9 July 1999

ELSEVIER

Chemical Physics Letters 307 (1999) 469–478

CHEMICAL
PHYSICS
LETTERS

www.elsevier.nl/locate/cplett

New development in RECEP (rapid estimation of correlation energy from partial charges) method

Sándor Kristyán [a,*], Gábor I. Csonka [b,1]

[a] Department of Radiation Oncology, Emory University, 145 Edgewood Avenue, SE, Atlanta, GA 30335, USA
[b] Department of Inorganic Chemistry, Technical University of Budapest, H-1521 Budapest, Hungary

Received 11 March 1999; in final form 12 May 1999

Abstract

We analyze the feasibility of extremely rapid estimation of correlation energy from the HF-SCF charge distribution in closed-shell molecules. In Kristyan's previous work [Chem. Phys. 224 (1997) 33-51] a simple linear relationship using atomic correlations was developed in order to calculate correlation energy of molecules. This method has been further refined in this Letter. The proposed method is analyzed for 18 molecules and ions and the new results are compared to the B3LYP, CCSD and G2 results. © 1999 Elsevier Science B.V. All rights reserved.

1. Introduction

The efficient calculation of correlation energy plays a central role in modern theoretical chemistry [1–4]. The exact solution of the Schrödinger equation is provided by the full configuration interaction (CI) method for some lowest lying states, and the approximate Hartree–Fock self-consistent field (HF–SCF) for the ground state energy. The very accurate full-CI method is practical only for very small molecules due to its high computing power demand [4,5] and factorial scaling with the system size. The HF–SCF calculations [4–8] for the ground state require much less computing power and scale at best linearly with the system size. However, the inherent approximations bear a systematic error in comparison to the accurate full-CI method. The term 'electron correlation energy' (E_{corr}) for the electronic ground state is usually defined as the difference between the exact non-relativistic full-CI energy of the system (E_0(CI)) and the HF–SCF basis set limit energy (E_0(HF–SCF)):

$$E_{corr} = E_0(\mathrm{CI}) - E_0(\mathrm{HF{-}SCF}) \ . \quad (1)$$

By the variational property of the HF–SCF approximation $E_{corr} < 0$ for systems with $N \geqslant 2$. The HF–SCF method provides an excellent starting point, which accounts for about 99% of the total energy of the molecule. However, the neglect of correlations between electrons (mean-field approximation), leads to rather poor description in the chemical sense because the energy of the chemical bonds is comparable to that 1% error of the HF–SCF method. This

---

[*] Corresponding author. Fax: +1 404 616 5689; e-mail: skrist2@emory.edu or ekristy@emory.edu

[1] Also corresponding author. Fax: +36 1 463 3642; e-mail: csonka@web.inc.bme.hu

PII: S0009-2614(99)00544-8

is why the inclusion of the electron correlation is critical for the accurate and quantitative evaluation of molecular energies. Yet, HF–SCF program packages are much simpler to handle than CI program packages, so correlation calculation methods are sought which require much less computational demands than the CI, but reach CI accuracy.

According to Eq. (1) the electron correlation effects are a measure of the systematic errors that are inherent in HF theory. However, one should keep in mind that while HF theory is well-defined and unique for closed-shell molecules, several versions of HF theory are used for open-shell molecules. Correlation energy for an open-shell molecule is usually defined with respect to unrestricted Hartree–Fock (UHF) theory where the spatial orbitals are different for α and β spins. However, some authors prefer to define it with respect to restricted Hartree–Fock (RHF) theory where the spatial orbitals for α and β spins are identical. Sometimes, it may be convenient to replace the HF approximation in the above definition with another well-defined approximation such as a multiconfiguration reference function. In the present Letter we study only closed-shell molecules and use RHF theory for such molecules.

Simple analysis of the nature of correlation energy (E_{corr}) leads to a frequently used rule of thumb: that correlation effects contribute about 23 kcal/mol for a pair of electrons in a well-localized orbital (cf. to the expected 1 kcal/mol chemical accuracy of quantum chemical calculations). Hartree–Fock calculations with large basis sets show that correlation effects contribute about 25 kcal/mol to the binding energy in H_2. For many pairs of electrons in close proximity, correlation effects become very large. For example, they contribute more than 100 kcal/mol to the bond energy in N_2. There exists a clear dependence between the magnitude of the correlation energy and the number of electrons. In this sense the electrons with antiparallel and parallel spins must be clearly differentiated because the HF method provides a good description for the electrons with parallel spins (Pauli exclusion) and poorer (average) description for the antiparallel spins. Although E_{corr} depends [9–11] mainly (quasi-linearly) on the number of electrons, N (more precisely on $N-1$), the considerably smaller dependence on the nuclear potential, $v(\mathbf{R},\mathbf{Z})$, must be considered in order to obtain chemical accuracy (with $(\mathbf{R},\mathbf{Z})$ defining the nuclear frame). As a consequence, the HF–SCF method itself is not accurate even for energy differences (e.g. in the case of frequency or activation energy calculations), because this systematic error (E_{corr}) does not cancel.

Some well-known, very expensive classical approximations for correlation energy are the second, third or fourth order Møller–Plesset perturbation methods (MP2, MP3 or MP4) [6–8,12], and the still more expensive coupled cluster methods (CCSD, or CCSD(T)) [8,13]. The applied basis set is also a source of various problems, as the MPx and CCSD methods show a rather slow convergence with respect to the basis set increase. Even divergence is not uncommon. In this respect the HF–SCF method shows a more advantageous behavior. The very poor convergence and the $O(n^5) - O(n^7)$ scaling [14] with the basis set increase led to the development of the so-called composed (or extrapolation) methods (e.g. G2) [8,15]. Considerably faster alternatives for the estimation of the correlation energy are the density functional (DFT) methods. The recent correlation functionals, e.g. Perdew–Wang and Lee–Yang–Parr (LYP) [16,17], or the hybrid adiabatic connection methods (ACM) e.g. B3PW91 or B3LYP [18–22]), and many other models [6–8,23–27], provide usually rather good results, and show a basis set convergence similar to HF–SCF method. However, it should be noted that the DFT methods also introduce new problems. There is no simple way to improve the results (cf. use of semi-empirical functionals), and numerical instabilities might occur as well.

We used the GAUSSIAN 94 program [8], and selected 6-311 + G(2d,p) as default basis set in different methods. We have performed HF–SCF, full and frozen core CCSD(T), G2 and B3LYP calculations. The natural population analysis (NPA) charges [8] were calculated via the HF–SCF method using the G2 geometries; the geometries were reoptimized in the B3LYP calculations. For various estimations of the molecular correlation energy we use the following notation: E_{corr}(CCSD), E_{corr}(G2), E_{corr}(B3LYP), E_{corr}(RECEP), while E_{corr} itself notes the accurate complete-CI value as defined in Eq. (1). For atoms we note the number of electrons (N), and atomic charge (Z) in the argument as well, e.g. $E_{\mathrm{corr}}(\mathrm{CI}, N, Z)$, and sometimes the spin states.

2. Rapid estimation of correlation energy from partial charges (RECEP) method

The RECEP method (introduced first in Ref. [11]) is a very remarkable estimation for correlation energy. It is based on the following ideas:

Empirically, the correlation energy of any N electronic systems can be roughly approximated [9–11] by the following quasi-linear relationship:

$$E_{\text{corr}} \approx a(N-1)\ , \tag{2}$$

where $-0.035 \geqslant a[\text{hartree}] \geqslant -0.045$.

The physical origin of $N-1$ in Eq. (2) is that the electron correlation increases nearly proportionally by the number of $N-1$ electrons around a selected electron [11]. The CI results for atoms and positively charged atomic ions ($1 \leqslant N \leqslant Z \leqslant 18$) and some A^- atomic anions in free space reported in Refs. [28–30] fall within the limits given by Eq. (2). The B3LYP results in Refs. [9–11] for neutral and positively charged closed-shell atoms show a similar agreement with Eq. (2). However, the quasi-linear relationship in Eq. (2) cannot reach the chemical accuracy (1 kcal/mol, about 0.0016 hartree/particle). This rough approximation does not make a difference between α and β spin electrons and does not deal with the polarization of electron cloud in molecular bond. The estimation of correlation energy for molecules by Eq. (2) misses the ($\mathbf{R}$,$\mathbf{Z}$) nuclear frame dependence; thus it adds a constant to the entire HF–SCF potential surface. For this latter reason, this approximation is certainly insufficient for the molecular hypersurfaces.

The RECEP method can be constructed from the basic idea that the electron correlation energy is roughly proportional to the number of electrons. The essence of the RECEP is the use of partial atomic charges, an integrated atomic quantity to estimate the number of electrons around the atoms. It is supposed that the quasi-linear dependence of the correlation energy on the (fractional) number of electrons is conserved in the molecules. The E_{corr}(RECEP) is calculated as follows: ‘par(A)’ denotes the partial charge on atom A in the molecule consisting of M atoms ($A = 1,2,3,\ldots,M$) in atomic units (i.e. -1 for electron, etc.). The molecular charge, Q, is simply $Q \equiv \Sigma_{(A=1,\text{M})} Z_A - N = \Sigma_{(A=1,\text{M})}\text{par}(A)$, where Z_A is the nuclear charge on the A-th nucleus and N is the total number of electrons. The N_A is defined as $N_A \equiv Z_A - \text{par}(A)$ and its ‘meaning’ is the electron ‘content’ around the nucleus A in the molecule. The different partial charges will be discussed later. The partial charges offer an amazing estimation for the correlation energy, considerably better than Eq. (2) [11]. The formula for RECEP to estimate the molecular E_{corr} value is the following:

$$E_{\text{corr}}(\text{RECEP}) \equiv \Sigma_{(A=1,M)} E_{\text{corr}}(N_A, Z_A)\ . \tag{3}$$

In Eq. (3), N_A is the electron content around atom A, calculated from partial charge par(A). The N_A values in Eq. (3) are not necessarily integer numbers, while the Z_As are. The summation in Eq. (3) runs for all atoms in the molecule. The two basic assumptions of Eq. (3) can be summarized as follows: the correlation energy is the sum of the atomic correlation energies and the atomic correlation energies can be estimated from the partial charges.

The proof of the first theorem is quite simple. The gradient vector field analysis of the electron density ($\rho(\boldsymbol{r})$) provides that the molecules can be cut apart to virial atoms [31]. The zero flux surfaces of the electron density give the borders for these virial atoms and in these atomic volumes the virial theorem is fulfilled. The total energy can be calculated from the sum of the virial atomic energies. If we introduce the $\rho(\boldsymbol{r})$ from HF–SCF and full-CI calculations and calculate the virial atomic energies respectively, the difference will provide the atomic correlation energies, thus the existence of such energies is proved for those cases. One technical difficulty is that the full-CI electron density is not generally known. Another more serious theoretical problem is that it is not impossible to imagine examples that contain domains without nuclei, thus assigning electrons to interatomic volumes; this warns us that the above-mentioned procedure is not always applicable. However, it is applicable for most of the known organic molecules and interesting results would probably arise from such kind of studies for atomic correlation energies in various molecules.

The proof of the second assumption is very difficult and we shall use it as a work hypothesis. This assumption is the extension of the N dependence of the atomic correlation energy for the non-integer partial charges. The essential problem: what is the definition of partial charges? The reader must be aware of the fact that partial charges are mathematical constructions that may help chemists to establish empirical rules. Indeed, the partial charges might differ substantially depending on the mathematical formula. Different methods contradict frequently each other or lead to 'unphysical' numbers in the so-called difficult cases. We recall here only four atomic charge definitions. First we mention the widely used Mulliken charges [2,8]. The deficiencies of this method are well-known: for example the Mulliken charges are oscillating with respect to the basis set increase and do not show convergent behavior. The charges derived from the electrostatic potential (ESP) [7,32] and from natural population analysis (NPA) [8,33] show considerably more stability with respect the increase of the basis set. Finally we mention the charges derived from Bader's population analysis [31]. This latter definition of partial charges is derived directly from the integration of $\rho(\boldsymbol{r})$ in the domains of virial atoms. While the derivation of the charges is well-defined in the theoretical sense, the derivation of such charges might suffer from serious numerical instabilities, and as mentioned above, there exist examples where charges are assigned to spatial domains without nuclei. We also note that Bader's population analysis leads (for example) to negative charges on hydrogen atoms of hydrocarbons while the NPA or ESP charges give approximately the same positive values on these hydrogen atoms – and in this way they can refer e.g. about dipol moments. We also note that the shapes of virial atoms in molecules are frequently quite different from the known spherical symmetry of free atoms in free space.

The use of the HF–SCF $\rho(\boldsymbol{r})$ calculated with a moderate basis set (e.g. 6-31G(d)) can be justified readily. For example, introducing HF–SCF $\rho(\boldsymbol{r})$ into a correlation energy functional (e.g. LYP) provides more than sufficiently adequate correlation energy [9,10,23]. Moreover, the value of E_{corr}(LYP) change only a little if we introduce HF–SCF limit electron density. Similar observations have been made for the LSDA electron density [18–23], which is successfully used for correlation energy calculations in gradient corrected functionals. This observation originates in the fact that the E_{corr} is an integrated quantity with respect to the $\rho(\boldsymbol{r})$; thus more accurate electron density causes little change in the E_{corr} (the one-electron density integrates to the number of electrons) [9,10].

The next question is how to select the values for the $E_{\mathrm{corr}}(N_A,Z_A)$ terms in Eq. (3). These terms can be calculated from accurate CI calculations comparing the correlation energy of the cation, neutral atom and anion, because in covalent bonds (for example the NPA) partial charges are generally between -1 and $+1$ in a neutral molecule. The CI data for integer (N_A,Z_A) pairs can be found in Refs. [28–30]. For the non-integer N_A values linear interpolation [11] might provide a reasonable correlation energy value. Probably the linear interpolation could be replaced by an interpolating function; however, this question will be investigated in future papers. It should be noted that these CI values are not the best choice for Eq. (3). (cf. appendix of Ref. [11]). The reason for this is the following: the atoms change their spin state in molecules, using the open-shell high spin multiplett correlation energies which would provide some bias. For example, in case of CH_4, the partial charge on C is between 0 and -1, i.e. the N_A for the C atom is between 6 and 7. The accurate CI calculations in Refs. [28–30] provide the correlation energies for high spin states, i.e. $E_{\mathrm{corr}}(CI,\ N_A=6,\ Z_A=6$, triplet, i.e. $1s^2 2s^2 2p_x 2p_y)$ and $E_{\mathrm{corr}}(CI,\ N_A=7,\ Z_A=6$, quartet, i.e. $1s^2 2s^2 2p_x 2p_y 2p_z)$ values for a linear interpolation. However, in a closed-shell (at equilibrium or close to equilibrium) methane there is no unpaired electron around the carbon atom; the molecule is singlet. As was mentioned before, the correlation energy is very sensitive to spin pairing effects because the opposite spin electrons have different correlation energy than the parallel spin electrons. For example the exact CI correlation energy changes only by -3 millihartrees (mh) between He and Li and it changes by a considerably larger extent between Li and Be (-49 mh) as the spin pairing occurs (cf. Fig. 1). The spin pairing effect can be easily followed on Fig. 1.

Thus instead of using the correlation energy of high spin atomic states we propose using the energy of the excited or low spin states, e.g. $E_{\mathrm{corr}}(\mathrm{CI},\ N_A=6,\ Z_A=6$, singlet, i.e. $1s^2 2s^2 2p_x^2)$ and $E_{\mathrm{corr}}(\mathrm{CI},\ N_A=6,\ Z_A=7$, doublet, i.e. $1s^2 2s^2 2p_x^2 2p_y)$. (It is seemingly contradictory to use excited state parameters – i.e. correlation of low spin atomic states – in approximating ground state molecules with Eq. (3).

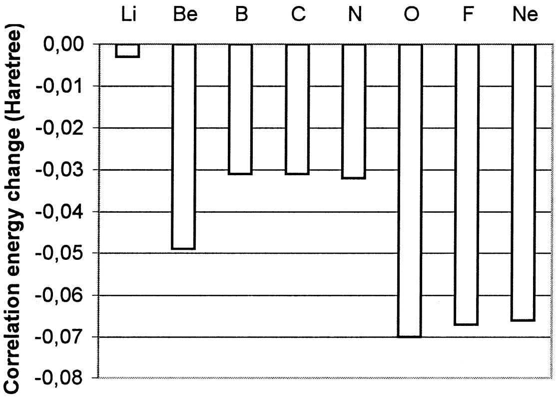

Fig. 1. The exact correlation energy differences between the neighboring elements going from left to right in the first row of the periodic table (for example Li shows the correlation energy difference between He and Li, Be shows the correlation energy difference between Li and Be, etc.). The plot is based on footnote b values of Table 1. This chart shows the important spin pairing effect in correlation energy (e.g. the pairing in 2s shell in Be, and in 2p shell from the oxygen atom (cf. Hund's rule) is manifesting).

But we recall Eq. (2) again, s.t. the number of electrons counts in the main event, and that in closed-shell molecules at ground state the spin pairing is maximal. We can also analogously mention that the CI method is superior to the HF–SCF method in approximating ground state E_0 value, because 'Slaterian excited states' are also involved [2]. And after all, our computational experience [11] strongly suggests that low spin atomic $E_{\mathrm{corr}}(N_A, Z_A)$ values give better results via Eq. (3) than high spin values.) This method we used in this Letter and was used in Ref. [11]. Correlation energies were calculated for the first row atoms and ions, forcing them to be in singlet or doublet states – where necessary, because accurate CI calculations of this kind are not available yet. (For B3LYP correlation energies the $E_{\mathrm{corr}}(\mathrm{B3LYP}) \equiv [E_0(\text{HF–SCF and B3}) + E_{\mathrm{corr}}(\mathrm{LYP})] - E_0(\text{HF–SCF})$ definition was used [11].[2]) We compared the high spin and low spin B3LYP correlation energies and obtained a correction factor for 6, 7 and 8 electronic systems: for 6, 7, and 8 electronic systems the correlation energy changes approximately by -0.019, -0.026, and -0.026 hartree, respectively. The exact CI values were corrected with these values for the 6, 7, and 8 electronic systems shown in Table 1 (footnote c). Table 1 also shows the corresponding low spin B3LYP atomic correlation energies. Similar values can be calculated for the whole periodic system, and other methods of correlation energy calculation can be used instead of full-CI or B3LYP (e.g. CCSD(T)). In this Letter we test two sets of atomic correlation energy values: RECEP-c will note the calculation for estimating molecular correlation energies calculated by Eq. (3) using NPA [8] charges and corrected CI atomic correlation energies for $E_{\mathrm{corr}}(N_A, Z_A)$, the values under footnote c in Table 1, and RECEP-d will note the calculation if values under footnote d in Table 1 are used for $E_{\mathrm{corr}}(N_A, Z_A)$ instead. Our experience is that using different correlation calculations for atomic $E_{\mathrm{corr}}(N_A, Z_A)$ causes only secondary effect in Eq. (3) (mainly for energy differences) due to the fact that the sum of partial charges is always exactly the molecular charge itself.

Further refined atomic correlation energies, $E_{\mathrm{corr}}(N_A, Z_A, \text{singlet})$ values (or such fictitious singlets as those for a neutral nitrogen atom [11], where singlet does not exist in atomic state, only doublet or quartet) could be obtained from known correlation energies of a larger set of smaller molecules (e.g. the molecule set used in G2 method) by a fit of Eq. (3) for its $E_{\mathrm{corr}}(N_A, Z_A, \text{singlet})$ parameters; and these fitted parameters could be used in the same way as (for example) the molecular mechanic type calculations use their force constants, etc. Clearly, these fitted $E_{\mathrm{corr}}(N_A, Z_A, \text{singlet})$ values, like the values under footnote c in Table 1, are no longer individual atomic parameters. These are atomic parameters in molecular bond, however; these $E_{\mathrm{corr}}(N_A, Z_A)$ values are very close to atomic singlet correlation energy values. Furthermore, it may eliminate the dilemma of the use of UHF or RHF correlation energy. On the other hand, the values under footnotes b and d in Table 1 are atomic parameters in free space.

The hydrogen atoms require special attention because their partial charges fall frequently between 0

---

[2] The $E_{\mathrm{corr}}(\mathrm{LYP})$ DFT functional [16,17] is proposed for Kohn–Sham orbitals, not for Slaterian orbitals, i.e. not for approximating E_{corr} in Eq. (1). But interestingly, if $E_{\mathrm{corr}}(\mathrm{LYP})$ is used simply for Eq. (1), it gives reasonable result (SK). This was not emphasized enough in Refs. [9,10] when it was compared to $E_{\mathrm{corr}}(\mathrm{MP2})$, although was recognized; Refs. [9,10] state only that $E_{\mathrm{corr}}(\mathrm{LYP})$ is much less sensitive for increasing basis set than e.g. $E_0(\text{HF–SCF})$.

Table 1
Atomic correlation energies (hartree) to use in Eq. (3)

Atom[a,b]	N	Z	Corrected exact[c]	B3LYP[d]
H	2	1	−0.0395	−0.0432
He	2	2	−0.0420	−0.0531
Li	2	3	−0.0435	−0.0491
	3	3	−0.0453	−0.0593
Be	2	4	−0.0443	−0.0425
	3	4	−0.0474	−0.0601
	4	4	−0.0943	−0.0994
B	4	5	−0.1113	−0.1060
	5	5	−0.1249	−0.1316
	6	5	−0.1640	−0.1765
C	5	6	−0.1388	−0.1400
	6	6	−0.1754	−0.1911
	7	6	−0.2087	−0.2258
	8	6	−0.2839	−0.2883
N	6	7	−0.1856	−0.2005
	7	7	−0.2143	−0.2373
	8	7	−0.2877	−0.3035
	9	7	−0.3314	−0.3622
O	7	8	−0.2202	−0.2445
	8	8	−0.2839	−0.3079
	9	8	−0.3314	−0.3619
	10	8	−0.4080	−0.4513
F	9	9	−0.3245	−0.3599
	10	9	−0.3995	−0.4430
Ne	10	10	−0.3905	−0.4338

[a] N: number of electrons. Z: nuclear charge.

[b] CI correlation energies for high spin ground states can be found in Refs. [28–30].

[c] Corrected (or optimized) CI correlation energies: −0.019 hartree correction for 6 electronic systems, −0.026 hartree correction for 7 and 8 electronic systems with respect to footnote b (RHF) values. These values no longer refer to the individual atomic ground states in free space; we suggest they be applied to Eq. (3) for $E_{\mathrm{corr}}(N_A,Z_A)$ to estimate molecular correlation energies. The calculated molecular correlation energies in this way are noted as RECEP-c in the text. These atomic parameters are probably close to the low spin state CI values, not yet available in the literature.

[d] Atomic correlation energies for singlet or doublet (not necessarily ground state) systems calculated as a difference of B3LYP/6-311 + G(3df,2pd) and HF/6-311 + G(3df,2pd) energies. (i.e. atomic spin states are as low as possible, e.g. the carbon is in $1s^2 2s^2 2p_x^2$ singlet low spin state, etc.) These values are also suggested for use in Eq. (3) for $E_{\mathrm{corr}}(N_A,Z_A)$ to estimate molecular correlation energies. The calculated molecular correlation energies in this way are noted as RECEP-d in the text.

[c,d] In case of C, N and O atoms, seemingly A^{-2} atomic anions are listed when $N = Z + 2$, however, these do not exist in gas phase in free space [11], although ab initio packages [8] can calculate their correlation energies. Indeed, these values are necessary in the calculation e.g. for some molecules listed in Tables 2 and 3. One should take these values as parameters for Eq. (3) instead as correlation energies.

and 1.0. In Ref. [11] the $E_{\mathrm{corr}}(0 \leqslant N_A < 1.0,\ Z_A = 1) = 0$ was used in Eq. (3) for these cases, and for $1 \leqslant N_A < 2$ the $E_{\mathrm{corr}}(N_A = 1.0,\ Z_A = 1) = 0$ and $E_{\mathrm{corr}}(N_A = 2.0,\ Z_A = 1) < 0$ values were used (hydrogen atom and H^- anion, respectively) in the linear interpolation for a negatively charged hydrogen atom in the molecule in question. In this way Eq. (3) yielded zero correlation for the H_2 molecule, which was essentially incorrect [11]. Here, we calculate the correlation energy for H atoms as follows: for N_A between 0 and 2, we linearly extrapolate the $E_{\mathrm{corr}}(N_A,Z_A)$ value from the two, $E_{\mathrm{corr}}(N_A = 0,\ Z_A = 1) = 0$ and $E_{\mathrm{corr}}(\text{B3LYP [8]},\ N_A = 2,\ Z_A = 1, \text{singlet}) = -0.043200$ hartree values via the actual N_A value in the molecule. In this way Eq. (3) yields reasonable correlation energy for the hydrogen molecule as well. This method is important not only for the hydrogen molecule, but for many other hydrogen-containing molecules, because in these molecules the electrons are correlated around the hydrogen atoms as well. This was also incorporated in RECEP-c and RECEP-d calculations.

Eq. (3) deals neither with the spin distribution of electrons (e.g. the LYP estimation does for the two cases: open ($N =$ odd) and closed ($N =$ even) system) nor the polarization of electron cloud, but this causes only little error. (More precisely, these properties pass somehow via the partial charges from the HF–SCF calculation.) Eq. (3) is very likely not 'variational estimation' for the correlation energy of the molecules. However, because the number of electrons around the atoms in a neutral molecule does not differ 'significantly' from the atomic states of the participating atoms, Eq. (3) should not be far from a possible 'variation-like' behavior. With the note on spin states above, Eq. (3) is suggested primarily to calculate stationary points (minimums and transition states) only [11], where the relative spin states are described reasonably plausibly by the Slater determinant formalism.

3. Test of RECEP estimation for correlation energy

First we selected several very important 10-electron systems, the protonated, neutral and deprotonated water and ammonia, and the methane and deprotonated methane. In these systems the correlation energy changes systematically depending on the number of protons and lone pairs. These molecules are ideal to represent the influence of the nuclear frame and lone pairs on the correlation energy while the $N = 10$, constant. We show the corresponding HF/6-311 + G(2d,p)//MP2/6-31G(d) total energies and various correlation energies in Table 2. All calculations were done with the G2 optimized geometries except the B3LYP/6-311 + G(2d,p) calculations, where the geometries were reoptimized. All correlation calculations show the same tendencies: for example the deprotonation of the neutral molecules changes the G2 correlation energies by -0.015 ± 0.002 (hartree), the protonation of one of the free lone pairs change the correlation energies by about $+0.0055$ (hartree). All methods (our RECEP estimation included) agree quite well. Moreover, if we go into the quantitative details we see an excellent agreement between the RECEP-d and G2 correlation energies (cf. Fig. 2). We note that the G2 method does not include the core correlation, thus the RECEP-d correlation energies are considerably more negative (by about 40 – 80 millihartrees, see the 0.0734 statistical analysis constant on Fig. 2). In order to show the approximate effect of the inclusion of the core correlation we show the full and frozen core CCSDT/6-311 + G(2d,p)//MP2/6-31G(d) correlation energies in Table 2. Fig. 2 shows the least square fit linear equation that transforms the RECEP-d values to G2 correlation energy values quantitatively; the R^2 (statistical correlation) is near to unity (0.99). This is our main point, that the estimation with Eq. (3) reaches the known high quality G2 estimation, and that while G2 is an expensive calculation, Eq. (3) practically needs no

Table 2
Molecular HF/6-311 + G(2d,p) total energies and correlation energies (hartree) calculated with various methods [8] for 10 electron molecules (fc = frozen core). The geometries were optimized on MP2/6-31G(d) level. Using the HF//6-311 + G(2d,p) NPA charges [8], the RECEP-c and -d correlation energy estimations were obtained by Eq. (3) with the corresponding values in Table 1

Molecules	HF–SCF energy	Correlation energy					
		CCSD-fc	CCSD-full	G2	RECEP-c	RECEP-d	B3LYP
CH_4	−40.2102	−0.2049	−0.2259	−0.2433	−0.2646	−0.2868	−0.3247
CH_3^-	−39.5195	−0.2140	−0.2346	−0.2558	−0.2870	−0.3043	−0.3380
NH_4^+	−56.5586	−0.2266	−0.2483	−0.2706	−0.3181	−0.3388	−0.3619
NH_3	−56.2151	−0.2312	−0.2526	−0.2766	−0.3282	−0.3478	−0.3687
NH_2^-	−55.5423	−0.2425	−0.2634	−0.2933	−0.3397	−0.3657	−0.3841
OH_3^+	−76.3314	−0.2446	−0.2662	−0.2945	−0.3451	−0.3763	−0.4002
OH_2	−76.0533	−0.2483	−0.2697	−0.2998	−0.3492	−0.3812	−0.4062
OH^-	−75.4056	−0.2595	−0.2805	−0.3155	−0.3718	−0.4080	−0.4222

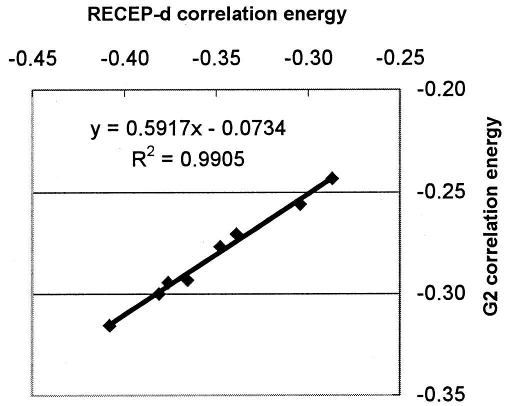

Fig. 2. Statistical data for RECEP-d vs. G2 correlation energies (hartree) listed in Table 2.

CPU time. For larger systems (e.g. with 50 − 100 electrons), G2 may have problems with disc space, CPU time or convergency, but until HF–SCF can be performed, the RECEP can readily be used for correlation calculation.

After the encouraging performance of the presently proposed RECEP-d method we include several more difficult cases and study the problems arising from this larger scope of molecules. Table 3 shows all the neutral molecules and ions selected for the present study. Fig. 3 shows linear fits for the RECEP and G2 relative correlation energies (in order to decrease the effect of the missing core correlation from the G2, we recalculated the correlation energies relative to LiH as zero). In LiH the core correlation is relatively important; thus it is expected that the G2 method will provide moderate performance. It can be observed that the LiH is rather off the linear regression curve. It should also be noted that the G2 method reproduces quite well the value of heat of formation, ΔH_f^0 (298 K), for LiH. We also note the case of the F^- ion. For this ion the full-CI calculations provide considerably more negative correlation energy than the G2 method, and certainly the latter is in error. Our estimations of the correlation energy for the fluoride containing molecules are considerably more negative than the G2 estimation. This can be partly explained by the inclusion of the core correlation, but this is not sufficient, because the difference between the G2 and full-CI estimates for the correlation energy is fairly large (cf. Table 3). As Curtiss and Pople noted recently: ‘‘Apparently there is some inherent prob-

Table 3
Number of electrons (N), NPA charges[a] (a.u.) on the non-hydrogen atom(s) in the molecule, RECEP-c, RECEP-d, and G2 correlation energies (hartree). The molecules are listed in monotonic order with N, and if N is the same, with G2 correlation energy values

Molecule	N	NPA charge	RECEP-c	RECEP-d	G2
LiH	4	0.812	−0.0799	−0.0901	−0.0343
BeH_2	6	1.218	−0.1107	−0.1258	−0.0696
BH_3	8	0.434	−0.1645	−0.1946	−0.1232
CH_4	10	−0.711	−0.2646	−0.2868	−0.2433
CH_3^-	10	−1.335	−0.2870	−0.3043	−0.2558
NH_4^+	10	−0.824	−0.3181	−0.3388	−0.2706
NH_3	10	−1.031	−0.3282	−0.3478	−0.2766
NH_2^-	10	−1.512	−0.3397	−0.3657	−0.2933
OH_3^+	10	−0.777	−0.3451	−0.3763	−0.2945
HF	10	−0.564	−0.3755	−0.4162	−0.2951
OH_2	10	−0.927	−0.3492	−0.3812	−0.2998
F^-	10	−1.000	−0.3995	−0.4430	−0.3150
OH^-	10	−1.362	−0.3718	−0.4080	−0.3155
LiF	12	−0.976	−0.4413	−0.4904	−0.3064
C_2H_2	14	−0.225	−0.3966	−0.4312	−0.3408
B_2H_6	16	0.017	−0.3693	−0.3926	−0.2724
C_2H_4	16	−0.339	−0.4395	−0.4774	−0.3544
C_2H_6	18	−0.510	−0.4839	−0.5251	−0.3745

[a] On H atoms, the NPA can be obtained from symmetry and molecular charge using the NPA values on non-hydrogen atoms listed; in case of LiF the NPA refers about Fluor atom; for Eq. (3) the NPA is a good choice for par(A).

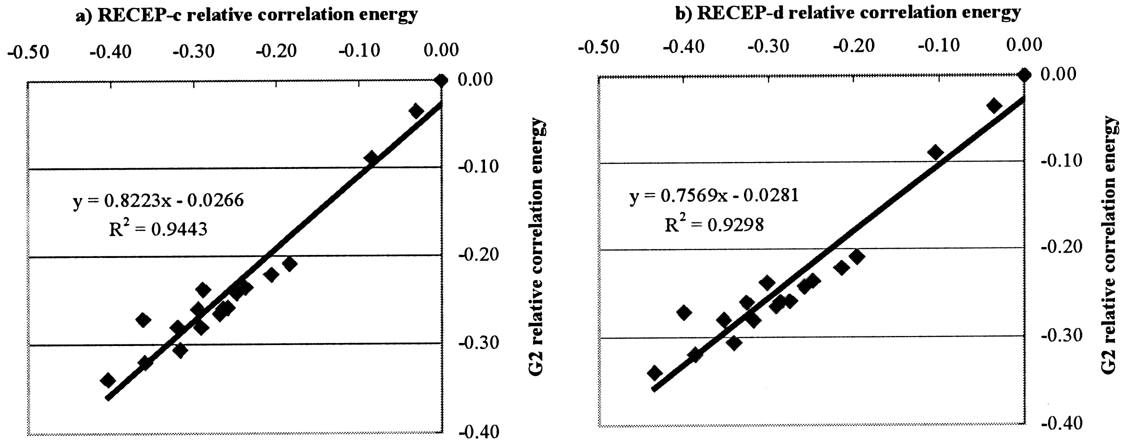

Fig. 3. RECEP-c (a) and RECEP-d (b) relative correlation energy (cf. Table 1) vs. G2 relative correlation energy (hartree) for the second set of molecules listed in Table 3. (Relative means: values deviating with respect to LiH molecule.)

lem in G2 theory with some of the fluorine molecules'' [34].

4. Conclusions

The quasi-linear dependence of the correlation energy on the partial NPA charges in molecules was developed and analyzed in Eq. (3), which is literally a density functional method based on a totally different technique than the known other methods. We propose two sets for the atomic correlation parameters to use in Eq. (3). One atomic correlation energy set is derived from the atomic full-CI and HF–SCF limit results, noted as RECEP-c (footnote c values in Table 1); another set is derived from the B3LYP calculations, noted as RECEP-d (footnote d values in Table 1). The partial atomic charges used in Eq. (3) were derived from the NPA analysis of HF/6-311 + G(2d,p) wave-function, readily available in (e.g.) the GAUSSIAN 94/DFT package [8]. The results show that both RECEP-c and RECEP-d parameter sets provide similar results for the selected 18 molecules and ions that contain first row elements. For the protonated, neutral, and deprotonated water and ammonia, and the methane and the deprotonated methane we obtained a good linear relationship between RECEP-d and G2 correlation energies ($R^2 = 0.99$), although the slope is far from the unity. Because of the missing core correlation from G2 there is a -73.4 mh shift between RECEP-d and G2 correlation values. Comparison of the G2 and RECEP relative correlation energies for the full set of 18 molecules and ions reveals, that there is a slightly better agreement between the RECEP-c and G2 values ($R^2 = 0.944$), than between the RECEP-d and G2 values ($R^2 = 0.930$). The slope is considerably closer to the unity in the former case as well.

In summary, the method in Ref. [11] has been refined in this Letter in four respects: First by using the partial charges from natural population analysis, which has good convergence properties with respect to the basis set increase. Second with an improved accounting for correlation energy for H atoms in the molecules. Third by suggesting new atomic parameter sets; and fourth by providing a theoretical background for partitioning the correlation energy for atoms in molecules.

5. Supplementary material

A small database of the geometries, NPA charges, and energies of 18 selected molecules and the RECEP atomic parameters is provided via `http://web.inc.bme.hu/mols/recep/`. The simple Fortran program for Eq. (3) with the RECEP-d $E_{\mathrm{corr}}(N_A, Z_A)$ parameters can be downloaded from `http://userwww.service.emory.edu/`

`~ekristy/` or can be obtained via e-mail (`skrist2@emory.edu`) or surface mail.

Acknowledgements

SK is thankful to the Domus Hungarica Scientarium et Artium 1998 summer grant for partial support of this Letter.

References

[1] P.W. Atkins, Molecular Quantum Mechanics, 2nd edn., New York, 1983.

[2] A. Szabo, N.S. Ostlund, Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory, New York, 1982.

[3] J. Sadlej, Semi-Empirical Methods of Quantum Chemistry, New York, 1985.

[4] S. Kristyan, Computers in Physics 8 (1994) 556.

[5] Z. Peng, S. Kristyan, A. Kuppermann, J. Wright, Phys. Rev. A 52 (1995) 1005.

[6] P. Pulay, K. Wolinski, T.P. Hamilton, J.M. Bofill, G. Fogarasi, TX90 Program System, University of Arkansas, Fayetteville, AR, 1982-1990. DFT subroutines were written by S. Kristyan in 1993-1994 for P. Pulay to build into the main program package.

[7] Biosym Technologies, 9685 Scranton Road, San Diego, CA 92121-2777, USA: Turbomole version 2.3 program, 1993 December; InsightII/Turbomole/Parameters/Properties pulldown: 'ESP Charges' for van der Waals fit to the molecular wave function to calculate partial charges in the molecule.

[8] GAUSSIAN 94/DFT, Revision F.2: M.J. Frisch, G.W. Trucks, H.B. Schlegel, P.M.W. Gill, B.G. Johnson, M.W. Wong, J.B. Foresman, M.A. Robb, M. Head-Gordon, E.S. Replogle, R. Gomperts, J.L. Andres, K. Raghavachari, J.S. Binkley, C. Gonzalez, R.L. Martin, D.J. Fox, D.J. Defrees, J. Baker, J.J.P. Stewart, J.A. Pople, Gaussian, Pittsburgh, PA, 1993.

[9] S. Kristyan, Chem. Phys. Lett. 247 (1995) 101.

[10] S. Kristyan, Chem. Phys. Lett. 256 (1996) 229.

[11] S. Kristyan, Chem. Phys. Lett. 224 (1997) 33.

[12] C. Moller, M.S. Plesset, Phys. Rev. 46 (1934) 618.

[13] J. Cizek, Adv. Chem. Phys. 14 (1969) 35.

[14] K. Raghavachari, J.B. Anderson, J. Phys. Chem. 100 (1996) 129.

[15] L.A. Curtiss, K. Raghavachari, G.W. Trucks, J.A. Pople, J. Chem. Phys. 94 (1991) 7221.

[16] C. Lee, W. Yang, R.G. Parr, Phys. Rev. B 37 (1988) 785.

[17] B. Miehlich, A. Savin, H. Stoll, H. Preuss, Chem. Phys. Lett. 157 (1989) 200.

[18] A.D. Becke, J. Chem. Phys. 88 (1988) 1053.

[19] A.D. Becke, J. Chem. Phys. 88 (1988) 2547.

[20] A.D. Becke, J. Chem. Phys. 96 (1992) 2155.

[21] A.D. Becke, J. Chem. Phys. 97 (1992) 9173.

[22] A.D. Becke, J. Chem. Phys. 98 (1993) 5648.

[23] R.G. Parr, W. Yang, Density-Functional Theory of Atoms and Molecules, New York, 1989.

[24] S. Kristyan, P. Pulay, Chem. Phys. Lett. 229 (1994) 175.

[25] S. Kristyan, J. Chem. Phys. 102 (1995) 278.

[26] P.A.M. Dirac, Proc. Cambridge Phil. Soc. 26 (1930) 376.

[27] E.P. Wigner, Trans. Faraday Soc. 34 (1938) 678.

[28] E.R. Davidson, S.A. Hagstrom, S.J. Chakravorty, V.M. Umar, C.F. Fischer, Phys. Rev. A 44 (1991) 7071.

[29] S.J. Chakravorty, S.R. Gwaltney, E.R. Davidson, F.A. Parpia, C.F. Fischer, Phys. Rev. A 47 (1993) 3649.

[30] E.R. Davidson's www home page: `http://ezinfo.ucs.indiana.edu/~davidson` and `http://phb.indiana.edu/~davidson/atom/tub.3.`

[31] R.F.W. Bader, Adv. Quant. Chem. 14 (1981) 63.

[32] L.E. Chirlian, M.M. Francl, J. Comput. Chem. 8 (1987) 894.

[33] A.E. Reed, R.B. Weinstock, F.J. Weinhold, J. Chem. Phys. 83 (1985) 735.

[34] L.A. Curtiss, K. Ragavachary, P.C. Redfem, J.A. Pople, J. Chem. Phys. 106 (1997) 1063.