

Correlation energy estimates in periodic extended systems using the localized natural bond orbital coupled cluster approach

N. Flocke and R. J. Bartlett^{a)}

Quantum Theory Project, Department of Chemistry and Physics, University of Florida, Gainesville, Florida 32611

(Received 31 October 2002; accepted 3 January 2003)

A new approach for the determination of correlation energies in periodic extended systems is proposed using the high transferability of amplitudes and integrals from natural bond orbital coupled cluster (NBO CC) calculations performed for small subunits. It is shown that the NBO CC calculations can in fact deliver detailed correlated wave function information for extended periodic systems. As an example we apply the ideas presented in this paper to determine an estimate for the valence correlation energy in diamond at the CCSD level. © 2003 American Institute of Physics. [DOI: 10.1063/1.1555123]

I. INTRODUCTION

The treatment of extended systems beyond the Hartree–Fock level still poses considerable difficulties.¹ The need for large one-electron basis sets for accurate size-extensive correlation calculations severely limits the size of the systems that can be treated. While it is possible to some extent to extrapolate subsystem correlation results in the case of 1D extended systems, the situation becomes hopeless for the 2D and 3D cases due to the quadratic and cubic increase in subsystems size. Electron correlation on the other hand is a rather “short-sighted” effect and there should be, *a priori*, no need to consider the whole spatial extent of extended systems but rather only small subregions at a time. This can be facilitated by properly transforming the traditional canonical HF orbitals into orbitals of a more spatial local nature. Then with proper algorithms, these localized orbitals effectively achieve “deoverlapping” of correlation space and restore to some extent the local nature of the original atomic basis set.

The usefulness of localizing electron correlation was already realized in the early days of quantum chemistry^{2,3} and resulted in the design of several localization schemes for HF orbitals. Localized orbitals, apart from possessing high interpretative value, effectively achieve the effect of electronic correlation being confined to small local spatial neighborhoods within the entire molecular framework and can result in correlation algorithms which scale linearly with the size of the system. The exponential scaling behavior with the number of electrons, which hampers the traditional correlation methods based on canonical HF orbitals, is thus avoided. Several research groups^{4–27} have been studying and designing algorithms based on the localized picture of electron correlation, and the technique holds great promises for correlation treatments of very large molecular systems in the near future.

The crucial aspect of all localized correlation methods is the neglect of most of the integrals and cluster amplitudes

(excitation parameters) due to their rapid decay with spatial distances between the localized orbitals involved. Several types of localized correlation treatments exist, depending upon the choice and construction of the localized orbitals. Actually, the oldest using CC theory is that of Laidig, Purvis, and Bartlett^{5,6} that exploited simple localized bond orbitals.²⁸ The next is based on unitary transformations applied to the occupied and virtual HF orbitals separately, yielding two sets of occupied and virtual localized molecular orbitals (LMO's).^{10,29–31} The third approach uses the original AO basis and reformulates the correlation equations in terms of that basis.²⁴ The AO basis constitutes the most local possible basis, however its nonorthogonality and nonseparability into occupied and virtual sets poses additional complications, resulting in more complex correlation equations with many more excitation parameters and the necessary inclusion of overlap matrix elements. A fourth localized correlation treatment was developed to overcome the difficulties associated with localization of the HF virtual space and can be considered a hybrid between the first two methods. The LMO/AO hybrid approach^{18,20,27} treats the occupied space as LMO's, but the virtual space is projected back to AO-type orbitals. The LMO/AO method is intermediate in complexity between the LMO and the pure AO approach.

All three LMO, AO, and LMO/AO methods are based on the HF reference function and hence lead to correlation energy invariance at each level of approximation. Following older work,^{5,6} a fifth approach to localized correlation treatment was suggested recently using natural bond orbitals (NBO's).³² The use of NBO's avoids the localization problems of the HF virtual space that are experienced during LMO construction by localizing the occupied and virtual HF spaces *together*. This in effect shifts some localization from the occupied to the virtual space and treats them on a more equal footing. The net result is a more uniformly localized overall correlation space at the expense of losing correlation energy invariance. In a recent paper it has been shown that a proper NBO CC treatment with inclusion of specific \hat{T}_1 amplitudes to correct for the NBO reference determinant can

^{a)} Author to whom correspondence should be addressed. Electronic mail: bartlett@qtp.ufl.edu

lead to a powerful localized correlation method exhibiting not only rapid exponential decay of amplitudes and integrals but also showing excellent transferability of amplitudes and integrals between similar molecular structures.³² In the present investigation we thus examine the possibility of using the NBO CC transferability properties to construct wave functions for extended systems via amplitudes and integrals obtained from NBO CC calculations on much smaller subunits. The prime example of the diamond lattice is chosen, primarily because of its simplicity (only one atom type and only two-center bonds defining its valence part) but also because of availability of correlation results from other investigations using different methods.

II. THEORY

The starting point for any NBO CC calculation is the construction of the NBO basis from a suitable one-particle density matrix. Details of this procedure about individual NBO stages of construction via natural atomic (NAO) (Ref. 33) and natural hybrid (NHO) (Ref. 34) orbitals can be found elsewhere and will not be repeated here.^{35,36} The resulting NBO basis can be classified into several types of functions according to their associated density matrix occupation numbers and number of atomic centers involved in their construction: core, lone-pair, two- or more-centered bond, anti-bond, and Rydberg NBO's. For a closed-shell molecule the first three would have occupation numbers close to 2, while the rest would have ≈ 0 . This fact enables the construction of a reference function in NBO space by assigning two electrons to each of the core, lone-pair and bond type NBO's and leaving the rest empty. Such a reference function reflects the way chemists view the structures of molecular compounds. However the exact HF density is only reproduced to a good approximation.

As such, the NBO CC approach to correlation is somewhat different from the usual MO CC or even the AO-based CC methods. While, due to the orthogonality of the NBO's, the same correlation equations result for both the NBO CC and the MO CC the energy correction relative to the reference will be different. To obtain comparable "correlation" corrections, the NBO CC needs to overcome the deficiency inherent in its reference function as compared to the HF reference. In a recent paper³² we showed how this is done most efficiently using Thouless theorem,³⁷ which connects the HF with the NBO reference function via a set of T_1 amplitudes obtainable at almost no cost. From a computational point of view, using such precalculated T_1 amplitudes as a start for NBO CC calculations enables one to perform NBO CC calculations which show almost the same number of iterations as their HF CC counterparts.

The NBO CC equations are identical to the MO based ones. In its spin-independent formulation³⁸ the cluster excitation operators are given in terms of the total spin preserving unitary group generators $E_{pq} = p_{\alpha}^{\dagger} q_{\alpha} + p_{\beta}^{\dagger} q_{\beta}$,

$$\hat{T}_1 = \sum_{ia} t_i^a E_{ai}, \quad (1)$$

$$\hat{T}_2 = \frac{1}{2} \sum_{ijab} t_{ij}^{ab} E_{ai} E_{bj}, \quad (2)$$

with obvious extensions to higher excitation clusters. The t 's are the corresponding excitation amplitudes with indices i, j denoting occupied and a, b virtual orbitals. The correlation energy is given by the following expression and is independent of the CC approximation level:

$$\Delta E = 2 f_i^a t_i^a + w_{ij}^{ab} (t_{ij}^{ab} + t_i^a t_j^b), \quad (3)$$

where f_i^a denote elements of the occupied-virtual block of the Fock matrix (nonzero in the NBO CC case) and antisymmetrized integrals,

$$w_{ij}^{ab} = 2 \langle ab | ij \rangle - \langle ba | ij \rangle \quad (4)$$

in Dirac notation have been used. Also summation over all repeated indices is assumed. For the present paper it is convenient to divide contributions to the correlation energy into a 1-particle ΔE_1 and a 2-particle ΔE_2 contribution,

$$\Delta E = \Delta E_1 + \Delta E_2, \quad (5)$$

$$\Delta E_1 = 2 f_i^a t_i^a + w_{ij}^{ab} t_i^a t_j^b, \quad (6)$$

$$\Delta E_2 = w_{ij}^{ab} t_{ij}^{ab}. \quad (7)$$

The quantity ΔE_2 represents the pairwise correlation of electrons in a molecular system and in the HF picture gives by far the most important contribution to the total electronic energy. In case of the NBO approach, it was found that the precalculated T_1 amplitudes connecting the HF with the NBO reference are already a very good approximation to the final set of NBO CCSD T_1 amplitudes. This fact, together with the almost invariant total CCSD energy results obtained for both HF and NBO cases,³² makes the magnitude of ΔE_2 in both HF and NBO cases almost identical. The NBO ΔE_2 values are lower by 1% as compared to the HF ones, indicating a slightly greater pair correlation in the NBO picture.

The T_2 amplitudes t_{ij}^{ab} and the corresponding antisymmetrized integrals w_{ij}^{ab} become highly transferable in the NBO case. This can ultimately be traced to the NBO construction procedure, which in principle localizes the entire set of one-electron basis functions in a molecule via analysis of the first-order density matrix. Thus the highly localized character is not only present in the occupied space consisting of core, lone-pair and bond NBO's, but is also characteristic for the virtual space comprising the anti-bond and Rydberg NBO's. Further both sets of amplitudes and integrals decay exponentially in magnitude with NBO separation in space. Both properties, high transferability and exponential value decay, opens the road to determine the correlation energy of extended systems by performing manageable NBO CC calculations on some small system subunits and calculating ΔE_2 for the whole system from the ΔE_2 's of the corresponding subunits,

$$\Delta E_2(\text{total}) = \sum_{i=\text{subunits}} \Delta E_2(i). \quad (8)$$

In the next section we will show how this idea is applied to calculate the correlation energy of a diamond crystal.

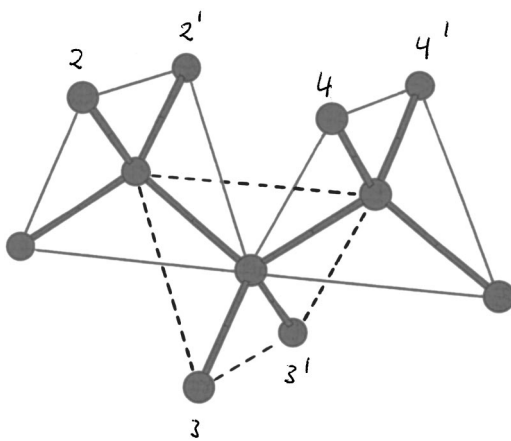


FIG. 1. Subunit of the diamond lattice, showing interconnected solid- and broken-line tetrahedra used to determine mono-, di-, and tri-tetrahedral excitation contributions to the diamond correlation energy per unit cell.

A. Correlation energy of diamond crystal using NBO CCSD results

In diamond each carbon atom is surrounded tetrahedrally by four nearest-neighbor carbons. Its structure can be viewed as composed of vertex-joined body-centered regular tetrahedra, where each vertex is surrounded tetrahedrally by four such tetrahedra. Note that although each body-centered tetrahedron has the right atom count of 2 carbon atoms (1 in the center and $4 \times 1/4$ for the vertices) it cannot be considered a primitive unit cell of diamond due to the impossibility of generating the diamond lattice from it by pure translations. However, for our case of calculating the correlation energy of diamond this basic body-centered tetrahedral building block proves to be best. The set of all carbon atoms in diamond are thus split into two sets: the set consisting of all the carbon atoms sitting at the center of all tetrahedra and the set of all carbon atoms sitting at all the vertices. Figure 1 shows a small piece of diamond illustrating this partitioning.

Let us now apply Eq. (8) to calculate the first approximation to the correlation energy of diamond. To do so we will restrict our subunits in the rhs of Eq. (8) to single tetrahedra only, i.e., we will neglect correlation contributions to ΔE_2 involving amplitudes and integrals over NBO's more than 2 centers apart. The smallest molecule containing such a carbon tetrahedral unit is neo-pentane C_5H_{12} , in which each vertex carbon is terminated by three hydrogen atoms.

Assuming transferability of NBO CC amplitudes and integrals, we solve the NBO CCSD equations for neo-pentane in a geometry in which the carbon atoms are arranged in the same way as in the diamond lattice. Each carbon-carbon length in diamond equals 1.54 \AA and tetrahedral angles between all vertices and the central carbons are assumed. The carbon-hydrogen distance was set to 1.10 \AA and all hydrogen-hydrogen as well as hydrogen-carbon angles were set equal to the tetrahedral angle. All amplitudes and integrals involving only the carbon NBO's were extracted and properly combined. Note that when evaluating the correlation energy of diamond per unit cell it is not appropriate to consider each amplitude/integral pair that enters ΔE_2 with a weight factor of one. Rather the evaluation of the correlation energy per unit cell is done in two steps (see Fig. 1). The first step considers only the contributions inside each solid-line tetrahedron. In this case all amplitude/integral pairs are given a weight factor of 1 with the exception of those pairs which involve only NBO's belonging to one vertex carbon, in which case the weight factor is $1/4$, because each such atom is shared by 4 neighboring solid-line tetrahedra. The second step considers the contribution arising due to the broken-line tetrahedra. To evaluate these contributions with the proper weight factors, we single out a vertex carbon in neo-pentane together with its associated NBO bond formed with the central carbon and consider only those excitations from that atom/bond unit to the rest of the tetrahedral structure. Excitations within that unit and to the central carbon must be excluded because they were already counted in the first step. The weight factors are evaluated as follows: suppose an excitation from the chosen atom/bond unit involves X other atom/bond units. These X other units belong each to a different solid-line tetrahedron (see Fig. 1) and thus a factor of $1/(1+X)$ is necessary to avoid overcounting. On the other hand there are 4 specific atom/bond units within each solid-line tetrahedron, so the final weight factor is $4/(1+X)$. Table I presents the results obtained from the all-electron NBO CCSD calculations of neo-pentane using the basis set combinations DZP/DZ, DZP/DZP, TZP/TZ, TZ2P/TZ, and CC-PVTZ/TZ for the carbon/hydrogen atoms. These valence double- and triple- ζ plus polarization functions basis sets^{39,40} have the following contraction patterns. For carbon we have DZP:($9s5p1d$)/[$4s2p1d$], TZP:($10s6p1d$)/[$5s3p1d$], TZ2P:($10s6p2d$)/[$5s3p2d$], and CC-PVTZ:($10s5p2d1f$)/[$4s3p2d1f$]. For hydrogen

TABLE I. Contribution of diamond correlation energy per unit cell from local excitations within each tetrahedra using NBO CCSD amplitudes and integrals from neo-pentane. All energies in Hartrees.

Excitation type	DZP/DZ	DZP/DZP	DZP/DZ(sp^3)	TZP/TZ	TZ2P/TZ	CC-PVTZ/TZ
same atom	-0.0225	-0.0224	-0.0226	-0.0232	-0.0235	-0.0115
same bond	-0.0437	-0.0435	-0.0430	-0.0384	-0.0383	-0.0382
atom-atom	-0.0005	-0.0005	-0.0006	-0.0015	-0.0014	-0.0044
bond-bond	-0.0355	-0.0353	-0.0340	-0.0352	-0.0345	-0.0330
atom-bond	-0.1231	-0.1225	-0.1266	-0.1453	-0.1587	-0.1846
core	-0.0318	-0.0317	-0.0317	-0.0372	-0.0404	-0.0336
valence	-0.1935	-0.1927	-0.1950	-0.2065	-0.2159	-0.2381
total	-0.2253	-0.2244	-0.2267	-0.2437	-0.2563	-0.2717

we have $DZ:(4s)/[2s]$ and $TZ:(5s)/[3s]$. All results were obtained using the QCPACK software.⁴¹

The total core plus valence correlation contribution was split into the following type of excitations: (i) those involving NBO's on one atom only (same-atom), (ii) the ones involving one bond only, i.e., excitations from a bond to its corresponding antibond (same-bond), (iii) excitations involving NBO's on different atoms (atom-atom), (iv) excitations involving NBO bonds to other NBO antibonds (bond-bond), (v) the rest of the excitations connecting NBO's of atoms and bonds (atom-bond), (vi) excitations involving at least one contribution from the carbon $1s$ core (core), and (vii) pure valence excitations (valence).

The first observation to make is that although the individual amplitude times integral contributions to the correlation energy coming from the local type of excitations (same-atom and same-bond) are by far the largest in magnitude, their very small occurrence number (for example, only 4 such excitation products for the same-bond case) prevents them from dominating the total correlation result. For a quantitative description of the correlation energy in diamond it is thus essential to consider all possible excitations and any attempt to separate dominant types of excitations will lead to considerable errors. The core contributions are far less than the valence ones (ratio approximately 1:6 for the double- ζ and 1:5 for the triple- ζ type basis set calculations). These values partly reflect the fact that no explicit core correlation basis functions were used in our chosen basis sets. Hence the core correlation energies presented here are not very meaningful. To gain more information about the core correlation contributions one would have to resort to basis sets like for example CC-PCVTZ.

As expected the increase of the quality of the basis sets on the carbon atoms will lead to larger valence correlation energy contributions. This is shown in the first and the last three columns of Table I, where the carbon basis set is improved from DZP to CC-PVTZ quality and the valence correlation energy increases in magnitude from -0.1935 to -0.2381 Hartrees. The considerable increase in the correlation energy contribution from TZP to CC-PVTZ indicates that further improvement of the carbon basis set including more polarization functions will lead to a further increase in the correlation energy. A very rough estimate for this first approximation in the diamond valence correlation energy per unit cell in the basis set limit, where correlation excitations are limited to within each tetrahedral unit, would be around -0.27 Hartrees.

Columns one, two, and three of Table I shows what happens when different basis sets are used for the hydrogen atoms (compare columns one and two) and when the angles between the hydrogens on all methyl groups are adjusted to 98° , such that the carbon NHO's forming the bond NBO's in neo-pentane are pure sp^3 NAO hybrids (compare columns one and three). The differences in correlation energy induced by these changes are only minor. Increase of the hydrogen basis set from DZ to DZP quality lowered the total correlation contribution by 0.9 mHartrees, while the change to pure sp^3 NHO hybrids increases the total correlation energy by 1.4 mHartrees. No big shifts were observed for the individual

type of excitations. The percentage of total correlation energy change was thus less than 0.7% for these cases compared to the original value in the first column. This shows that the quality of the basis set on the terminating hydrogens as well as their spatial arrangement around the carbon atom has little influence on the correlation results, as long as these "variables" remain within reasonable limits, i.e., the basis set imbalance between the carbon and hydrogens should not become too great (as, for example, when mixing triple- ζ for carbon with double- ζ for hydrogen) and the angles between the hydrogens and the carbons as well as the hydrogen-carbon bond lengths should not depart significantly from the common experimental values for saturated hydrocarbons.

How can we improve the first correlation energy approximation for diamond? Looking at Fig. 1, the next larger diamond subunit would be the one formed by all carbons from a solid line tetrahedron plus the ones from the broken line tetrahedron, i.e., a di-tetrahedral structure involving 8 carbon atoms. Terminating the dangling bonds with hydrogens we thus arrive at the saturated hydrocarbon 2,2', 3,3'-tetramethylbutane C_8H_{18} . A NBO CCSD calculation on this structure would hence give us information about magnitude and importance of correlation contributions involving new di-tetrahedral excitations of longer range that interconnect NBO's from the solid- with the broken-line tetrahedra. Note that due to the two-particle nature of both the amplitudes and the integrals in Eq. (7), it is sufficient for one excitation to be of di-tetrahedral type in order to classify the whole amplitude-integral product as di-tetrahedral. The weights for those excitations lying completely within one tetrahedron (both the solid and the broken line tetrahedron are equivalent by symmetry in C_8H_{18} , hence only one is considered) were already determined for the neo-pentane case. The weights for the di-tetrahedral contributions are deduced as follows. Each di-tetrahedral excitation can be associated with only one definite solid-and broken-line tetrahedron pair, that is, no di-tetrahedral excitation can belong at the same time to more than one such pair. Each solid-line tetrahedron gives rise to 4 unique such di-tetrahedral pairs, hence the weight factor for diamond is 4 for all di-tetrahedral excitations. The di-tetrahedral amplitude-integral products can again be classified into several categories: those that have no local excitations belonging to either the solid- or broken-line tetrahedron (nonlocal), excitation products which have only one such local excitation (one-local), excitation products in which one excitation belongs to the solid-line tetrahedron and the second one to the broken-line tetrahedron (two-local), excitation products involving only excitations between bonds and antibonds (bond only) and the valence contributions. Table II summarizes the results obtained for the di-tetrahedral diamond correlation energy contributions per unit cell as well as the correlation contribution due to a single tetrahedral unit (referred from now on as mono-tetrahedral) from all-electron NBO CCSD calculations on C_8H_{18} using the basis set combinations DZP/DZ and TZP/TZ. The same geometry specifications were used as for neo-pentane with the addition of a staggered structure of the methyl groups around the central carbon-carbon bond.

The di-tetrahedral contribution to the diamond correla-

TABLE II. Contribution of diamond correlation energy per unit cell from mono- and di-tetrahedral excitations using NBO CCSD amplitudes and integrals from 2,2',3,3'-tetramethylbutane C_8H_{18} . All energies in Hartrees.

Excitation type	DZP/DZ	TZP/TZ
di-tetrahedral		
nonlocal	-0.0003	-0.0004
one-local	-0.0108	-0.0138
two-local	-0.0283	-0.0319
bond only	-0.0136	-0.0155
valence	-0.0386	-0.0446
total	-0.0394	-0.0460
mono-tetrahedral		
bond only	-0.0771	-0.0721
valence	-0.1933	-0.2022
total	-0.2254	-0.2394

tion energy is far less than the mono-tetrahedral contribution, reflecting the exponential decay in magnitude for both NBO amplitudes and integrals. The largest contribution to the di-tetrahedral correlation energy comes from those excitation products consisting of two local tetrahedral excitations. The contribution of the nonlocal excitation products is almost negligible, being smaller by almost two orders of magnitude compared to the two-local ones. The ratio of numbers of excitation products for the non-, one-, and two-local cases are roughly 1:4:2 (the exact numbers for the DZP/DZ case are 156816:719928:357858), hence the importance of terms is not related to how many terms are present in each category but rather to the absolute values of the individual terms. The mono-tetrahedral correlation value remains virtually unchanged for the DZP/DZ case, but deviates somewhat by 4.3 mHartrees when using the TZP/TZ basis (compare Table II results with first and fourth column of Table I). The almost invariant results for the mono-tetrahedral correlation contribution when moving from one model structure to the next is a strong indication of the high transferability of the individual NBO excitation products entering the correlation energy expression. Increase of correlation contribution when moving from DZP/DZ to TZP/TZ is quite pronounced and larger percentage-wise than the corresponding mono-tetrahedral increase.

With inclusion of the di-tetrahedral component all excitation products are considered, whose underlying NBO's are not longer than 4 carbon centers apart. The next step in refining the diamond correlation energy would be to include all those excitation products which involve at least one excitation ranging over 5 carbon centers. From Fig. 1 we see that this would correspond to using as a model the saturated hydrocarbon tetra-tertbutyl-methane $(C_4H_9)_4C$ (including two more solid-line tetrahedra with carbon atoms 3 and 3' as centers in Fig. 1) for our NBO CCSD calculations, but the use of the large 327 basis functions DZP/DZ basis set combination would lead to diminishing returns. Alternatively, we use the model structure as presented in Fig. 1, leading to the model saturated hydrocarbon 2,2', 3,3', 4,4'-hexamethylpentane $C_{11}H_{24}$ with three interconnected tetrahedra. Extra tri-tetrahedral long-range excitations arise within this model (apart from the mono- and di-tetrahedral excita-

TABLE III. Contribution of diamond correlation energy per unit cell from mono-, di-, and tri-tetrahedral excitations using NBO CCSD amplitudes and integrals from 2,2',3,3',4,4'-hexamethylpentane $C_{11}H_{24}$ and its stepped-down model 2,2',4,4'-tetramethylpentane C_9H_{20} . Results obtained for the DZP/DZ basis set combination. All energies in Hartrees.

Excitation type	$C_{11}H_{24}$	C_9H_{20}
tri-tetrahedral		
nonlocal	-0.0005	-0.0003
one-local	-0.0168	-0.0140
two-local	-0.0144	-0.0136
valence	-0.0310	-0.0272
total	-0.0318	-0.0279
di-tetrahedral		
valence	-0.0411	...
total	-0.0420	...
mono-tetrahedral (solid)		
valence	-0.1934	-0.1936
total	-0.2256	-0.2255
mono-tetrahedral (broken)		
valence	-0.1915	...
total	-0.2235	...

tions) connecting atomic and bond NBO's between the two solid-line tetrahedra. Two points require some close attention when using the above mentioned model hydrocarbon. The first concerns the terminating hydrogens on the 2,2' and the 4,4' methyl groups. These four methyl groups must each be rotated by 30° in order to avoid a highly strained molecule due to excessive crowding of the hydrogens. Thus, as opposed to the former smaller hydrocarbon models, a purely staggered structure for $C_{11}H_{24}$ cannot be used. The second point arises due to the relation between the model hydrocarbon and the diamond structure. In diamond the methyl carbon atoms 2 and 4 (and likewise 2' and 4') would be bridged by a third carbon, forming the diamond six-membered ring structures. Hence an excitation product linking only carbon 2 and 4 atomic NBO's would not be of the long-range tri-tetrahedral type, but rather would belong to the mono-tetrahedral type considered earlier. Care must thus be exercised to eliminate those excitation products between the solid-line tetrahedra which belong to either the mono- or the di-tetrahedral type. The weights of the tri-tetrahedral excitations are determined in a similar way as for the di-tetrahedral ones. Each tri-tetrahedral excitation is shared between two solid-line tetrahedra (giving rise to a factor of 1/2), of which there are 12 pairs surrounding each solid-line tetrahedron (factor of 12). The same type of excitations arise among all broken-line tetrahedron pairs, of which there are the same number of pairs in a diamond lattice as for the solid-line case (factor of 2). Each tri-tetrahedral excitation is unique and belongs to only one specific tetrahedral pair (factor of 1). The overall weights are thus all equal and equal to $1 * 2 * 12 / 2 = 12$. The contributions to the correlation energy due to these long-range tri-tetrahedral excitations as well as the mono-tetrahedral and di-tetrahedral contributions obtained from the $C_{11}H_{24}$ hydrocarbon model are shown in Table III for the DZP/DZ basis set combination using the same carbon-carbon and carbon-hydrogen bond lengths as

for neo-pentane. The tri-tetrahedral excitations were also classified according to the following three categories: (i) nonlocal, with no excitation confined within one tetrahedron only, (ii) one-local, having only one mono-tetrahedral excitation, and (iii) two-local, made up of two mono-tetrahedral excitations.

Together with the results for the $C_{11}H_{24}$ model we also show the corresponding results for the model 2,2', 4,4'-tetramethylpentane C_9H_{20} , which can be considered a stepped-down model of the $C_{11}H_{24}$ model. The reasons we include the results for C_9H_{20} are on one side one more further check on transferability of the mono-tetrahedral results and on the other we also wanted to see of what order are the errors when using only stepped-down models.

The first point we notice in Table III is the fact that the correlation energy contribution due to the tri-tetrahedral excitations is of the same order of magnitude as the contribution due to the di-tetrahedral ones. Despite the fact that the magnitude of the amplitudes and integrals decrease exponentially with NBO distance, the 3D nature of the diamond lattice opposes this advantage by a cubical increase in the number of long-range excitations as we move from a central solid-line tetrahedron outwards. The tri-tetrahedral excitations give thus a substantial correlation energy contribution in the diamond lattice, which is solely related to the sheer number of such excitations and not their magnitude. The largest individual tri-tetrahedral excitation product was about a factor of 1/1000 smaller in magnitude than the total contribution.

The second observation parallels the one already made for the di-tetrahedral contributions. The number of two-local type of tri-tetrahedral excitations is far less than the number of one- and nonlocal ones (ratio 1:5:3), yet their correlation contribution is almost equal to the most numerous one-local excitations. However, their overall importance will decrease for higher than tri-tetrahedral excitations, so again we cannot select a specific subset or subsets of excitations as being always dominant for a particular long-range excitation.

Excellent transferability is observed for the mono-tetrahedral correlation contributions for $C_{11}H_{24}$ as well as for the C_9H_{20} model when compared with results from Tables I and II. The di-tetrahedral contributions differ somewhat from those in Table II, but are within reasonable limits. Note the two kinds of mono-tetrahedral correlation contributions the $C_{11}H_{24}$ model allows us to obtain. The one stemming from the broken-line tetrahedron is more closely related to the actual diamond structure (two tert-butyl substitutions instead of only one on the central carbon) and hence we took this value as our final mono-tetrahedral DZP/DZ result.

The tri-tetrahedral results for the stepped-down C_9H_{20} model are, as expected, lower than the ones for the $C_{11}H_{24}$ model. The excitation products missed by the stepped-down model are those involving excitation parts of the di-tetrahedral type, hence we would expect, under the condition of ideal transferability of the individual excitation products, only noticeable changes in the one-local tri-tetrahedral contributions. The two-local ones, involving necessarily only local excitations from the two solid-line tetrahedra, should not change at all. This is reflected in the Table III results, where

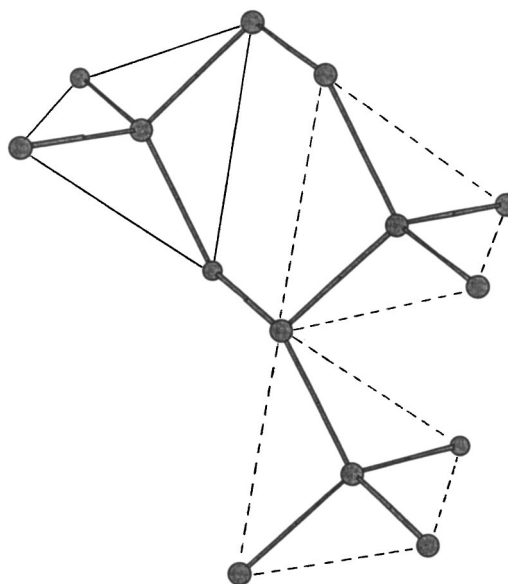


FIG. 2. Subunit of the diamond lattice used to determine tetra-tetrahedral excitation contributions to the diamond correlation energy per unit cell.

the one-local changes are much more pronounced, while the two-local ones show a slight change due to nonideal transferability of the excitation products between the two models.

We now move farther out from the central tetrahedral unit. In Fig. 2 we have drawn the diamond subunits which lead to saturated hydrocarbons still accessible to our NBO CCSD calculations. They all consider excitation products between NBO's which are 6 carbon centers apart (tetra-tetrahedral excitations). Two model hydrocarbons emerge from Fig. 2. The first one considers the long-range excitations between the plane-joined solid- and broken-line tetrahedra and leads to the 1,1', 4,4'-tetramethylcyclohexane $C_{10}H_{20}$ model. The other one is derived from the solid-/broken-tetrahedra pair which is line-joined and gives the hydrocarbon 2,2', 5,5'-tetramethylhexane $C_{10}H_{22}$. Since all different broken-line tetrahedra surrounding a solid-line tetrahedron can be obtained uniquely from Fig. 2 by twice a \hat{C}_3 rotation around the solid-line tetrahedral face facing the observer and since there are 4 such faces, the weight factor of each tetra-tetrahedral excitation product is 12. Note that both hydrocarbon models are considerably stepped-down models from the ideal diamond subunit derived from considering all broken-line tetrahedra around the solid-line tetrahedron (a structural beast containing 101 carbon atoms!). Hence we will be able to calculate only a fraction of the complete tetra-tetrahedral contribution to the diamond correlation energy. Table IV shows the mono- and tetra-tetrahedral diamond correlation contributions obtained for both $C_{10}H_{20}$ and $C_{10}H_{22}$ models using the DZP/DZ basis set combination.

Again the mono-tetrahedral contributions show excellent transferability properties and are almost identical with those obtained previously from the smaller models. The tetra-tetrahedral contributions are considerably smaller than the tri-tetrahedral ones, however, as mentioned before, the results of Table IV represent only a fraction of the overall tetra-tetrahedral contribution in diamond.

TABLE IV. Contribution of diamond correlation energy per unit cell from mono- and tetra-tetrahedral excitations using NBO CCSD amplitudes and integrals from 1,1',4,4'-tetramethylcyclohexane $C_{10}H_{20}$ and 2,2',5,5'-tetramethylhexane $C_{10}H_{22}$. Results obtained for the DZP/DZ basis set combination. All energies in Hartrees.

Excitation type	$C_{10}H_{20}$	$C_{10}H_{22}$
	tetra-tetrahedral	
valence	-0.0037	-0.0039
total	-0.0039	-0.0040
	mono-tetrahedral	
valence	-0.1931	-0.1932
total	-0.2250	-0.2251

Finally, Table V summarizes our valence correlation energy results for diamond at two basis set levels, namely, using a pure DZP/DZ basis throughout each model and using the corresponding largest basis set results for each model. Table V also includes for comparison results for the valence correlation energy per unit cell for diamond obtained by other workers using other kinds of methods.

B. Comparison with other methods

In this section we compare our diamond correlation results with those calculated by other methods. The first method we discuss is the increment method,⁴² in which the valence diamond correlation energy per unit cell was estimated including one- to three-body increments at energy level corresponding to a CEPA-0=LCCD (Ref. 43) level of approximation in which only one- to three-bond excitations from a Foster-Boys LMO localized HF space²⁹ were allowed. Like the NBO excitation products, the individual *energy* increments were found to be highly transferable from one hydrocarbon model to the next. Our DZP/DZ result in Table V is virtually identical, but slightly lower in magnitude than the DZP result obtained via the increment method. Note that our result includes excitation products from fourth-

nearest-neighbor bonds (see Fig. 2), whereas the results in Ref. 42 include only up to third-nearest-neighbor increments. However, the individual energy increment contributions seem to not always be negative, as the positive major three-body increment value shows. In contrast, our individual mono- through tetra-tetrahedral-type correlation energy contributions are found to be always negative. A direct comparison between different increment contributions and our results does not seem to be appropriate. For example, adding the DZP increments restricted to be within our mono-tetrahedral subunit (i.e., the result of Table III and the first results of Tables IV and V in Ref. 42), we would obtain a value of -0.224 which is much higher in magnitude than our average mono-tetrahedral valence result of -0.193 . On the other hand, the observed 20% increase in increment magnitude when changing from DZP to TZ2P basis correlates with our 23% increase in valence mono-tetrahedral correlation contribution when changing the carbon basis set from DZP (first column in Table I) to CC-PVTZ (last column in Table I). Note that the basis described as TZ2P in Ref. 42 corresponds actually to our CC-PVTZ basis set for the carbon atom. Again, the increment TZ2P valence correlation result can be considered identical with our “best basis” result from Table V.

The second method⁴⁴ is based on the old intuitive idea of estimating the overall correlation energy of a system as a sum of individual well localized correlation energy contributions between certain molecular electronic local substructures such as lone-pairs and bonds. The idea is based on introducing, formally, a pairwise interaction correlation operator of unknown structure and to express the correlation energy of the system in terms of the trace of the product between the second order density and the correlation operator matrix, both assumed to be over localized SCF MO's. Further simplification using the locality of the SCF MO's and the assumption that similar orbital pairs give comparable correlation operator matrix elements depending on orbital nature and their spatial separation, leads to a correlation energy expression in the form of a bilinear sum of products between so called group electron pair populations and group correlation strengths. The former are known exactly from SCF theory, while the latter are treated empirically and are fitted to reproduce theoretically calculated correlation energies of a large set of organic molecules. Note that, as opposed to the increment and NBO CC method, this extrapolation procedure avoids completely the necessity of constructing actual LMO's. Only the existence and possibility of constructing localized MO's is used as an argument to make the bilinear correlation energy expression plausible and to justify the neglect of group contributions involving localized MO's which are spatially far apart. Based on the extrapolation formulas obtained in Ref. 44, an average valence correlation energy of -0.313 for diamond per unit cell was given for a CC-PVTZ basis set quality, a value a little lower in magnitude than our “best basis” result of -0.321 , which is based on CC-PVTZ quality NBO CC results for the mono-tetrahedral excitations. One must keep in mind, however, that the CC-PVTZ extrapolated result was obtained including only up to vicinal-vicinal LMO pairwise correlations, which

TABLE V. Valence correlation energy of diamond per unit cell including up to tetra-tetrahedral excitations and results from other workers. All energies in Hartrees.

Method	
NBO CCSD (all DZP/DZ)	-0.271
NBO CCSD (best basis)	-0.321
Increment DZP ^a	-0.275
Increment TZ2P ^a	-0.322
Local ansatz (calculated) ^b	-0.228
Local ansatz (estimated) ^b	-0.324
Quantum MC ^c	-0.300
SCF/TZ fit (3.1) ^d	-0.312
SCF/TZ fit (4.1) ^d	-0.314
SCF/CB fit (3.3) ^d	-0.328
SCF/CB fit (4.2) ^d	-0.328
Experiment	-0.348

^aReference 42.

^bReference 45.

^cReference 46.

^dReference 44.

would correspond to our di-tetrahedral contributions. Unfortunately we could not evaluate the di-tetrahedral NBO CC contribution at the CC-PVTZ level, so we have no direct comparison between the extrapolation and the NBO CC method. However, since the extrapolation parameters are based on CCSD(T) calculations and somehow through the extrapolation fitting procedure incorporate part of the long-range correlation interactions, we expect them to do better than our NBO CC results. The (B) and (BB1) groups defined in Ref. 44 correspond, respectively, to intrabond and nearest-neighbor interbond excitation parameters and are thus entirely within our mono-tetrahedral diamond subunit. Taking the (B) and (BB1) parameter values given in Table VII in Ref. 44 and neglecting the longer range vicinal bond interactions (VV), the extrapolation result for the CC-PVTZ diamond valence correlation energy is still -0.310 per unit cell, whereas our mono-tetrahedral contribution of -0.238 is by far lower in magnitude. The energy extrapolation formula in Ref. 44 hence considerably overestimates the intrabond and nearest-neighbor interbond excitations, a result which is to be expected since they have to account for the considerable correlation contributions arising due to third-nearest and higher neighbor excitations.

The increment and extrapolation methods are both based on partial or total energy results of suitable small model systems. An entirely different approach comes from the local ansatz (LA) (Ref. 45) and quantum MC methods,⁴⁶ which construct a variational correlated wave function for different diamond cluster sizes by operating with an exponential correlation operator on the SCF determinant. The correlation operator is constructed in such a way as to contain the essential correlation effects like charge and spin fluctuations. In the LA method the correlation operators are defined in terms of local atomic functions constructed from the SCF determinant, ensuring rapid convergence of matrix elements with increasing cluster size. From Table V we can see that the calculated diamond valence correlation result is considerably lower than our DZP/PZ result. The calculated LA result was obtained using a DZP-type basis set and larger than second-nearest neighbor effects were neglected. It should hence be comparable in magnitude with our DZP/DZ results including just the mono- and di-tetrahedral contributions, and this is indeed the case, since we obtain, by adding up the mono- and di-tetrahedral valence contributions from our Table II, a value of -0.232 , which differs from the LA value only by about 4 mHartree. Another direct comparison can be made on the basis of how the LA result was decomposed into “interatomic” and “intra-atomic” correlation contributions. The “interatomic” part was defined as originating from correlation effects due to local bond-type atomic functions and hence would correspond to our “bond only” results in Table II. The “intra-atomic” part constitutes the valence minus the “bond only” results. We obtain for the “bond only” result a value of -0.0907 and for the valence minus “bond only” case a value of -0.1412 , which are both in very good agreement with the “interatomic” and “intra-atomic” values of -0.0925 and -0.1352 from Table I in the LA paper.⁴⁵ It seems thus that the low value obtained by the LA method is not so much a result of the poor LA configuration selection

scheme as claimed in Ref. 42, but rather is due to the modest size of the DZP basis set employed and the neglect of third-nearest and higher neighbor correlation effects.

III. CONCLUSION

In this paper we presented a new way of calculating *wave functions and correlation energies* for extended systems by performing NBO CC calculations on smaller subunits and using the resulting high transferability property of both CC T_2 amplitudes and integrals. The idea was applied to determine the valence correlation energy per unit cell in a diamond crystal and to gain some insight about how correlation effects are distributed within the diamond lattice at the local NBO picture. Although high transferability of NBO T_2 amplitudes was previously detected on some small saturated hydrocarbons and some of their substituted derivatives, there was *a priori* no definite conclusion if NBO CC transferability would be sufficient to allow us to make a reasonable estimate of the correlation energy for an extended system like diamond. In this respect the results obtained for the diamond crystal in the present investigation can be considered very encouraging. Not only are the results in very good agreement with those obtained by other methods under similar circumstances like basis sets used, but also transferability occurs on every single correlation energy contribution when the underlying models are different. We would like to emphasize that the present method based on NBO CC results uses in fact transferability at the *wave function* level and not at the energy level as for example is the case for the increment or the extrapolation method. This in turn implies that also the density matrices for extended systems becomes accessible from a few simple model studies and hence all properties of the extended system. The advantage of NBO's over the usual LMO's is that their localization is obtained for both the occupied (Lewis) and virtual (non-Lewis) orbitals on an equal footing, hence there are no problems in properly localizing the virtual space. The price to pay for this better localized NBO space, namely the slightly poorer reproduction of the HF density, is easily overcome at the CC level of correlation, where a proper choice of \hat{T}_1 amplitudes via Thouless theorem interconnects the HF determinant with the NBO determinant. The NBO's are thus valuable “probes” for gaining insight into the electron correlation phenomenon in extended systems. Further investigations are however needed, particular on extended systems known to possess some degree of electronic delocalization, to see if the NBO CC method can become a universal tool in estimating correlation energies for extended systems. A further enhancement on the size of systems that can be treated via NBO CC would come from formulating and implementing NBO CC algorithms that take into account localized properties of the amplitudes and integrals from the beginning. All results obtained in the present paper were obtained with a standard CC code and thus were limited to systems of fairly small size. Investigations about such localized NBO CC algorithms are underway.

ACKNOWLEDGMENT

Support by the U.S. National Science Foundation under KDI Award No. DMR-9980015 is greatly acknowledged.

- ¹J. Q. Sun and R. J. Bartlett, in *Topics in Current Chemistry* (Springer, Berlin, 1999), Vol. 203, p. 121.
- ²O. Sinanoğlu, *Adv. Chem. Phys.* **6**, 315 (1964).
- ³R. K. Nesbet, *Adv. Chem. Phys.* **9**, 321 (1965).
- ⁴W. Meyer, *J. Chem. Phys.* **64**, 2901 (1976).
- ⁵W. D. Laidig, G. D. Purvis III, and R. J. Bartlett, *Int. J. Quantum Chem., Quantum Chem. Symp.* **6**, 561 (1982).
- ⁶W. D. Laidig, G. D. Purvis III, and R. J. Bartlett, *Chem. Phys. Lett.* **97**, 209 (1983).
- ⁷P. Pulay, S. Saebø, and W. Meyer, *J. Chem. Phys.* **81**, 1901 (1984).
- ⁸W. Meyer, R. Ahlrichs, and C. E. Dykstra, in *Advanced Theories and Computational Approaches to the Electronic Structure of Molecules*, edited by C. E. Dykstra and R. J. Bartlett (Reidel, Dordrecht, 1984), p. 19.
- ⁹M. Takahashi and J. Paldus, *Phys. Rev. B* **31**, 5121 (1985).
- ¹⁰W. D. Laidig, G. D. Purvis III, and R. J. Bartlett, *J. Phys. Chem.* **89**, 2161 (1985).
- ¹¹G. Stollhoff and P. Vasilopoulos, *J. Chem. Phys.* **84**, 2744 (1986).
- ¹²B. Kirtman and C. E. Dykstra, *J. Chem. Phys.* **85**, 2791 (1986).
- ¹³W. Meyer and L. Frommhold, *Phys. Rev. A* **33**, 3807 (1986).
- ¹⁴J. Almöf, *Chem. Phys. Lett.* **181**, 319 (1991).
- ¹⁵S. Saebø, *Int. J. Quantum Chem.* **42**, 217 (1992).
- ¹⁶M. Häser and J. Almöf, *J. Chem. Phys.* **96**, 489 (1992).
- ¹⁷M. Häser, *Theor. Chim. Acta* **87**, 147 (1993).
- ¹⁸S. Saebø and P. Pulay, *Annu. Rev. Phys. Chem.* **44**, 213 (1993).
- ¹⁹B. Kirtman, *Int. J. Quantum Chem.* **55**, 103 (1995).
- ²⁰C. Hampel and H. J. Werner, *J. Chem. Phys.* **104**, 6286 (1996).
- ²¹P. E. Maslen and M. Head-Gordon, *Chem. Phys. Lett.* **283**, 102 (1998).
- ²²P. E. Maslen and M. Head-Gordon, *J. Chem. Phys.* **109**, 7093 (1998).
- ²³P. Y. Ayala and G. E. Scuseria, *J. Chem. Phys.* **110**, 3660 (1999).
- ²⁴G. E. Scuseria and P. Y. Ayala, *J. Chem. Phys.* **111**, 8330 (1999).
- ²⁵M. S. Lee, P. E. Maslen, and M. Head-Gordon, *J. Chem. Phys.* **112**, 3592 (2000).
- ²⁶S. Hirata, I. Grabowski, M. Tobita, and R. J. Bartlett, *Chem. Phys. Lett.* **345**, 475 (2001).
- ²⁷M. Schütz and H. J. Werner, *J. Chem. Phys.* **114**, 661 (2001).
- ²⁸J. P. Malrieu, P. Claverie, and S. Diner, *Theor. Chim. Acta* **13**, 18 (1969).
- ²⁹S. F. Boys, *Rev. Mod. Phys.* **32**, 296 (1960).
- ³⁰C. Edmiston and K. Ruedenberg, *Rev. Mod. Phys.* **35**, 457 (1963).
- ³¹J. Pipek and P. G. Mezey, *J. Chem. Phys.* **90**, 4916 (1989).
- ³²N. Flocke and R. J. Bartlett, *Chem. Phys. Lett.* **367**, 80 (2003).
- ³³A. E. Reed, R. B. Weinstock, and F. Weinhold, *J. Chem. Phys.* **83**, 735 (1985).
- ³⁴J. P. Foster and F. Weinhold, *J. Am. Chem. Soc.* **102**, 7211 (1980).
- ³⁵A. E. Reed, L. A. Curtiss, and F. Weinhold, *Chem. Rev.* **88**, 899 (1988).
- ³⁶F. Weinhold, "Natural bond orbital methods," in *Encyclopedia of Computational Chemistry*, edited by P. v. R. Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollman, H. F. Schaefer III, and P. R. Schreiner (Wiley, Chichester, 1998), Vol. 3, pp. 1792–1811.
- ³⁷D. J. Thouless, *The Quantum Mechanics of Many-Body Systems* (Academic, New York, 1961).
- ³⁸J. Paldus, in *Methods in Computational Molecular Physics, Series B*, edited by S. Wilson and G. H. F. Dierksen (Plenum, New York, 1991), Vol. 293.
- ³⁹T. H. Dunning, *J. Chem. Phys.* **53**, 2823 (1970); **90**, 1007 (1989).
- ⁴⁰Basis sets were obtained from the Extensible Computational Chemistry Environment Basis Set Database, Version 7/30/02, as developed and distributed by the Molecular Science Computing Facility, Environmental and Molecular Sciences Laboratory which is part of the Pacific Northwest Laboratory, P.O. Box 999, Richland, Washington 99352, USA, and funded by the U.S. Department of Energy.
- ⁴¹QCPACK, a highly modular Quantum Chemical PACKAge written by Norbert Flocke. The package performs standard *ab initio* calculations and has SCF, CI, and CC capabilities together with NAO and NBO generation programs.
- ⁴²H. Stoll, *Phys. Rev. B* **46**, 6700 (1992).
- ⁴³R. J. Bartlett, *Annu. Rev. Phys. Chem.* **32**, 359 (1981).
- ⁴⁴L. Bytautas and K. Ruedenberg, *Mol. Phys.* **100**, 757 (2002).
- ⁴⁵G. Stollhoff and K. P. Bohnen, *Phys. Rev. B* **37**, 4678 (1988).
- ⁴⁶S. Fahy, X. W. Wang, and S. G. Louie, *Phys. Rev. B* **42**, 3503 (1990).