basis-sets to try to avoid BSSE altogether. For example Schwenke and Truhlar and Frisch et al. concluded that "... in general the extra expense of a counterpoise correction is not warranted, and it is better to increase the basis set to the maximum size affordable for noncounterpoise-corrected calculations." However, considering that in spite of the use of very large basis sets, none of the recent landmark correlated calculations is free from BSSE, this does not represent a realistic strategy.

An entirely opposite approach to the problem of BSSE is to give the monomer calculations the same flexibility that is open to the monomers in the dimer calculation. This, in essence, is the philosophy of the function counterpoise (FCP) procedure. In spite of early criticism of the FCP approach (for a summary, see Ref. 1), there is now a growing consensus that this procedure is a sensible tool for avoiding BSSE in calculated interaction energies as well as...
in properties derived from these, such as equilibrium geometries. 

Nevertheless, there remains a widespread feeling that the BSSE-removal provided by the FCP recipe is only approximate. These hesitations are fueled by occasional findings where the FCP-corrected result is further removed from some desired value (usually an experimental one) than the uncorrected value. However, while one may agree that applying FCP may not be a sufficient remedy in such cases, these examples say very little about the merits of the FCP recipe as such. In fact, the only valid criterion for judging a FCP-corrected result is whether it is close to (or perhaps identical to) the result that should have been expected for the basis set and ab initio method employed in the supermolecular calculation. Thus, "the problem of judging the FCP approach is in knowing the right answer." Attempts to validate the FCP approach in this manner have been made by studying FCP-corrected interaction energies \( \Delta E \) for series of basis sets which differ in the size of the BSSE, but which reasonably may be expected to yield very similar \( \Delta E \) when corrected properly.

A more rigorous and direct assessment of the FCP is to identify the leading basis-set dependent interaction terms for the system at hand and to calculate these by BSSE-free perturbation methods along with the supermolecular interaction energies. The two sets of results should vary with the basis in a parallel fashion if the BSSE removal was successful. Some years ago we reported such parallel calculations of the BSSE-free second-order Möller-Plesset dispersion energy \( E^{(2)}_{\text{MP,disp}} \) and the counterpoise-corrected coupled electron pair approximation (CEPA-1) interaction energies for He, and we were able to show that only the FCP recipe gave acceptable results. Notably, the virtuals-only counterpoise (VCP) recipe, that is repeatedly advocated as a correction scheme superior to the FCP approach, most recently by Yang and Kestner, was shown to fail badly. In our view this proves that the criticism of the FCP variant, which invokes Pauli exclusion principle arguments, is unjustified and results from a misinterpretation of the intermolecular exchange effects.

At the time we were not in a position to show that the FCP-results were exactly right, since \( E^{(2)}_{\text{MP,disp}} \) modeled the dispersion energy present in the CEPA calculations only approximately and, moreover, the dispersion energy is not the only basis-set dependent term in the CEPA interaction energy.

In the present paper we report two extensions of our previous work which allow us to assess the accuracy of the FCP approach more definitively. First, we have obtained new CEPA interaction energies for He, in which only the intramolecular correlation effects are accounted for. The basis-set dependence of the resulting \( E_{\text{int}}^{(1)} \) values is dominated by that of the first-order interaction between correlated He atoms. Second, we have determined the corresponding first-order energies \( E^{(1)} \) directly, employing a BSSE-free perturbation approach. The comparison of these quantities allows us to conclude that the FCP approach yields—to machine precision—exactly the result that should be expected for a given basis and method. As in our previous work we limit ourselves to He, since for this system it is possible to perform a full configuration interaction (CI) calculation on the monomer and evaluate the corresponding first-order energy in any reasonable basis set.

In Sec. II we recall the expressions for the perturbation first-order electrostatic and exchange and the second-order Möller-Plesset dispersion energies as well as the partitioning of the CEPA-1 interaction energy into intra- and intermonomer components. The series of one-electron basis sets used in this study is described in Sec. III. In Sec. IV we compare our perturbation energies \( E^{(1)} \) and \( E^{(2)}_{\text{MP,disp}} \) with the supermolecular CEPA-1 interaction energies corrected for BSSE using the counterpoise method both before and after partitioning into the intra- and intermonomer components. The study is summarized in Sec. V. Throughout the paper we use hartree and kelvin as energy units, 1 \( E_h = 315 777 \text{ K} \).

II. METHODS

In our study we use both the symmetry-adapted perturbation theory of intermolecular forces and the supermolecular CEPA-1 approach to calculate the interaction energy for He,

A. Counterpoise procedures

In the counterpoise (CP) procedure the energy of the dimer is calculated using a supermolecular method and the interaction energy is obtained by subtracting monomer energies resulting from monomer calculations that are performed in functional spaces that simulate the variational freedom available for the monomers in the dimer. There are a few variants of the counterpoise method that differ in the choice of these functional spaces. In the original Boys and Bernardi FCP variant the monomers take advantage of the whole dimer-centered basis set (DCBS).

In the alternative VCP procedure of Daudey et al., the energy of each monomer is calculated using its monomer-centered basis set (MCBS), augmented only by the virtual self-consistent field (SCF) orbitals of the partner monomer.

For comparison, we define a scheme no-CP in which no counterpoise correction is applied.

B. First-order interaction energy

The first-order interaction energy is defined within the symmetry-adapted perturbation theory of intermolecular forces as

\[
E^{(1)} = \frac{\langle \phi_A \phi_B | V_{\text{eff}} | \phi_A \phi_B \rangle}{\langle \phi_A \phi_B | \phi_A \phi_B \rangle},
\]

where \( \phi_A \) and \( \phi_B \) are the unperturbed wave functions of \( A \) and \( B \), respectively, \( V \) is the interaction operator, and \( \text{eff} \) is the antisymmetrizer for the dimer. Since the monomer Hamiltonians are absent in Eq. (1), \( E^{(1)} \) is free of BSSE. The first-order interaction energy is usually split into the classical electrostatic interaction energy \( E_{\text{clast}}^{(1)} \) and the non-
classical exchange repulsion energy $E_{\text{exch}}^{(1)}$ that results from imposing the antisymmetry condition on the wave function of the whole complex

$$E^{(1)} = E_{\text{elst}}^{(1)} + E_{\text{exch}}^{(1)}$$

(2)

$$E_{\text{elst}}^{(1)} = \langle \psi^A \psi^B | V | \psi^A \psi^B \rangle,$$

(3)

and $E_{\text{exch}}^{(1)}$ is defined by Eqs. (1)-(3).

In the present study the monomer wave functions $\psi^A$ or $\psi^B$ are taken to be correlated (full CI) wave functions, employing either the MCBS of $A$ or $B$ only or the entire DCBS.

The $E^{(1)}$ results reported in this study were obtained with the program COREX first devised by Chalasinski. This program can handle interactions between two-electron systems only.

### C. Second-order Möller–Plesset dispersion energy

The second-order Möller–Plesset dispersion energy for the ground state of $\text{He}_2$ is defined as

$$E_{\text{MP\_disp}}^{(2)} = 4 \sum_{n=1}^{\text{virt}} \sum_{m=1}^{\text{virt}} \left( \frac{1}{r_{12}} \right)^2 \left( \left| 1 \right|^2 - \left| 1 \right|^2 \right)$$

(4)

where the occupied $\{1\}$, $\{m\}$ and virtual $\{n\}$ orbitals diagonalize the Fock operators of atoms $A$ and $B$, providing the orbital energies $\epsilon_{1 \sigma}$, $\epsilon_{1 \sigma}$, $\epsilon_{n \sigma}$, and $\epsilon_{n \sigma}$, respectively. The variational character of $E_{\text{MP\_disp}}^{(2)}$ (Ref. 43) permits optimization of non-linear parameters for polarization basis functions as well as using non-MCBS-type basis sets to obtain the monomer orbitals involved in Eq. (4). Like $E^{(1)}$, $E_{\text{MP\_disp}}^{(2)}$ is free of BSSE.

In the dispersion energy of Eq. (4) the inter–intra correlation coupling is completely neglected. $E_{\text{MP\_disp}}^{(2)}$ is related to the uncoupled Hartree-Fock dynamic polarizabilities. The resulting dispersion interaction for $\text{He}_2$ is underestimated by $-5$ K at $R=5.6$ a.u. The dispersion interaction $E_{\text{MP\_disp}}^{(2)}$ is contained in the supermolecular CEPA-1 energy. One cannot, however, expect a close agreement between the CEPA-1 and $E_{\text{MP\_disp}}^{(2)}$ dispersion energies since in the former some inter-intra correlation terms are taken into account. The dispersion energy is the component of the interaction energy that is the most sensitive to the quality of the basis sets used. We monitor this basis set dependence by calculating $E_{\text{MP\_disp}}^{(2)}$ and we expect a qualitatively similar dependence for the CEPA-1 dispersion interaction to be observed.

When optimizing the exponents of polarization sets for the dispersion energy we have used monomer SCF orbitals obtained with the appropriate MCBS and for every angular quantum number $l$ we optimized the partial wave component $E_{\text{MP\_disp}}^{(2)}(1,l)$. On the other hand, we now need $E_{\text{MP\_disp}}^{(2)}$ to estimate the accuracy of the dispersion energy that is operative in the supermolecular CEPA-1 approach and so we now use the DCBS to obtain the SCF spectra of the monomers. Using the language of valence bond theory, charge transfer structures are then allowed for in Eq. (4). These significantly improve the quality of $E_{\text{MP\_disp}}^{(2)}$ making up for the incompleteness of the monomer spectra obtained with MCBS’s.

The $E_{\text{MP\_disp}}^{(2)}$ results reported in this study were obtained with the program INTACT of Groen and van Lenthe.48

### D. CEPA-1 interaction energy

The CEPA-1 method may be considered either as a modification of the configuration interaction singles and doubles method (CISD) aiming at size consistency or as a simplified version of the inherently size consistent coupled cluster singles and doubles method. The CEPA-1 method is not only size consistent but it also provides the exact (full-CI) energy for the system of $N$ separated electron pairs, e.g., for $\text{He}_2$. The method is invariant with respect to a transformation from localized to delocalized internal orbitals for a system of separated electron pairs, and is generally close to unitary invariance. In fact, we verified that the CEPA-1 dimer total energies for $\text{He}_2$ are numerically invariant with respect to Foster-Boys localization of the occupied dimer SCF orbitals.

Taking advantage of the size consistency, the CEPA interaction energy is defined as

$$E_{\text{int}}^{\text{CEPA}} = E_{\text{dim}}^{\text{CEPA}} - E_A^{\text{CEPA}} - E_B^{\text{CEPA}},$$

(5)

where $E_{\text{dim}}^{\text{CEPA}}$ and $E_{\text{X}}^{\text{CEPA}}$ are the dimer and the monomer ($X=A,B$) CEPA energies, respectively. In the case of a two-electron monomer $X$ the energy $E_{\text{CEPA}}^{\text{X}}$ is identical to the full-CI (CISD) energy of $X$.

In addition to the full CEPA interaction energy, we use the Foster-Boys localized dimer internal orbitals to calculate an interaction energy that is more readily comparable to the energy $E^{(1)}$ from the perturbation theory. In this case we impose the restriction that the double excitations have to be localized on one of the monomers, i.e., the dispersion-type double excitations are omitted in the dimer wave function. This approach was also used by Meyer et al. to calculate the “intramonomer” correlation contribution to the interaction energy for the $\text{HeH}_2$ and $\text{He}_2$ systems. For a system of $N$ separated electron pairs, e.g., for $\text{He}_N$, the CEPA-intra method still provides the exact (full-CI) energy. For finite separations between the two-electron systems $A$ and $B$ the CEPA-intra interaction energy, $E_{\text{int}}^{\text{CEPA\_intra}}$

$$E_{\text{int}}^{\text{CEPA\_intra}} = E_{\text{dim}}^{\text{CEPA\_intra}} - E_A^{\text{CEPA\_intra}} - E_B^{\text{CEPA\_intra}},$$

(6)

covers the first-order interaction energy between the correlated monomers as well as some deformation terms. The $E^{(1)}$ energy for correlated monomers [cf. Eq. (1)] may be considered to be the first approximation to the $E_{\text{int}}^{\text{CEPA\_intra}}$ energy. This arises because $E^{(1)}$ is based upon the Heitler–London dimer wave function $\langle \text{HF}_A + \text{HF}_B + \text{S}_A + \text{S}_B \rangle$, where HF is the Hartree–Fock reference and S(D) stands for the single (double) excitations. This wave function contains exactly the same quadrupole excitations that CEPA is designed to account for. This is the reason that $E_{\text{int}}^{\text{CEPA\_intra}}$ and $E^{(1)}$ are so comparable.
The CEPA results reported in this study were obtained with the ATMOL system of codes.\textsuperscript{50}

III. BASIS SETS

The Gaussian basis sets used in this work were obtained as combinations of two basis subsets. The first subset denoted \( CR_n \) \((n=0,...,4)\) was designed to describe the correlation contribution to the atomic energy of the ground state \(^1\text{S}_0\) helium atom. The second subset denoted \( DS_m \) \((m=1,...,5)\) was created to describe the dispersion interaction energy for the helium dimer in its ground \(^1\text{S}_e^+\) electronic state. When advancing in the \( CR_n \) sequence we expect to improve the quality of (a) the helium atom energy, hence the BSSE will decrease; (b) the first-order interaction energy \( E^{(1)} \). On the other hand, when progressing in the \( DS_m \) sequence we will observe (a) improvement in the dispersion interaction energy leading to more attractive CEPA interaction energies; (b) increase of BSSE since the polarization functions with low exponents centered on monomer A can effectively contribute to the lowering of the monomer B energy. Throughout the calculations, Gaussian orbitals with spherical harmonics for angular parts were used.

A. \( CR_n \) subsets

These basis sets were optimized for the CISD energy of the ground state helium atom.

For the radial correlation energy the \( k \)-term contracted sets of symmetry \( s \) were obtained from the 15-term regularized even-tempered (RET15) primitive set\textsuperscript{31} by leaving \( k-1 \) primitives with low exponents uncontracted and by contracting the remaining primitives with high exponents into one basis function. The SCF energy of the helium atom calculated with the RET15 basis set differs by 0.88 \( \mu \text{H} \) from the Hartree–Fock limit value.\textsuperscript{52} Formally, this basis set then leaves room for a BSSE of 0.6 K at the SCF level of theory, but in practice the SCF BSSE never exceeds 0.005 K in our study. The RET15 basis set secures correct values of the first-order electrostatic and exchange components of the SCF interaction energy calculated with DCBS monomer wave functions.\textsuperscript{35} In addition, the completely uncontracted RET15 set reproduces 99.98\% of the radial correlation energy of the helium atom.\textsuperscript{53}

Saturation of the angular correlation energy requires extended polarization sets. The exponents \( \xi_{ik} \) of the \( k \)-term uncontracted basis sets of symmetry \( l \) \((l=1,...,4)\) were restricted to follow the even-tempered relation\textsuperscript{51}

\[
\xi_{ik} = \alpha_k \beta_k \quad i=1,...,k
\]  

and the parameters \( \alpha_k \) and \( \beta_k \) were optimized for the correlation energy of the helium atom. The optimal parameters \( \alpha_k \) and \( \beta_k \) are reported in Table I. Table I also shows the related partial correlation energy contributions to the atomic energy, \( E_{ik} \), when adding the \((l,k)\) set to a set saturated up to \((l-1)\) symmetry and these are compared with the limiting values of \( E_{ik} \) obtained by Bunge.\textsuperscript{53} There are remarkable regularities in the optimal exponents. In particular, for fixed \( k \), the optimal values of the exponents increase as \( l \) increases and simultaneously the spacing between the exponents decreases. This is visualized in Fig. 1 for two- and one-term sets with different \( l \)s. Clearly, an extrapolation is possible to obtain more extended basis sets (larger \( l \)s and \( k \)s).

The \( CR_n \) sets were built from differently contracted RET15 \( s \)-symmetry sets and the uncontracted polarization sets from Table I using the criterion that for every symmetry \( l \) the error in the partial correlation energy should not exceed a judiciously chosen threshold value. The resulting \( CR_n \) sequence \((n=0,...,4)\), the related thresholds, and the remaining errors in the atomic correlation energy for every \( CR_n \) set are shown in Table II. Even the most extended basis sets optimized for the monomer energy still leave ample room for BSSE at the correlated level of theory. For the high-\( l \) symmetries we found that the \( f \)- and in particular \( g \)-symmetry functions optimized for the atomic

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
\( l \) & \( k \) & \( \alpha_k \) & \( \beta_k \) & \( -\Delta E_{ik} \) \\
\hline
1 & 1 & 1.07295 & 1.0 & 1.709(2) \\
1 & 2 & 0.18911 & 4.01413 & 2.073(2) \\
1 & 3 & 0.17333 & 3.24071 & 2.014(3) \\
1 & 4 & 0.16383 & 2.78672 & 2.014(4) \\
1 & 5 & 0.16700 & 2.52000 & 2.014(7) \\
\hline
\hline
1 & \( \infty \) & 2.149(2) & 2.149(2) & 2.014(2) \\
2 & 1 & 1.95910 & 1.0 & 1.607(3) \\
2 & 2 & 0.34626 & 3.50198 & 2.071(3) \\
2 & 3 & 0.31853 & 2.88507 & 2.019(2) \\
2 & 4 & 0.27300 & 2.66200 & 2.022(3) \\
2 & \( \infty \) & 2.254(2) & 2.254(2) & 2.022(2) \\
3 & 1 & 2.0754 & 1.0 & 3.565(4) \\
3 & \( \infty \) & 0.54181 & 3.16636 & 4.819(4) \\
3 & \( \infty \) & 5.330(4) & 5.330(4) & 4.819(4) \\
4 & 1 & 3.4274 & 1.0 & 1.140(4) \\
4 & \( \infty \) & 1.650(4) & 1.650(4) & 1.140(4) \\
\hline
\end{tabular}
\caption{The parameters \( \alpha_k \) and \( \beta_k \) that define the exponents of the \( k \)-term basis set of symmetry \( l \) suitable for the atomic energy of the \(^1\text{S}_0\) He. The related partial energy contributions to the atomic correlation energy \( \Delta E_{ik} \) and the limiting \( \Delta E_{ik} \) values in \( E_{ik} \).}
\end{table}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig1}
\caption{The exponents \( \xi \) of the one- and two-term \( l \)-symmetry \((l=1,...,4)\) basis sets optimized for the atomic correlation energy \( CR \) and for the dispersion interaction energy \( DS \). For two-term sets of \( p \)-symmetry, the low exponent of \( CR \) and the high exponent of \( DS \) practically coincide.}
\end{figure}

J. Chem. Phys., Vol. 98, No. 6, 15 March 1993
energy do not significantly affect the helium dimer interaction energy. The basis set resulting from omitting the two $f$ and one $g$ functions from the CR4 set is labeled CR4'.

### B. $DSm$ subsets

The $DSm$ basis sets derive from the basis sets that were reported in our study of the partial wave expansion and the damping phenomenon for the second-order Møller–Plesset dispersion interaction energy in the ground $1\Sigma^+_g$ state of the helium dimer.\(^{47}\) The exponents of the dispersion-optimized one and two term polarization sets for each symmetry $(p,d,f,...)$, taken from Table I of Ref. 47, are visualized in Fig. 1. As for the CR exponents, the variation with $l$ shows a remarkable regularity. This time, however, the values of the optimal exponents decrease with increasing $l$. A significant observation from Fig. 1 is that the ranges of the dispersion- and atomic-correlation-type exponents become more separated as $l$ increases.

The $DSm$ sets were built from the uncontracted polarization sets\(^{47}\) using the criterion that the error in $E_{\text{MP}_{\text{disp}}}^{(2)}$ caused by using a finite size $(k=1,...,3)$ polarization set for symmetry $l$ is about the same for each $l$ and below a chosen threshold value. The resulting $DSm$ sequence $(m=1,...,5)$, the threshold values, and the remaining errors in the total second-order Møller–Plesset dispersion energy for a given $DSm$ set are shown in Table III. The $E_{\text{MP}_{\text{disp}}}^{(2)}$ energies were obtained with the monomer orbitals calculated with DCBS.

### IV. RESULTS AND DISCUSSION

Calculations were performed for the van der Waals minimum distance in He$_2$, $R=5.6$ a.u. We do not report results involving the $DS1$ basis set since the error in $E_{\text{MP}_{\text{disp}}}^{(2)}$ is rather large (see Table III). Instead of the CR4 basis set we use the CR4' set.

#### A. First order energies $E^{(1)}$ for the CRnDSm sequence

A detailed presentation of $E^{(1)}$ and its components, for some selected basis sets, is given in Table IV. DCBS $E^{(1)}$ results for the remaining sets are shown in Table VI. Table IV is organized such that basis sets containing $s$ functions only are treated first, basis sets having only $s$ and $p$ functions come next, etc. The entries marked "limit" refer to sets which were developed by extending the $s$, $p$, and $d$ parts beyond those of the largest CR$n$ set (i.e., CR4') until no further changes in $E^{(1)}$ were observed (at the DCBS level). The $spd$-limit, however, is an estimate based on the $spd$-limit and the $f$-effect observed for the CR3-results, relative to the corresponding CR2-results. All of these limits are thought to be converged to within 0.01 K (at the DCBS level). They follow the pattern first set out by Chałasiński et al.\(^{54}\) That is, starting from the Hartree–Fock limiting value of 9.69 K the in–out correlation introduced at the $s$-limit makes the atom more diffuse, thereby increasing $E^{(1)}$. On the other hand, the angular terms in the $sp$, $spd$, and $spdf$ sets cause a contraction of the atom, and a progressive lowering of $E^{(1)}$. In view of the convergence in the $s$, $sp$, $spd$, $spdf$ sequence, functions beyond $f$ are not expected to contribute significantly and so we estimate the complete-basis limit for $E^{(1)}$ to be $10.61 \pm 0.01$ K. Knowledge of this quantity is essential in selecting basis sets for full He$_2$ potential calculations that aim at an accuracy at the 0.1% level. Our result is in line with previous estimates by Rybak et al.\(^{55}\) who obtained $E^{(1)}$-values of 10.51, 10.61,
TABLE IV. Convergence of the first-order energy $E^{(1)}$ (in K) and its components to the basis set limit. The required correlated CISD monomer wave functions have been obtained using either the monomer basis set (MCBS, denoted M), or the entire dimer basis (DCBS, denoted D). He$_2$ at $R=5.6$ a.u.

<table>
<thead>
<tr>
<th>Basis set</th>
<th>$P_{el-ex}$</th>
<th>$E^{(1)}_{el-ex}$</th>
<th>$E^{(2)}_{disp}$</th>
<th>$E^{(1)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>D</td>
<td>M</td>
<td>D</td>
</tr>
<tr>
<td>RET15→8s</td>
<td>-2.39</td>
<td>-0.46</td>
<td>-2.01</td>
<td>-2.02</td>
</tr>
<tr>
<td>s-limit$^a$</td>
<td>-2.45</td>
<td>-2.28</td>
<td>-2.01</td>
<td>-2.01</td>
</tr>
<tr>
<td>CR0</td>
<td>-2.08</td>
<td>+0.80</td>
<td>-1.75</td>
<td>-1.78</td>
</tr>
<tr>
<td>spd-limit$^b$</td>
<td>-2.12</td>
<td>-2.05</td>
<td>-1.75</td>
<td>-1.75</td>
</tr>
<tr>
<td>CR1</td>
<td>-2.05</td>
<td>-0.56</td>
<td>-1.72</td>
<td>-1.72</td>
</tr>
<tr>
<td>CR2</td>
<td>-2.05</td>
<td>-0.74</td>
<td>-1.72</td>
<td>-1.72</td>
</tr>
<tr>
<td>CR4$^c$</td>
<td>-2.05</td>
<td>-0.97</td>
<td>-1.72</td>
<td>-1.72</td>
</tr>
<tr>
<td>CR0,DS2</td>
<td>-2.12</td>
<td>+0.32</td>
<td>-1.78</td>
<td>-1.81</td>
</tr>
<tr>
<td>CR1,DS2</td>
<td>-2.08</td>
<td>-1.91</td>
<td>-1.75</td>
<td>-1.74</td>
</tr>
<tr>
<td>CR2,DS2</td>
<td>-2.06</td>
<td>-2.17</td>
<td>-1.73</td>
<td>-1.72</td>
</tr>
<tr>
<td>CR4,DS2$^d$</td>
<td>-2.06</td>
<td>-0.67</td>
<td>-1.72</td>
<td>-1.72</td>
</tr>
<tr>
<td>spd-limit$^d$</td>
<td>...</td>
<td>-2.07</td>
<td>...</td>
<td>-1.72</td>
</tr>
<tr>
<td>CR$^e$</td>
<td>-2.04</td>
<td>-0.74</td>
<td>-1.72</td>
<td>-1.72</td>
</tr>
<tr>
<td>CR3,DS2$^d$</td>
<td>-2.05</td>
<td>...</td>
<td>-1.73</td>
<td>...</td>
</tr>
<tr>
<td>spd-limit$^d$</td>
<td>...</td>
<td>-2.08</td>
<td>...</td>
<td>-1.71</td>
</tr>
</tbody>
</table>

$^a$RET30→12s, monomer CISD DCBS energy = $-2.879 \times 10^2$ E$_h$.

$^b$RET25→11s+6p. The p exponents are 0.15, 0.36, 0.864, 2.074, 4.977, and 11.940. The monomer CISD energy = $-2.900 \times 10^2$ E$_h$ (DCBS).

$^c$Same set as above + 7d. The d exponents are 0.20, 0.60, 1.80, 5.40, 16.20. The monomer CISD energy = $-2.902 \times 10^2$ E$_h$ (DCBS).

$^d$Based on the spd-limit results and on the increments observed going from CR2 to CR3 and from CR2,DS2 to CR3,DS2.

$^e$Based on the spd-limit results and on the increments observed going from CR2 to CR3 and from CR2,DS2 to CR3,DS2.

A final observation concerns the electrostatic energies $E_{el-ex}$ in Table IV. Although on the whole these are more stable to changes in the basis set than are the $E^{(1)}_{el-ex}$ results, this does not go for the separate terms that make up $E^{(1)}_{el-ex}$. As an illustration we quote in Table IV the perturbation part, $P_{el-ex}$, of the electron–electron repulsion term (i.e., its deviation from the value 4/R a.u.). This quantity is well-behaved in MCBS, but the DCBS results are anomalous for all but the largest basis sets employed. The reason for this is that in DCBS the monomers acquire some small artificial multipole moments, and the interaction of the A-atom multipoles with the B-atom monopole (and vice versa) causes $P_{el-ex}$ to deviate from its proper value. However, since each He is a neutral system, these deviations are canceled by corresponding artifacts in the electron–nuclear attraction terms, and the total $E^{(1)}_{el-ex}$ is not affected to any appreciable extent.

B. Second-order Möller–Plesset dispersion energies for the CRnDSm sequence

A selection of the BSSE-free perturbation energies $E_{MP,disp}^{(2)}$, which were obtained using DCBS monomer wave functions, has been included in Table V. The energies become significantly lower when going from one DS set to another, and they also become a little lower as the CR set goes from $n=0$ to 4. These results are consistent with the results of Table III and with the variational character of $E_{MP,disp}^{(2)}$. The best results in the CRnDSm sequence are still some 0.4 K above the basis-set limit, which is currently estimated to be $-17.07$ K.

We have verified that $E_{MP,disp}^{(2)}$ results obtained with the slightly different monomer wave functions that would be consistent with the VCP scheme differ by no more than 0.02 K from those given in Table V. Hence our final conclusions (see below) about the failure of the VCP scheme are not biased by our use of DCBS results throughout.
The negative of $E_{\text{CEPA}}$ energies (K) calculated with the series of the CRnDSm basis sets and not corrected for BSSE were obtained from Eq. (8), using the $E_{\text{disp}}$ values from Table VI. The most striking observation is that the perturbation estimates follow closely the FCP results. The agreement is quite good, i.e., within 0.4 K, or even within 0.05 K, if one excludes the CR0DSm series. There, the estimated results are systematically too repulsive. We had hoped, by taking the basis set dependence of $E_{\text{CEPA}}$ into account, to find a better agreement than in Ref. 33, where we compared $E_{\text{int}}$ to a perturbation estimate based only on $E_{\text{disp}}$. But in fact, the present discrepancies though different from those in Ref. 33 are hardly smaller. The reason for this is, as we discuss at the end of this section, that BSSE is not treated correctly. Therefore, it is not by itself a sufficient guarantee that the FCP approach is fully correct. The BSSE-removal by the counterpoise approach was successful.

Anticipating our final conclusion that the FCP method is correct we calibrate our estimate using the FCP $E_{\text{int}}$ to coincide with the CEPA interaction energy in this basis. The resulting estimated interaction energies are reported in the last column of Table V. The most striking observation is that the perturbation estimates follow closely the FCP results. The agreement is quite good, i.e., within 0.4 K, or even within 0.05 K, if one excludes the CR0DSm series. There, the estimated results are systematically too repulsive. We had hoped, by taking the basis set dependence of $E_{\text{CEPA}}$ into account, to find a better agreement than in Ref. 33, where we compared $E_{\text{int}}$ to a perturbation estimate based only on $E_{\text{disp}}$. But in fact, the present discrepancies though different from those in Ref. 33 are hardly smaller. The reason for this is, as we discuss at the end of this section, that BSSE is not treated correctly. Therefore, it is not by itself a sufficient guarantee that the FCP approach is fully correct. The BSSE-removal by the counterpoise approach was successful.

Anticipating our final conclusion that the FCP method is correct we calibrate our estimate using the FCP $E_{\text{int}}$ to coincide with the CEPA interaction energy in this basis. The resulting estimated interaction energies are reported in the last column of Table V. The most striking observation is that the perturbation estimates follow closely the FCP results. The agreement is quite good, i.e., within 0.4 K, or even within 0.05 K, if one excludes the CR0DSm series. There, the estimated results are systematically too repulsive. We had hoped, by taking the basis set dependence of $E_{\text{CEPA}}$ into account, to find a better agreement than in Ref. 33, where we compared $E_{\text{int}}$ to a perturbation estimate based only on $E_{\text{disp}}$. But in fact, the present discrepancies though different from those in Ref. 33 are hardly smaller. The reason for this is, as we discuss at the end of this section, that BSSE is not treated correctly. Therefore, it is not by itself a sufficient guarantee that the FCP approach is fully correct. The BSSE-removal by the counterpoise approach was successful.

The experimental\cite{57} and the best theoretical predictions\cite{1,58} for the interaction energy at $R=5.6$ a.u. cover the narrow range of $-10.87 \pm 10.95$ K. The CEPA method neglects triple and higher excitations that stabilize the complex by more than 1 K.\cite{46} Hence, it is expected that the basis set limit for the CEPA interaction energy should be around $-9.9$ K. However, the CEPA results sink well below this value for the CRn set (n=0,...;3) series. More interesting, for the CR0 basis set the VCP results drop below this value as well, by as much as 9 K. One would rather expect that the properly CP-corrected interaction energy should be above the CEPA-limit energy because of the unsaturation of the dispersion energy when using the DSm basis sets (cf. Table III). We therefore conclude, as in Ref. 33, that the VCP approach yields unphysical results, which do not correspond to the basis set and correlation method employed in a given CRnDSm CEPA calculation. In some cases the VCP-results may be close to experiment, as in the recent paper by Yang and Kestner,\cite{22} but such agreement is fortuitous and should not be construed as evidence for the correctness of VCP.

By contrast, the FCP results all lie well above the estimated $E_{\text{int}}$-limit of $-9.8$ K. However, this is not by itself a sufficient guarantee that the FCP approach is fully correct. Therefore, it is expected that the dominant components of the interaction energy, viz. $E^{(1)}$ and $E^{(2)}_{\text{disp}}$, are correct. Therefore, the CEPA-intra approach is expected to be correct. The BSSE-removal by the counterpoise approach was successful.

Anticipating our final conclusion that the FCP method is correct we calibrate our estimate using the FCP $E_{\text{int}}$ to coincide with the CEPA interaction energy in this basis. The resulting estimated interaction energies are reported in the last column of Table V. The most striking observation is that the perturbation estimates follow closely the FCP results. The agreement is quite good, i.e., within 0.4 K, or even within 0.05 K, if one excludes the CR0DSm series. There, the estimated results are systematically too repulsive. We had hoped, by taking the basis set dependence of $E_{\text{CEPA}}$ into account, to find a better agreement than in Ref. 33, where we compared $E_{\text{int}}$ to a perturbation estimate based only on $E_{\text{disp}}$. But in fact, the present discrepancies though different from those in Ref. 33 are hardly smaller. The reason for this is, as we discuss at the end of this section, that BSSE is not treated correctly. Therefore, it is not by itself a sufficient guarantee that the FCP approach is fully correct. The BSSE-removal by the counterpoise approach was successful.

These considerations led us to search for a supermolecular approach which produces an interaction energy that is easier to model in parallel perturbation calculations. This can be achieved by adopting, for the dimer calculation, the CEPA-intra approach in which localized dimer internal orbitals are used and the double excitations are restricted to be of the intramonomer-type. Then, the dispersion as well as the inter-intra correlation coupling terms are automatically eliminated. In this approach, the interaction energy contains only the first-order interaction energy $E^{(1)}$ and a small deformation energy arising as the monomers adapt to the Pauli repulsion in the dimer. The method is still prone to BSSE. Values of the FCP-corrected $E$ energies together with the $E^{(1)}$ DCBS energies for the CRnDSm series are reported in Table VI. The two sets of results are strongly correlated, as is visualized in Fig. 2. As before, by calibrating against the CR4'DS4 result, a perturbation estimate of the CEPA-intra result for other sets may be made as follows:

$$E_{\text{int, estimated}}(nm) = E^{(1)}(nm) - 0.207 K,$$

where $-0.207 K$ is the limiting value of the deformation energy, viz. the difference between $E_{\text{int, estimated}}$ and $E^{(1)}$ for the CR4'DS4 basis set. These estimates are shown in the
TABLE VI. The $E^{(1)}$ and the FCP-corrected $E_{\text{CEPA-intra}}$ energies (K) calculated with the series of CR$n$DS$m$ basis sets. Hel at $R = 5.6$ a.u. Perturbation estimates of the interaction energies were obtained from Eq. (9).

<table>
<thead>
<tr>
<th>Basis set</th>
<th>$E^{(1)}$</th>
<th>$E_{\text{CEPA-intra}}$</th>
<th>Estimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR0DS2</td>
<td>11.045</td>
<td>10.778</td>
<td>10.838</td>
</tr>
<tr>
<td>CR0DS3</td>
<td>10.978</td>
<td>10.727</td>
<td>10.771</td>
</tr>
<tr>
<td>CR0DS4</td>
<td>10.971</td>
<td>10.719</td>
<td>10.764</td>
</tr>
<tr>
<td>CR1DS2</td>
<td>10.976</td>
<td>10.708</td>
<td>10.769</td>
</tr>
<tr>
<td>CR1DS3</td>
<td>10.717</td>
<td>10.505</td>
<td>10.510</td>
</tr>
<tr>
<td>CR1DS4</td>
<td>10.712</td>
<td>10.499</td>
<td>10.505</td>
</tr>
<tr>
<td>CR2DS2</td>
<td>10.655</td>
<td>10.447</td>
<td>10.448</td>
</tr>
<tr>
<td>CR2DS3</td>
<td>10.651</td>
<td>10.442</td>
<td>10.444</td>
</tr>
<tr>
<td>CR2DS4</td>
<td>10.650</td>
<td>10.441</td>
<td>10.443</td>
</tr>
<tr>
<td>CR4' DS2</td>
<td>10.639</td>
<td>10.432</td>
<td>10.432</td>
</tr>
<tr>
<td>CR4' DS3</td>
<td>10.637</td>
<td>10.429</td>
<td>10.430</td>
</tr>
<tr>
<td>CR4' DS4</td>
<td>10.636</td>
<td>10.429</td>
<td>10.429</td>
</tr>
<tr>
<td>spd-limitb</td>
<td>10.633</td>
<td>10.426</td>
<td>10.426</td>
</tr>
</tbody>
</table>

*a* Used for calibrating the "estimated" column.  
* [1ls6p5d] basis set from Table IV.

The FCP-corrected CEPA-intra results deviate at most 0.06 K from the perturbation estimates, and as the CR set grows larger the agreement improves quickly, the deviation from CR4' onwards being at most 0.001 K. The initial deviations of up to 0.06 K may be rationalized by supposing that for these sets the deformation energy has not yet reached the limiting value assumed in Eq. (9). (In fact the deformation energy is more negative for the smaller sets in which $E^{(1)}$ is too repulsive.) The final deviations of around 0.001 K appear to be caused mainly by the finite numerical precision of our data. The smallness of these deviations is all the more striking if one recalls that the BSSE for the larger CR sets still has values which range from 0.20–2.82 K (cf. Table V). We conclude that the FCP method does fully remove these BSSE effects and yields interaction energies that are precisely correct, not only for the CR2 and CR4' sets where a perturbation estimate including a converged deformation component is at hand to prove it, but for all other sets as well.

Before concluding this section we return to the question why the perturbation estimate $E^{(1)} + E^{(2)}_{\text{disp}}$ reproduced the full, unpartitioned, CEPA interaction energies only moderately well. Having established that the intracorrelation part of $E^{\text{CEPA-inter}}$ is accurately modeled by the perturbation term $E^{(1)}$, it follows that the remainder of $E^{\text{CEPA-inter}}$, viz. $E^{(2)}_{\text{disp}}$, which mainly consists of the CEPA dispersion energy, is not modeled quite so well by the perturbation estimate $E^{(2)}_{\text{disp}}$. In Fig. 3 we have plotted the latter two quantities against each other, and it turns out that the CEPA-inter energies for the CR0DSm sets lie on a different line (more negative CEPA energies) than those for the other CR$n$DS$m$ (n > 0) sets. This suggests that the quality with which a CR set describes the intramonomer correlation has a direct influence on the CEPA dispersion energy, but not on $E^{(2)}_{\text{disp}}$. This arises because intra-inter correlation coupling terms are present in CEPA, but not in the MP dispersion calculation. We conclude that, as for $E^{(1)}$, the CR part of the basis should be at least of CR2-quality in calculations aiming at a precision of 0.1% in the total He interaction energy.

V. SUMMARY

A series of one-electron basis sets is documented that approach in a systematic way the basis set limit for the potential energy curve in He2. Every basis is composed of two parts (i) a CR set of functions optimized for the atomic correlation energy in He; (ii) a DS set of functions optimized for the dispersion interaction in He2.

With more extended basis sets optimized for the atomic correlation energy only, we determined the spd-limit to be 10.63 ± 0.01 K for the correlated first-order interaction energy in He2 at $R = 5.6$ a.u. The spdf limit is extrapolated to be 10.61 ± 0.01 K. The correlation compon-
In our previous paper various variants of the counterpoise method were tested, comparing the counterpoise corrected supermolecular interaction energies, $E_{\text{int}}$, for several basis sets to the BSSE-free $E_{\text{MP,disp}}$ energy. In the present paper, this approach is extended by including the BSSE-free correlated $E^{(1)}$ as well. For consistency with the supermolecular approach, the DCBS $E^{(1)}_{\text{disp}}$ and $E^{(1)}$ are used. By calibrating the two sets of data at one selected basis set it is then possible to make a BSSE-free estimate of the $E_{\text{int}}$ values to be expected for the other basis sets. The Boys–Bernardi (FCP) counterpoise corrected $E_{\text{int}}$ are found to differ at most 0.4 K from this perturbation-theory estimate.

The accuracy of the FCP scheme was established more definitively by comparing the CP corrected $E_{\text{int}}$ values to the BSSE-free $E^{(1)}$ energies for the series of the $\text{CRnDSm}$ basis sets. $E_{\text{int}}^{\text{CEPA-intra}}$ should contain the same interaction terms as $E^{(1)}$, except for a small deformation term, that is rather insensitive to the choice of basis set. The comparison therefore provides a more sensitive test of the correctness (or otherwise) of a chosen CP procedure than does the comparison performed previously. The $E_{\text{int}}^{\text{CEPA-intra}}$ results were found to be consistent to 0.06 K with the $E^{(1)}$ energies, if the FCP method was used. The agreement improves to 0.001 K, i.e., to machine precision, from CR4 onwards, even though the BSSE never becomes smaller than 0.20 K. This proves conclusively that, in He$_2$, the FCP approach removes all BSSE-artefacts from the interaction energy.

A side effect of the supermolecular approach, not corrected by the FCP counterpoise scheme, is that the DCBS description of the monomers worsens the first-order electrostatic interaction energy $E_{\text{clat}}$ compared to the MCBS description. Analysis of the penetration components of $E^{(1)}$ showed, however, that while this so-called secondary BSSE deteriorates the separate components, nevertheless $E^{(1)}_{\text{clat}}$ itself converges rapidly to a precision of 0.01 K.

We conclude that the Boys and Bernardi function counterpoise method is a correct approach to cope with BSSE at both the SCF and correlated levels of theory. It yields interaction energies that are consistent with the quality of the \textit{ab initio} method and the basis set employed. The perturbation components of the supermolecular interaction energy are reproduced at the level of monomer wave functions obtained in the dimer-centered basis set. The use of bond functions does not present any special problem, since always the complete basis set of the complex is used. For larger systems than He$_2$, when approximate methods are employed, the proper counterpoise does not follow in a straightforward way from the original Boys–Bernardi paper. Both orbital and configuration spaces should match exactly. In full CI calculations the Boys and Bernardi function counterpoise method can be applied as originally suggested.

**ACKNOWLEDGMENTS**

Many discussions with Dr. G. Chalański are gratefully acknowledged. M. G. wishes to thank the University of Utrecht for supplying the funds for his stay in the Netherlands and Mrs. Lidy ten Doesschate for her warm hospitality. This stay was organized on the basis of the Cultural Treaty between Poland and the Netherlands. The hospitality of Professor Jack Simons from the University of Utah, where a part of this work was done, is also appreciated. Dr. E. Earl is acknowledged for reading the manuscript. We thank Dr. V. R. Saunders and Dr. M. F. Guest for making available the ATOMI programs. This research was made possible by a grant of supercomputertime on the Cyber 205 of SARA (Amsterdam) from the foundation SURF through the National Funding of Supercomputers (NFS). M. G. was supported by NSF Grants Nos. CHE-8814765 and CHE-9116286 and the KBN Grant No. 2 0556 91 01.


G. Kaplan, Intermolecular Interactions (Elsevier, Amsterdam, 1987).


