Combined bond-polarization basis sets for accurate determination of dissociation energies

Part 3: Basis set superposition error in polyatomic systems

J. M. L. Martin[1,2], J. P. François[1], and R. Gijbels[2]

[1]Limburgs Universitair Centrum, Department SBM, Universitaire Campus,
B-3610 Diepenbeek, Belgium

[2]University of Antwerp (UIA), Department of Chemistry, Universiteitsplein 1,
B-2610 Wilrijk, Belgium

Summary. The effect of bond functions on the basis set superposition error (BSSE) is investigated at both SCF (self consistent field) and correlated levels for a number of basis sets using the pairwise additive function counterpoise (PAFC), the site-site function counterpoise (SSFC), and the newly proposed successive reaction counterpoise method (SRCP). BSSEs using bond functions are shown to be roughly twice those without bond functions, whereas the latter may still be quite sizeable. The addition of f functions dramatically decreases the bond function BSSE. The results obtained support the empirical decision in our earlier papers to neglect BSSE altogether.

Key words: Basis set — Basis set superposition error (BSSE) — Bond functions — Site-site function counterpoise (SSFC) — Pairwise additive function counterpoise (PAFC) — Successive reaction counterpoise (SRCP)

1. Introduction

For overcoming the so-called "molecular polarization problem" in the choice of basis sets for ab initio calculations, a promising approach, the so-called combined bond-polarization basis sets, was introduced recently [1–5]. (In the subsequent discussion, papers [3–5] will be referred to as Parts 1, 2, and 4, respectively.) The literature and theoretical discussion of the subject can be found in these papers (especially Part 2), and will not be repeated here.

The main objection against bond functions in correlated *ab initio* calculations, the exaggerated basis set superposition error (BSSE) [6], was shown to be greatly reduced by sufficiently saturating the polarization complement with respect to the atomic correlation energy [Part 2]. However, this was only shown

for the diatomic first-row hydrides. On the basis of these results, basis set superposition error was neglected entirely for both diatomic and polyatomic hydrides, resulting in excellent agreement with experiment (2 kcal/mol or better). The evaluation of polyatomic BSSEs is no trivial matter; in this paper, an effort is made to investigate BSSEs separately.

2. Theoretical aspects

The familiar counterpoise method of Boys and Bernardi [7] can be formulated as follows. Suppose it is required to calculate the interaction energy between the systems (atoms, molecules) A and B, usually called monomers, with energies $E(A)$ and $E(B)$. The interaction energy ΔE of the supersystem AB at some nuclear geometry is given by

$$\Delta E = E(AB) - E(A) - E(B). \tag{1}$$

In this equation, however, the effect of basis set superposition is not taken into account. Following Boys and Bernardi's idea, that "the full set of expansion functions used in the dimer calculation must also be used in the monomer calculations", the interaction energy ΔE should be written as

$$\Delta E' = E(AB) - E(AG_B) - E(BG_A), \tag{2}$$

where $E(AG_B)$ represents the energy of monomer A with the basis functions for B added to the basis set (so-called "ghost orbitals"), and analogously for $E(BG_A)$. The BSSE is then defined as follows:

$$\text{BSSE} = \Delta E - \Delta E' = E(AG_B) - E(A) + E(G_AB) - E(B). \tag{3}$$

The Boys and Bernardi method has received firm support over the years [8] from theoretical studies, in contrast to the results of empirical studies [9] which have indicated that it is not quantitative, and suggest that it can only be used as a crude indicator for basis set saturation.

In a very recent paper, however, Mayer and Vibok [10] computed SCF interaction energies using a "chemical Hamiltonian" formalism [11] which is rigorously free of BSSE, and concluded that the interaction energy and its BSSE correction are not additive; to quote these authors, "the non-additivity of the BSSE" (and the interaction energy, author's note) "makes meaningless the polemy about whether or not the Boys–Bernardi scheme overestimates BSSE: in fact, it can overestimate or underestimate the value of BSSE (see also [12]) or even give a wrong sign. This rules out all the conventional schemes used for correcting BSSE". Mayer also gave a theoretical proof [13] for the nonadditivity of BSSE and interaction energy in the simplest possible case. Summing things up, no definitive conclusion as to the validity of the counterpoise procedure appears to have been reached at this stage.

The generalization to polyatomic systems (or many-body molecular interactions) is not unique. Wells and Wilson [14] present two possible approaches. In the pairwise additive function counterpoise (PAFC) approach, the basis set

superposition error is approximated by a sum of Boys–Bernardi BSSEs on pair interactions:

$$\text{PAFC} = \sum_{i \neq j} (E(A_i G_j) - E(A_i)), \tag{4}$$

where the summation indices i and j run over all the monomers in the system.

In the site-site function counterpoise (SSFC) approach, however, it is assumed that all orbitals in the many-body system are available to any of the monomers, and thus the total basis set superposition error may be written as:

$$\text{SSFC} = \sum_{i} (E(A_i G_{jkl\cdots}) - E(A_i)), \tag{5}$$

where i again runs over all the monomers, and $G_{jkl\cdots}$ represents the ghost orbitals on all the other monomers. The SSFC expression has the advantage that a meaningful decomposition of the total interaction energy is still possible. Its main disadvantage is the necessity to evaluate the quadratically increasing number of different terms in Eq. (5), so the PAFC (Eq. 4) is certainly more economical. Wells and Wilson concluded from SCF calculations on helium dimers and trimers that the PAFC tends to overestimate the BSSE (with respect to the SSFC approach), and that the difference between either approach was to some extent a function of the geometric arrangement of the monomers in the adduct. A recent example of an application of SSFC may be found in [15].

To our knowledge no systematic comparison of both approaches at SCF or correlated levels has as yet been undertaken. The present work attempts to provide such a comparison for a moderately-sized class of molecules, the open- and closed-shell first-row hydride species in their respective ground states, for six different basis sets. Three of these are standard [16] (6-311 + G(d, p), 6-311 + G($2d, p$), and 6-311 + G($2df, p$)), the remaining three basis sets contain a single sp-shell halfway along each bond [Part 2] (6-311 + G(d, p)B, 6-311 + G($2d, p$)B and 6-311 + G($2df, p$)B). Electron correlation is included using full fourth-order Møller–Plesset theory (MP4(SDTQ) with a frozen core) [17].

All these basis sets have the common property, that the basis set for hydrogen is nearly saturated: in two-atom counterpoise calculations, its energy was lowered by 10 μE_h or less due to the basis functions of the heavy atom, whereas the BSSE in H_2 is even much smaller. Henceforth, to a very good approximation, the terms in Eqs. (4) and (5) containing only hydrogen atoms may be neglected. For AH_n this leads to the following simplified expressions:

$$\text{PAFC} = n(E(AG_H) - E(A) + E(HG_A) - E(H)), \tag{6}$$

$$\text{SSFC} = E(AG_{H_n}) - E(A) + n(E(HG_A) - E(H)). \tag{7}$$

In other terms, the PAFC of AH_n is exactly n times the counterpoise-BSSE for AH, whereas the SSFC is not. Therefore, one is here comparing plain additive with nonadditivity.

If the basis set of hydrogen were completely saturated, Eqs. (6) and (7) could be further simplified to:

$$\text{PAFC} = n(E(AG_H) - E(A)), \tag{8}$$

$$\text{SSFC} = E(AG_{H_n}) - E(A). \tag{9}$$

A third possibility besides the PAFC and SSFC approaches is the stepwise formation of the polyatomic molecule. In that case, the total formation reaction can be written as a sequence of binary formation reactions, and the BSSE for each of them can be computed using the normal counterpoise procedure; this eliminates all worries about nonadditivity. For general cases, this approach has the disadvantage that its application is not unique; for AH_n molecules however, there is an obvious way to proceed. We illustrate it by considering water.

The first step in the formation of water from its atoms will be

$$O(^3P) + H(^2S) \rightarrow OH(^2\Sigma^+). \tag{10a}$$
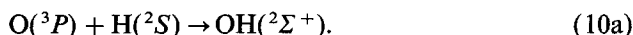
The counterpoise energy for this step is of course easily written as

$$\text{BSSE} = E(OG_H) - E(O) + E(HG_O) - E(H). \tag{10b}$$

Next comes the second association

$$OH + H \rightarrow H_2O \tag{10c}$$

for which the counterpoise energy can be written as

$$\text{BSSE} = E(OHG_H) - E(OH) + E(HG_{OH}) - E(H). \tag{10d}$$

The total BSSE then becomes

$$\text{BSSE} = (E(OHG_H) - E(OH)) + (E(HG_{OH}) - E(H))$$
$$+ (E(OG_H) - E(O)) + (E(HG_O) - E(H)). \tag{10e}$$

Of course, this definition is not unique. One might just as well consider the following scheme:

$$H + H \rightarrow H_2, \tag{11a}$$

$$H_2 + O \rightarrow H_2O, \tag{11b}$$

which leads to the following counterpoise expression

$$\text{BSSE} = (E(OG_{H_2}) - E(O)) + (E(H_2G_O) - E(H_2)) + 2(E(HG_H) - E(H)). \tag{11c}$$

However, if we consider the H basis set to be saturated, the second and third terms of Eq. (11c) vanish, reducing the equation to the form of Eq. (9) for the SSFC method under the same assumption. Henceforth, we will be using Eq. (10e) as a basis set superposition error estimate, which we will term the successive reaction counterpoise method (SRCP).

For more than triatomic systems, geometry changes can present additional problems. For example in the $CH_3 + H \rightarrow CH_4$ reaction, CH_3 has point group

D_{3h}, whereas CH_4 has point group T_d. We may split the reaction in two imaginary steps:

1) the geometry change $CH_3(D_{3h}) \rightarrow CH_3(C_{3v})$, for which no BSSE applies;

2) the reaction $CH_3(C_{3v}) + H \rightarrow CH_4(T_d)$, for which the counterpoise amounts to

$$BSSE = (E(CH_3(C_{3v})G_H) - E(CH_3(C_{3v}))) + (E(HG_{CH_3(C_{3v})}) - E(H)). \quad (12)$$

If more than three atoms are involved, a number of supplementary reaction schemes can be constructed. We will avoid these here, since only the scheme outlined in Eq. (10) corresponds to the sequential association which is relevant here.

SRCP will be by far the most expensive method considered here, and requires an operational definition based on the reaction scheme. On the other hand, it is entirely defined in terms of two-body BSSEs, and will thus not suffer from many-body nonadditivity effects.

3. Computational methods

All calculations were carried out using the GAUSSIAN 86 program system [18] running on a MicroVAX 2000 workstation under the VMS 4.7 operating system. To remain consistent with Parts 1, 2 and 4, HF/6-31G* geometries [19] were used throughout the present study.

For the systems CHG_H and OHG_H, the SCF criterion had to be loosened from the standard value of 1.0D-9 (Euclidean norm of the difference between consecutive density matrices) to 1.0D-5 and 1.0D-6, respectively, in order to obtain convergence. Both with and without DIIS [20], as well as when doing steepest-descent direct minimization SCF [21], the iteration started cycling indefinitely when better convergence was requested. This effect did not occur in any of the other systems.

4. Results, discussion, and conclusions

Total energies for the atomic systems may be found in Part 2, and are not repeated here. The BSSE increments per bond according to the SSFC method (Eq. 7) are given in the first part of Table 1 for the $6\text{-}311 + G(d, p)$ B, $6\text{-}311 + G(2d, p)$ B, and $6\text{-}311 + G(2df, p)$ B basis sets, and in the second part for the corresponding polarization-only basis sets, i.e. $6\text{-}311 + G(d, p)$, $6\text{-}311 + G(2d, p)$, and $6\text{-}311 + G(2df, p)$. The BSSE increments per bond for the PAFC method (Eq. 6) are of course all identical to the BSSE for the first step. Table 2 presents the same analysis as Table 1, but now for the SRCP method and omitting the $6\text{-}311 + G(2d, p)$ and $6\text{-}311 + G(2d, p)$ B basis sets. Finally, in Table 3 binding energies (without BSSE correction) at the corresponding levels of theory have been given, as well as experimental values corrected for zero point energy. The numerical values in Table 3 for the $6\text{-}311 + G(2d, p)$ and

Table 1. Basis set superposition error increments (kcal/mol) as estimated by the site-site function counterpoise method (SSFC)

Species	6-311 + $G(d, p)$B				6-311 + $G(2d, p)$B				6-311 + $G(2df, p)$B			
	UHF	MP2	MP3	MP4	UHF	MP2	MP3	MP4	UHF	MP2	MP3	MP4
HF	1.02	5.61	5.88	6.12	0.90	4.28	4.49	4.64	0.47	2.32	2.24	2.34
OH	0.35	4.27	4.53	4.71	0.33	3.26	3.30	3.39	0.22	1.83	1.72	1.76
OH_2	0.81	6.70	7.16	7.29	0.57	5.11	5.37	5.44	0.25	2.87	2.81	2.84
NH	0.08	3.13	3.18	3.29	0.09	2.44	2.29	2.31	0.09	1.44	1.27	1.28
NH_2	0.09	2.87	2.89	2.98	0.09	2.19	2.04	2.05	0.10	1.28	1.13	1.13
NH_3	0.10	2.45	2.42	2.49	0.10	1.94	1.80	1.80	0.10	1.20	1.05	1.05
CH	0.34	1.68	1.46	1.47	0.09	1.36	1.12	1.10	0.09	0.97	0.79	0.78
CH_2	−0.16	2.65	2.62	2.62	0.07	2.29	2.11	2.04	0.08	1.33	1.16	1.12
CH_3	0.21	1.95	1.85	1.86	0.09	1.52	1.35	1.31	0.09	0.91	0.78	0.76
CH_4	0.20	0.81	0.60	0.60	0.07	0.60	0.41	0.40	0.07	0.46	0.30	0.30
BH	0.13	0.96	0.71	0.66	0.05	0.91	0.67	0.62	0.05	0.66	0.48	0.46
BH_2	0.21	1.93	1.71	1.64	0.07	1.49	1.22	1.10	0.07	0.89	0.69	0.63
BH_3	0.10	0.56	0.36	0.35	0.03	0.41	0.22	0.21	0.03	0.27	0.15	0.14
BeH	0.02	0.45	0.21	0.17	0.02	0.52	0.27	0.22	0.02	0.39	0.21	0.20
BeH_2	0.01	0.28	0.09	0.07	0.02	0.28	0.12	0.09	0.02	0.20	0.09	0.09
LiH	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
H_2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Species	6-311+G(d,p)				6-311+G(2d,p)				6-311+G(2df,p)			
	UHF	MP2	MP3	MP4	UHF	MP2	MP3	MP4	UHF	MP2	MP3	MP4
HF	0.76	2.55	2.76	2.93	0.60	1.34	1.39	1.47	0.45	0.99	1.00	1.07
OH	0.28	1.77	1.97	2.12	0.27	1.04	1.07	1.12	0.20	0.66	0.64	0.67
OH$_2$	0.32	2.75	2.99	3.14	0.16	1.51	1.58	1.66	0.20	1.05	1.02	1.07
NH	0.06	1.24	1.36	1.46	0.06	0.83	0.78	0.80	0.07	0.50	0.44	0.44
NH$_2$	0.06	1.23	1.33	1.42	0.07	0.81	0.76	0.77	0.08	0.50	0.44	0.45
NH$_3$	0.07	1.16	1.23	1.30	0.08	0.80	0.74	0.76	0.08	0.51	0.45	0.46
CH	0.30	0.73	0.68	0.70	0.07	0.49	0.42	0.41	0.08	0.34	0.28	0.28
CH$_2$	-0.15	1.25	1.34	1.38	0.05	1.03	0.95	0.92	0.06	0.62	0.54	0.52
CH$_3$	0.19	1.01	1.03	1.07	0.09	0.72	0.65	0.64	0.09	0.51	0.45	0.44
CH$_4$	0.17	0.38	0.31	0.31	0.07	0.42	0.34	0.33	0.07	0.24	0.19	0.19
BH	0.11	0.41	0.33	0.32	0.04	0.31	0.22	0.20	0.01	0.03	0.02	0.02
BH$_2$	0.21	1.08	1.02	1.03	0.05	0.77	0.63	0.57	0.07	0.65	0.50	0.46
BH$_3$	0.07	0.25	0.18	0.17	0.03	0.20	0.12	0.11	0.03	0.17	0.11	0.11
BeH	0.01	0.16	0.05	0.03	0.01	0.15	0.05	0.03	0.01	0.12	0.05	0.05
BeH$_2$	0.01	0.15	0.06	0.05	0.01	0.16	0.07	0.06	0.01	0.09	0.04	0.04
LiH	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
H$_2$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 2. Basis set superposition error increments (kcal/mol) as estimated by the successive reaction counterpoise method (SRCP)

Species	6-311 + $G(d,p)$B				6-311 + $G(df,p)$B			
	UHF	MP2	MP3	MP4	UHF	MP2	MP3	MP4
HF	1.02	5.61	5.88	6.12	0.47	2.32	2.24	2.34
OH	0.35	4.27	4.53	4.71	0.22	1.83	1.72	1.76
OH_2	0.58	5.41	5.68	5.86	0.27	2.17	2.05	2.15
NH	0.08	3.13	3.18	3.29	0.09	1.44	1.27	1.28
NH_2	0.19	3.29	3.37	3.50	0.13	1.43	1.29	1.33
NH_3	0.52	4.05	4.12	4.24	0.22	1.71	1.56	1.63
CH	0.34	1.68	1.46	1.47	0.09	0.97	0.79	0.78
CH_2	0.26	2.73	2.67	2.69	0.12	1.32	1.15	1.14
CH_3	0.12	1.60	1.61	1.70	0.12	0.80	0.73	0.76
CH_4	0.18	1.83	1.82	1.90	0.12	1.01	0.91	0.95
BH	0.13	0.96	0.71	0.66	0.05	0.66	0.48	0.46
BH_2	0.08	1.36	1.09	1.03	0.05	0.68	0.48	0.46
BH_3	0.03	0.72	0.64	0.62	0.03	0.35	0.25	0.24
BeH	0.02	0.45	0.21	0.17	0.02	0.39	0.21	0.20
BeH_2	0.10	0.27	0.29	0.29	0.04	0.18	0.18	0.18
LiH	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
H_2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Species	6-311 + G(d, p)				6-311 + G(2df, p)			
	UHF	MP2	MP3	MP4	UHF	MP2	MP3	MP4
HF	0.76	2.55	2.76	2.93	0.45	0.99	1.00	1.07
OH	0.28	1.77	1.97	2.12	0.20	0.66	0.64	0.67
OH$_2$	0.41	2.51	2.71	2.87	0.23	0.89	0.87	0.94
NH	0.06	1.24	1.36	1.46	0.07	0.50	0.44	0.44
NH$_2$	0.19	1.51	1.64	1.76	0.10	0.59	0.54	0.57
NH$_3$	0.47	2.12	2.24	2.35	0.17	0.82	0.77	0.82
CH	0.30	0.73	0.68	0.70	0.08	0.34	0.28	0.28
CH$_2$	0.23	1.49	1.56	1.61	0.09	0.64	0.57	0.57
CH$_3$	0.09	0.80	0.86	0.92	0.09	0.43	0.41	0.43
CH$_4$	0.16	0.92	0.98	1.04	0.09	0.48	0.45	0.49
BH	0.11	0.41	0.33	0.32	0.01	0.03	0.02	0.02
BH$_2$	0.03	0.72	0.64	0.62	0.03	0.35	0.25	0.24
BH$_3$	0.17	0.59	0.60	0.63	0.07	0.31	0.28	0.29
BeH	0.01	0.16	0.05	0.03	0.01	0.12	0.05	0.05
BeH$_2$	0.09	0.22	0.23	0.24	0.03	0.13	0.14	0.14
LiH	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
H$_2$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 3. Theoretical and experimental dissociation energies (kcal/mol; no BSSE correction)

Species	6-311 + G(d,p)B				6-311 + G(2d,p)B				6-311 + G(2df,p)B				Experimental
	UHF	MP2	MP3	MP4	UHF	MP2	MP3	MP4	UHF	MP2	MP3	MP4	[3,22]
HF	97.5	143.3	137.4	141.0	97.4	144.3	138.0	141.5	96.7	142.9	136.4	139.7	141.2
H_2O	155.3	230.8	225.2	229.5	155.9	232.1	225.7	229.9	155.3	230.6	223.9	228.0	232.2
OH	67.8	103.7	102.6	104.2	68.3	104.8	103.2	104.8	68.1	104.2	102.6	104.1	106.6, 106.6 ± .3
NH_3	199.2	287.0	286.0	289.6	200.2	288.6	286.6	290.3	200.3	288.4	286.7	290.4	297.3 ± .1
NH_2	115.5	171.4	172.6	174.7	116.9	173.0	173.6	175.6	117.0	173.2	174.2	176.2	180.9 ± 1.5
NH	50.1	76.1	77.6	78.6	50.9	76.9	78.1	79.1	51.0	77.0	78.5	79.5	78.8 ± 4, ≤ 84.7
CH_4	327.0	409.5	412.3	414.1	327.5	410.3	412.6	414.4	327.5	409.7	412.2	414.3	419.8 ± .1
CH_3	242.0	299.4	301.8	302.5	242.3	300.1	302.1	302.8	242.3	299.6	301.9	302.7	307.2 ± .2
CH_2	154.5	186.8	187.7	187.5	154.5	187.1	187.7	187.5	154.5	186.6	187.5	187.4	189.8 ± .6
CH	56.0	77.4	79.6	80.3	56.5	78.1	80.1	80.7	56.6	78.2	80.5	81.2	84.4 ± .1, 83.9
BH_3	231.0	273.5	276.8	276.5	231.2	273.7	277.0	276.8	231.2	273.6	277.3	277.2	278.1 ± 10
BH_2	143.1	166.9	167.8	167.0	143.3	167.0	168.0	167.2	143.4	166.8	168.1	167.5	196.2 ± 15
BH	61.6	79.4	81.7	82.2	61.7	79.9	82.2	82.6	61.7	79.8	82.2	82.7	81.6 ± 2, 82.2
BeH_2	125.2	143.9	145.3	144.5	125.3	144.1	145.8	145.1	125.2	143.2	145.2	144.5	
BeH	50.3	52.7	50.9	49.4	50.4	52.8	51.3	50.0	50.3	52.0	50.8	49.4	49.8, 55.8 ± 7
LiH	33.9	49.6	53.5	54.7	34.0	49.8	53.7	54.8	34.0	50.2	54.0	55.2	58.0
H_2	83.4	101.5	105.4	106.4	83.4	101.5	105.4	106.4	83.4	101.5	105.4	106.4	109.5

| Species | 6-311+G(d,p) | | | | 6-311+G(2d,p) | | | | 6-311+G(2df,p) | | | | Experimental |
	UHF	MP2	MP3	MP4	UHF	MP2	MP3	MP4	UHF	MP2	MP3	MP4	[3, 22]
HF	96.2	137.7	131.5	135.0	96.3	138.6	132.0	135.4	96.2	140.0	133.4	136.7	141.2
H$_2$O	154.5	221.5	215.3	219.5	154.7	222.5	215.6	219.7	154.6	225.3	218.6	222.6	232.2
OH	67.4	99.2	97.7	99.3	67.7	100.0	98.2	99.6	67.8	101.6	100.0	101.4	106.6, 106.6 ± .3
NH$_3$	198.5	275.5	273.9	277.3	199.2	277.6	275.2	278.6	199.5	281.8	280.1	283.7	297.3 ± .1
NH$_2$	115.1	163.8	164.6	166.5	116.3	165.6	165.9	167.8	116.6	168.8	169.7	171.7	180.9 ± 1.5
NH	49.8	72.2	73.5	74.4	50.6	73.1	74.1	75.1	50.8	74.8	76.3	77.3	78.8 ± 4, ⩽ 84.7
CH$_4$	326.5	398.0	400.7	402.0	326.9	400.4	402.5	404.1	327.0	403.3	406.0	407.8	419.8 ± .1
CH$_3$	241.4	290.6	292.7	293.2	241.8	292.2	294.0	294.5	242.0	294.9	297.2	298.0	307.2 ± .2
CH$_2$	154.0	181.1	181.7	181.5	154.2	181.8	182.2	181.9	154.3	183.4	184.4	184.3	189.8 ± .6
CH	55.8	74.2	76.3	76.9	56.4	75.1	77.1	77.7	56.5	76.4	78.7	79.4	84.4 ± .1, 83.9
BH$_3$	230.5	267.1	270.4	270.1	231.1	268.5	271.8	271.5	231.1	270.3	274.1	274.0	278.1 ± 10
BH$_2$	142.8	162.5	163.5	162.6	143.2	163.5	164.4	163.6	143.3	164.6	166.0	165.4	196.2 ± 15
BH	61.5	77.1	79.4	79.9	61.6	77.8	80.1	80.5	61.6	78.5	81.0	81.5	81.6 ± 2, 82.2
BeH$_2$	124.9	141.1	142.6	141.8	125.3	142.0	143.5	142.8	125.3	142.8	145.0	144.5	
BeH	50.1	51.2	49.5	48.0	50.3	51.6	50.0	48.6	50.3	51.9	50.8	49.6	49.8, 55.8 ± 7
LiH	33.8	48.1	51.5	52.6	33.9	48.5	52.0	53.0	33.9	49.3	53.1	54.2	58.0
H$_2$	83.4	100.8	104.5	105.4	83.4	100.8	104.5	105.4	83.4	100.8	104.5	105.4	109.5

6-311 + G($2df, p$) basis sets differ from those given in [22] in that they were computed directly, whereas an additivity approximation was used in [22]. Noticeable nonadditivity errors occur only for the 6-311 + G($2df, p$) basis set; they are largest in ammonia (1.5 kcal/mol), methane (1.2 kcal/mol), methyl radical (0.9 kcal/mol), and water (0.8 kcal/mol). All other nonadditivity errors are less than or equal to 0.5 kcal/mol.

First and foremost, it is obvious that the basis set superposition error globally increases with increasing atom number of the central atom, both with and without bond functions: this was also seen in Part 2. The largest BSSEs in this study are seen for the water molecule, becoming an alarming 12.00 kcal/mol with the 6-311 + G(d, p)B basis set.

The BSSEs also greatly increase upon inclusion of electron correlation, except of course for H_2 and LiH, where the separated atoms have no correlation energy in the frozen-core approximation adopted here. For the lighter hydrides the BSSEs also exhibit a marked decrease from MP2 to MP3, much less from MP3 to MP4; for the heaviest hydrides there is a slight increase. The decrease is more marked for the larger than for the smaller basis sets: the difference between BSSEs with and without bond functions is also affected, and decreases for all cases with the greatest basis set employed. It was generally observed that BSSEs decrease upon proceeding from MP3 to MP4(DQ), whereas they again increase upon proceeding from MP4(DQ) to MP4(SDTQ). Extrapolating for higher-order effects according to the formula suggested in Part 4 resulted in BSSE changes of only about 0.01 kcal/mol, which enables the authors to say that, at least for this type of molecules, BSSE is adequately described at the MP4 level.

BSSEs are more than doubled when bond functions are added: this increase is somewhat less spectacular than expected from the work of Bauschlicher. It is remarked here, that the MP4/6-311 + G(d, p) BSSE still amounts to 5.26 kcal/mol for H_2O, which although less than half the 12.00 kcal/mol found when bond functions are included, is still anything but desirable. The largest MP4/6-311 + G($2df, p$) BSSE is 1.74 kcal/mol for H_2O, where only 2.86 kcal/mol is added on top by using the bond functions; the difference in binding energy amounts to 10.0 kcal/mol for the smaller, and 5.4 kcal/mol for the larger basis set. For NH_3, the picture is more clear-cut: 4.58 kcal/mol extra BSSE results from addition of bond functions to the smaller basis set, and 2.11 form addition to the larger basis set. This compares with the large differences in binding energy of 12.3 and 6.7 kcal/mol, respectively. For CH_4, the extra BSSE is 1.53 kcal/mol with the largest basis set, and 3.09 with the smallest basis set; this should be compared to extra binding energies of 6.5 and 12.1 kcal/mol, respectively.

If the BSSE differences caused by the bond functions are studied, it becomes clear that they are not markedly affected by the addition of a second set of d functions; they are, on the other hand, from the boron hydrides on, decreased by a factor of two or more when a set of f functions is added. For H_2, LiH and the beryllium hydrides, the d functions are already a second polarization space for the separated atoms, so f functions (which are then third polarization space) do not contribute appreciably.

For the BSSE without bond functions however, doubling the d functions decreases the BSSE more significantly than the subsequent addition of a set of f functions. The total BSSEs when using bond functions are decreased by a factor of about three when extra d and f functions are added at the same time. The final values are a bit large in some instances, but still acceptable when compared with the other errors inherent in the calculations.

Considering now the difference between the three methods, Wells and Wilson expected the PAFC to be larger than the SSFC, and thus overestimate the basis set superposition error. From Tables 1 and 2, it is quickly seen that PAFC may both over- and underestimate the SSFC BSSE. Additivity holds better for the UHF than for the correlated levels. The additivity also improves with the basis set: at the UHF/6-31 + G($2d, p$) and ($2df, p$)B levels, where the BSSE is also smallest, additivity is best. The SRCP, on average, tends to yield the highest BSSEs of all three methods, except for the case of H_2O, where it is in between the SSFC and PAFC results. Generally, the agreement between SSFC and SRCP appears to be better than either that between SSFC and PAFC, or between SRCP and PAFC.

Wells and Wilson also observed that additivity is strongly influenced by the symmetry of the system. This is particularly noticeable for CH_3 and CH_4; however, a tendency towards lowered BSSEs due to nonadditivity is seen in all cases where the point group of the adduct is of a more symmetric type than that of its lower hydride monomer (e.g., CH_3, BH_3: $C_{2v} \rightarrow D_{3h}$, CH_4: $D_{3h} \rightarrow T_d$, but not NH_3: $C_{2v} \rightarrow C_{3v}$). (Also, the wavefunction for CG_{H_4} (i.e., carbon atom in the basis set for methane) has only C_{3v} symmetry; if it were constrained to T_d symmetry (which is unfortunately impossible with GAUSSIAN 86), the BSSE would be slightly lower in the CH_4 case.) The same argument applies for BeH_2 ($C_{\infty v} \rightarrow D_{\infty h}$); the effect is only seen here when bond functions are included. The SRCP method, on the other hand, does not exhibit this phenomenon (except for BH_3 and CH_3); apparently some BSSE might be involved in the geometrical rearrangement.

The bond function contributions are also strongly nonadditive, reaching 1.56 kcal/mol for H_2O in the 6-311 + G(d, p)B basis set. Here, an overestimate is generally observed, except for H_2O, where PAFC underestimates the SSFC result.

It was found in Parts 1 and 2 that computed dissociation energies using the 6-31 + G(d, p)B and 6-311 + G(d, p)B basis sets and an isogyric reaction cycle involving molecular hydrogen [22] were too large by amounts up to 2 kcal/mol at the MP4 level when using bond functions; the overshoot was somewhat smaller with the 6-311 + G($2df, p$)B basis set. It would be natural to ascribe this effect to the basis set superposition error (which is of course reduced in the 6-311 + G($2df, p$)B basis set). However, in Part 4 it was found that a fair part of the overestimate was not due to BSSE, but to truncation of the Møller–Plesset series: when an augmented coupled cluster method was used instead, mean absolute errors were reduced to 0.42 kcal/mol for the 6-311 + G($2df, p$)B basis set, and 0.68 kcal/mol for 6-311 + G(d, p)B. The tendency towards overestimation remained, while it should also be noted that the use of optimum geometries

at the levels involved would further increase the computed dissociation energy by at most 0.3 kcal/mol [22]. This produces a net overestimation of 1 kcal/mol or less, which is definitely small if compared to the sometimes overly large BSSE values found by all three methods here. This is hard to explain from the viewpoint that the counterpoise method produces an accurate description of basis set superposition error.

The only possible explanation would be that BSSE and basis set incompleteness error follow the same trend, and thus compensate each other. Since both the basis set incompleteness error and the BSSE increase from left to right in periodic table, this seems plausible, at least for the case of diatomic molecules; in the case of polyatomic molecules however, this argument becomes more difficult to sustain.

In the light of Mayer's conclusion, however, that BSSE and interaction energy are generally nonadditive, our findings are readily understood. As far as the additive part of the BSSE is concerned, the above argument may still apply; the nonadditive part, of course, is also found in the true interaction energy. Since it is hardly desirable to compensate for the latter, and the former error wil cancel to a great extent with the basis set incompleteness error, it may be best *not* to compensate at all for BSSE in dissociation energy calculations using bond functions, until perhaps a correlated extension of Mayer's BSSE-free method becomes generally available. The most powerful argument supporting this conclusion is and remains the excellent agreement between calculated and experimental values using this approximation, since in the end, that is what quantum chemistry is all about. Also, it is by no means general practice to include BSSEs in dissociation energy calculations without bond functions. As is seen above, BSSEs according to all three generalizations of the counterpoise method are also anything but small in that case: so there hardly is any good reason either for dismissing bond function basis sets as BSSE-prone, or for advocating that BSSE corrections are more appropriate for calculations with bond functions than for calculations without them. Ultimately, in the opinion of the authors, the chief usefulness of counterpoise-related methods in dissociation energy calculations as a qualitative indicator for basis set incompleteness.

Finally, however, the authors wish to stress that these conclusions do not necessarily apply for the case of weak molecular interactions, where bond functions, with their inherent reliance on a more or less directional covalent-type bonding, would be of limited use anyway.

Note added in proof

It was recently shown (Martin, François, Gijbels, Chem. Phys. Lett., submitted) that, when correlated (CID/6-31G*) geometries are used, most of the remaining overestimate in the dissociation

energies is absorbed through the isogyric reactions. Using the CCD + ST(CCD)/6-311 + G($2df$, p)B// CID/6-31G* theoretical model and an isogyric reaction cycle [22], a mean absolute error of 0.12 kcal/mol per bond could be reached for the first-row hydrides with precisely known experimental dissociation energies. This further supports our conclusion that neglecting BSSE has no adverse effects on the quality of the results.

References

1. Wright JS, Williams RJ (1983) L Chem Phys 79:2893–2902; Wright JS, Donaldson DJ, Williams RJ (1984) J Chem Phys 81:397–406
2. Wright JS, Buenker RJ (1984) Chem Phys Lett 106:570–574; Wright JS, Williams RJ (1983) J Chem Phys 78:5264–5266
3. Martin JML, François JP, Gijbels R (1989) J Comp Chem 10:152–162
4. Martin JML, François JP, Gijbels R (1989) J Comp Chem 10: in press
5. Martin JML, François JP, Gijbels R (1989) Chem Phys Lett 157:217–223
6. Bauschlicher CW (1980) CI superposition error caused by bond functions, ICASE Report No. 80-9, NASA Langley Research Center, Hampton, VA; Bauschlicher CW (1985) Chem Phys Lett 122:572–576
7. Boys SF, Bernardi F (1970) Mol Phys 19:533–566
8. Hobza P, Zahradnik R (1988) Chem Rev 88:871–897 and references therein; Van Lenthe JH, van Duijneveldt-van de Rijdt JGCM, van Duijneveldt FB (1987) Adv Chem Phys 69:521–566
9. Schwenke DW, Truhlar DG (1985) J Chem Phys 82:2418–2426; (1986) 84:4113(E); Frisch MJ, Del Bene JE, Binkley JS, Schaefer III HF (1986) J Chem Phys 84:2279–2289
10. Mayer I, Vibok A (1987) Chem Phys Lett 140:558–564
11. Mayer I (1983) Int J Quantum Chem 23:341–363
    Mayer I, Vibok A (1987) Chem Phys Lett 136:115–121
12. Loushin SK, Liu S, Dykstra CE (1986) J Chem Phys 84:2720–2725; Loushin SK, Dykstra CE (1987) J Comp Chem 8:81–83. The closely related subject of "secondary basis set superposition error" is discussed in: Karlström G, Sadlej AJ (1982) Theor Chim Acta 61:1–9; Latajka Z, Scheiner S (1987) J Chem Phys 87:1194–1204
13. Mayer I (1987) Theor Chim Acta 72:207–210
14. Wells BH, Wilson S (1983) Chem Phys Lett 101:429–434
15. Waite J, Papadopoulos MG (1989) Theor Chim Acta 75:53–65
16. Frisch MJ, Pople JA, Binkley JS (1984) J Chem Phys 80:3265–3269 and references therein
17. Frisch MJ, Krishnan R, Pople JA (1980) J Chem Phys 72:4244–4245 and references therein
18. Binkley JS, Frisch MJ, Raghavachari K, DeFrees DJ, Schlegel HB, Whiteside RA, Fluder EM, Seeger R, Fox DJ, Head-Gordon M, Topiol S (1987) GAUSSIAN 86 release C, Carnegie-Mellon University, Pittsburgh, PA
19. Defrees DJ, Levi BA, Pollack SK, Hehre WJ, Binkley JS, Pople JA (1979) J Am Chem Soc 101:4085–4089; (1980) 102:2513(E)
    DeFrees DJ, Raghavachari K, Schlegel HB, Pople JA (1982) J Am Chem Soc 104:5576–5580
20. Pulay P (1982) J Comp Chem 3:556–560
21. Seeger R, Pople JA (1976) J Chem Phys 65:265–271
22. Pople JA, Frisch MJ, Luke BT, BInkley JS (1983) Int J Quantum Chem Symp 17:307–320