

Perturbed atoms in molecules and solids: The PATMOS model

Inge Røeggen and Bin Gao

Citation: *J. Chem. Phys.* **139**, 094104 (2013); doi: 10.1063/1.4818577

View online: <http://dx.doi.org/10.1063/1.4818577>

View Table of Contents: <http://jcp.aip.org/resource/1/JCPSA6/v139/i9>

Published by the [AIP Publishing LLC](#).

Additional information on *J. Chem. Phys.*

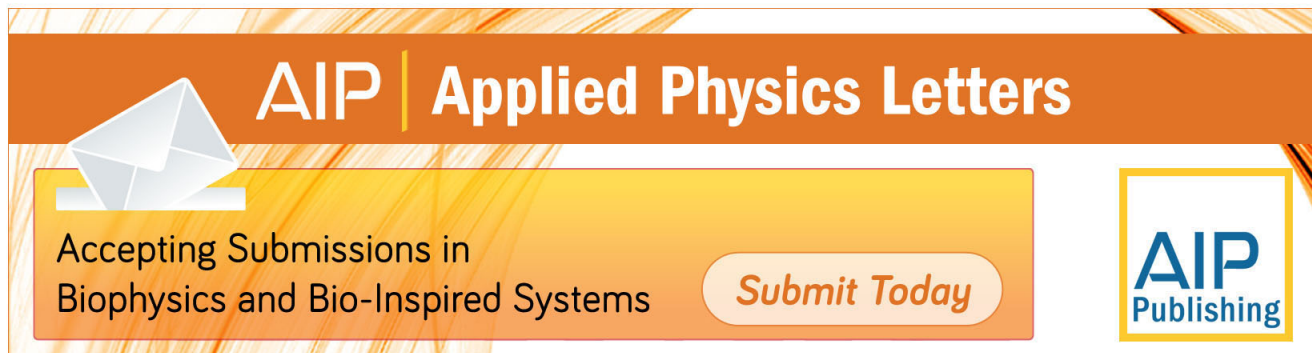
Journal Homepage: <http://jcp.aip.org/>

Journal Information: http://jcp.aip.org/about/about_the_journal

Top downloads: http://jcp.aip.org/features/most_downloaded

Information for Authors: <http://jcp.aip.org/authors>

ADVERTISEMENT



AIP | Applied Physics Letters

Accepting Submissions in
Biophysics and Bio-Inspired Systems

Submit Today

AIP
Publishing

Perturbed atoms in molecules and solids: The PATMOS model

Inge Røeggen^{1,2,a)} and Bin Gao¹

¹Centre for Theoretical and Computational Chemistry (CTCC), Department of Chemistry, University of Tromsø, N-9037 Tromsø, Norway

²Department of Physics and Technology, University of Tromsø, N-9037 Tromsø, Norway

(Received 13 March 2013; accepted 31 July 2013; published online 3 September 2013)

A new computational method for electronic-structure studies of molecules and solids is presented. The key element in the new model – denoted the perturbed atoms in molecules and solids model – is the concept of a perturbed atom in a complex. The basic approximation of the new model is unrestricted Hartree Fock (UHF). The UHF orbitals are localized by the Edmiston-Ruedenberg procedure. The perturbed atoms are defined by distributing the orbitals among the nuclei in such a way that the sum of the intra-atomic UHF energies has a minimum. Energy corrections with respect to the UHF energy, are calculated within the energy incremental scheme. The most important three- and four-electron corrections are selected by introducing a modified geminal approach. Test calculations are performed on N₂, Li₂, and parallel arrays of hydrogen atoms. The character of the perturbed atoms is illustrated by calculations on H₂, CH₄, and C₆H₆. © 2013 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4818577>]

I. INTRODUCTION

The atomic hypothesis of Democritus, refined and confirmed by the work of Dalton and Boltzmann in the 19th century, is perhaps one of the most fruitful scientific ideas ever formulated. Feynman¹ asks the following question in volume one of his well-known books *The Feynman Lectures on Physics*. What statement would contain most information in the fewest words? His answer is the following: “I believe it is the *atomic hypothesis* – that all things are made of atoms – little particles that move around in perpetual motion, attracting each other when they are a small distance apart, but repelling upon being squeezed into each other. In that one sentence there is an enormous amount of information about the world, if just a little thinking and imagination are applied.”

In chemistry and physics, a huge body of experimental results can be interpreted by the notion that matter consists of atoms. In spite of this tremendous success, the fundamental theory of chemistry – quantum mechanics – cannot rigorously support the notion of a well-defined state of an atom in a complex. The reason is easy to understand. Quantum mechanics is a holistic theory. Hence, even the states of non-interacting subsystems in a complex – if previously entangled – are not defined. A trivial case is the dissociation of the hydrogen molecule in the ground state. This particular feature of quantum mechanics is usually denoted Einstein-Podolsky-Rosen correlations.² Nevertheless, we do not consider a proper wave function for the whole universe. Non-interacting complexes are described by simple products of wave functions.

The first attempt to identify an atomic state in a molecular wave function can be traced back to the work of Moffitt.³ By using properly antisymmetrized product functions of atomic wave functions for the isolated subsystems as zero-order functions, he suggested in principle a perturbation approach for

the molecular problem where the interaction between the atoms was treated as a perturbation. In this approach, the atoms in a sense keep their individuality during molecular formation. However, it seems to be extremely difficult to apply this approach to a solid.

The one-electron density is the basis for several approaches defining atoms in a complex. In density functional theory,^{4–8} an atom in a molecule is defined by the following suppositions: First, the sum of atomic densities is equal to the molecular density. Second, the atoms are minimally promoted from their ground state. Third, the atoms in the molecule have equal chemical potentials. By using this approach, Palke⁵ and Guse⁶ have performed calculations identifying the hydrogen atom in the hydrogen molecule. Rychlewski and Parr⁹ introduced a wave function approach for defining an atom in molecule. Their approach was restricted to homonuclear diatomic molecules.

Bader and co-workers^{10–17} define atoms in a molecule by partitioning the one-electron density. The key element is the zero-flux condition yielding closed surfaces separating different parts of the real space. In this approach, an atom is a region of the real space that contains a single nucleus. Properties of the atoms, such as electron population and energies, can then be calculated by integrating over the appropriate part of the real space. Bader's^{10–17} theory makes it also possible to obtain a unique network of bonds for each geometry. This network is represented by a molecular graph. A molecular structure is defined as an equivalence class of molecular graphs. Bader's^{10–17} definition of an atom in a molecule is both creative and fruitful. However, it is very different from the common connotation that an atom is many-particle system comprising a nucleus and a number of electrons.

The present work has its origin in a conviction that the most appropriate root function in an *ab initio* calculation of the electronic structure of a complex, is likely to be based on one spatial orbital for each valence electron participating

^{a)}Electronic mail: inge.roeggen@uit.no

in the bonding. Hence, the bonding electrons should ideally be associated with a group function of the spin optimized Hartree-Fock (SOHF) type.^{18–20} All possible fragmentation of a complex can be qualitatively correctly described by this type of wave function. Furthermore, for a huge class of molecules and solids, the SOHF-orbitals are essentially distorted or perturbed atomic type orbitals. Accordingly, they can be assigned to the atoms of the complex. Unfortunately, the determination of the SOHF group function is numerically very demanding for electron groups with more than a few electrons, and practically impossible for electron groups with more than, say, 14–16 electrons.²⁰ In this work, we will therefore adopt a simpler procedure.

The basic approximation of the new model is unrestricted Hartree-Fock (UHF). The UHF orbitals can describe qualitatively any partitioning of a complex into fragments since there is a separate spatial orbital for each electron. In order to define atoms in a complex, the UHF orbitals have to be localized. The localized orbitals are chosen as the orthogonal Hartree orbitals which are expressed in terms of the optimized UHF orbitals. The orbitals obtained by this procedure are attributed to the nuclei in such a way that the sum of the intra-atomic components of the UHF energy has a minimum. A nucleus and the attributed localized UHF orbitals, are then denoted a perturbed atom of the complex in question. The perturbed atom can be a neutral entity or an ion.

As in the energy incremental method,^{21–26} we include in an additive way intra-atomic correlation energy, diatomic correlation energy, and if necessary polyatomic correlation corrections.

Within this framework, a complex may be considered as a collection of interacting perturbed atoms. This model is therefore denoted the PATMOS model (Perturbed AToms in MOleculEs and Solids).

In this work, we will argue that the concept of perturbed atoms in a complex can be useful in different ways. It can serve interpretative purposes and simplify electronic structure calculations.

The preliminary version of the model is restricted to the ground state of a complex.

The structure of the article is as follows. Section II is devoted to the theoretical framework, i.e., defining the model. Section III is concerned with UHF applied to extended systems. In Sec. IV, we present some test calculations related to the accuracy of the model. The character of the perturbed atoms is illustrated by calculations on the hydrogen molecule, methane, and benzene in Sec. V.

II. THE PATMOS MODEL

The PATMOS model is based on three basic assumptions. First, the root function is a UHF wave function. Second, a specific localization scheme for the UHF orbitals allowing the definition of perturbed atoms. Third, the energy incremental scheme.^{21–26} The total energy is written as a sum of intra-atomic terms, diatomic terms, and so on. At the highest level, the model is equivalent to a full configuration interaction (FCI) model. The FCI level is of course of limited in-

terest. In practice, it can never be reached except for systems with very few electrons and/or small basis sets.

A. Determination of the PATMOS orbitals

Localization of the UHF orbitals is a key issue in our approach. There are several localization scheme which can be used for obtaining localized orbitals. The Boys localization scheme,²⁷ the Pipek-Mezey scheme,²⁸ and the Edmiston-Ruedenberg scheme²⁹ are the most commonly adopted ones. These localization schemes give typically very similar results. The main difference is that the Pipek-Mezey scheme does not mix σ - and π -type orbitals. In this work, we prefer the Edmiston-Ruedenberg scheme since it is equivalent to obtaining a set of basis set restricted Hartree orbitals.

1. A basis set restricted Hartree model

Let $\{\psi_i^{\text{UHF};\alpha}; 1 \leq i \leq N_\alpha\}$ and $\{\psi_i^{\text{UHF};\beta}; 1 \leq i \leq N_\beta\}$ denote, respectively, the optimized α - and β -type UHF orbitals. Similarly, we have Hartree orbitals: $\{\psi_i^{\text{H};\alpha}; 1 \leq i \leq N_\alpha\}$ and $\{\psi_i^{\text{H};\beta}; 1 \leq i \leq N_\beta\}$. The basis set restricted Hartree orbitals are orthonormal, and they are expressed in terms of the UHF orbitals by a unitary transformation

$$\psi_i^{\text{H};\alpha} = \sum_{k=1}^{N_\alpha} \mathbf{U}_{ki}^\alpha \psi_k^{\text{UHF};\alpha}, \quad (1)$$

and

$$\psi_i^{\text{H};\beta} = \sum_{k=1}^{N_\beta} \mathbf{U}_{ki}^\beta \psi_k^{\text{UHF};\beta}. \quad (2)$$

The Hartree wave function and the corresponding Hartree energy

$$\Psi^{\text{Hartree}} = \left(\prod_{i=1}^{N_\alpha} \psi_i^{\text{H};\alpha} \right) \left(\prod_{j=1}^{N_\beta} \psi_j^{\text{H};\beta} \right), \quad (3)$$

$$\begin{aligned} E^{\text{Hartree}} &= \sum_{i=1}^{N_\alpha} \langle \psi_i^{\text{H};\alpha} | \mathbf{h} \psi_i^{\text{H};\alpha} \rangle + \sum_{i=1}^{N_\beta} \langle \psi_i^{\text{H};\beta} | \mathbf{h} \psi_i^{\text{H};\beta} \rangle \\ &+ \sum_{i=1}^{N_\alpha} \sum_{j=1}^{N_\beta} [\psi_i^{\text{H};\alpha} \psi_i^{\text{H};\alpha} | \psi_j^{\text{H};\beta} \psi_j^{\text{H};\beta}] \\ &+ \sum_{i < j}^{N_\alpha} [\psi_i^{\text{H};\alpha} \psi_i^{\text{H};\alpha} | \psi_j^{\text{H};\alpha} \psi_j^{\text{H};\alpha}] \\ &+ \sum_{i < j}^{N_\beta} [\psi_i^{\text{H};\beta} \psi_i^{\text{H};\beta} | \psi_j^{\text{H};\beta} \psi_j^{\text{H};\beta}]. \end{aligned} \quad (4)$$

In Eq. (4), \mathbf{h} denotes the one-electron Hamiltonian, and Mulliken notation is adopted for the two-electron integrals. Since the first three sums in Eq. (4) are invariant by unitary transformations, it follows that minimizing E^{Hartree} is equivalent to minimizing the Coulombic repulsion energies associated with each set of Hartree orbitals. Hence, the basis set

restricted Hartree orbitals are identical to the localized orbitals obtained by the Edmiston-Ruedenberg localization scheme.

2. Definition of perturbed atoms

By having constructed and tested several approaches for defining perturbed atoms in a complex, we arrived at the following unequivocal procedure based on minimizing the sum of the intra-atomic components of the UHF energy with respect to the distribution of localized spin orbitals:

$$\begin{aligned}
 E_{\text{intra}}^{\text{UHF}} = & \sum_{A=1}^{N_{\text{atoms}}} \left\{ \sum_{i=1}^{N_{\alpha}^A} \langle \psi_i^{A;\alpha} | \mathbf{h}_A \psi_i^{A;\alpha} \rangle + \sum_{i=1}^{N_{\beta}^A} \langle \psi_i^{A;\beta} | \mathbf{h}_A \psi_i^{A;\beta} \rangle \right. \\
 & + \sum_{i < j}^{N_{\alpha}^A} ([\psi_i^{A;\alpha} \psi_i^{A;\alpha} | \psi_j^{A;\alpha} \psi_j^{A;\alpha}] \\
 & - [\psi_i^{A;\alpha} \psi_j^{A;\alpha} | \psi_i^{A;\alpha} \psi_j^{A;\alpha}]) \\
 & + \sum_{i < j}^{N_{\beta}^A} ([\psi_i^{A;\beta} \psi_i^{A;\beta} | \psi_j^{A;\beta} \psi_j^{A;\beta}] \\
 & - [\psi_i^{A;\beta} \psi_j^{A;\beta} | \psi_i^{A;\beta} \psi_j^{A;\beta}]) \\
 & \left. + \sum_{i=1}^{N_{\alpha}^A} \sum_{j=1}^{N_{\beta}^A} [\psi_i^{A;\alpha} \psi_i^{A;\alpha} | \psi_j^{A;\beta} \psi_j^{A;\beta}] \right\}, \quad (5)
 \end{aligned}$$

where \mathbf{h}_A is the one-electron Hamiltonian associated with nucleus A , charge Z_A , and nuclear position \mathbf{R}_A , i.e.,

$$\mathbf{h}_A(\mathbf{r}) = -\frac{1}{2}\nabla^2 - \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}|}. \quad (6)$$

The spin orbitals $\{\psi_i^{A;\alpha}; 1 \leq i \leq N_{\alpha}^A\}$ and $\{\psi_i^{A;\beta}; 1 \leq i \leq N_{\beta}^A\}$ are localized spin orbitals associated with nucleus A .

The minimization procedure runs over all possible distributions of the spin orbitals with respect to the nuclei. To each distribution of the spin orbitals, there is associated a value of the functional $E_{\text{intra}}^{\text{UHF}}$. The optimal distribution is then the one which has the lowest value of $E_{\text{intra}}^{\text{UHF}}$. This particular distribution defines the perturbed atoms. A perturbed atom is then characterized by a nucleus and a set of spin orbitals, i.e., a set of α -type orbitals and/or a set of β -type orbitals. The number of spin orbitals associated with each nucleus, i.e., $N_{\alpha}^A + N_{\beta}^A$, determines whether the perturbed atom is a neutral entity or an ion.

The number of different distributions of the spin orbitals might be huge for a large molecule. However, in practice only a small fraction needs to be considered. First, we start by calculating the charge centroids of the spin orbitals. Then we calculate the distance between a charge centroid and the position of all nuclei. The spin orbital in question is then associated with the closest nucleus, but with a restriction that a neutral molecule should have neutral atoms in the initial distribution. Second, the core orbitals are kept fixed. Interchanging core orbitals between atoms yields a distribution far away from the optimal distribution. Third, a procedure for interchanging

and/or shifting valence orbitals between neighbouring nuclei yields the optimal distribution.

An example: For the equilibrium structure of the N_2 molecule, the Edmiston-Ruedenberg localization yields three equivalent α type bond orbitals and three equivalent β type bond orbitals. The spatial parts of these orbitals coincide pairwise, i.e., overlap completely, but this does not affect the minimization procedure. With three orbitals on each nucleus, we can for these six spin orbitals construct 20 $(6!/(3!)^2)$ different distributions. Among these distributions, there will be two different distributions corresponding to the lowest value of $E_{\text{intra}}^{\text{UHF}}$: three α -type orbitals on one nucleus and three β -type orbitals on the other or vice versa. The 18 distributions characterized by a mixture of α - and β -type orbitals on each nucleus, have higher values of $E_{\text{intra}}^{\text{UHF}}$. We can choose either of the optimal distributions. When the spatial orbitals of an $\alpha\beta$ -pair starts to split, then there will be just one distribution corresponding to the minimum.

We would also like to stress that the valence orbitals associated with a nucleus are not localized on this nucleus in a strict sense of the word. However, they are essentially localized in the region between the nucleus in question and the nearest neighbouring nuclei, see Fig. 4, the benzene case discussed in Sec. V C. But the localized orbitals have the proper ‘‘asymptotic’’ character, i.e., they move with the nuclei in a fragmentation or dissociation process, see Fig. 2 in Sec. V A.

Our definition of the perturbed atoms depends on the chosen localization procedure, i.e., the Edmiston-Ruedenberg procedure. In principle, it is possible to eliminate this dependence on the localization scheme. By having obtained the atoms as described, we can minimize $E_{\text{intra}}^{\text{UHF}}$ further by rotating localized orbitals belonging to different atoms. Contrary to the advocated approach, this improved procedure will also change the orbitals. The suggested procedure is computationally complicated, but feasible. It will be explored in future works.

B. Intra-atomic correlation terms

1. The general solution

Let \mathbf{H} denote the Hamiltonian of a N -electron complex where only Coulombic interactions are included

$$\mathbf{H} = \sum_{i=1}^N \mathbf{h}(\mathbf{r}_i) + \sum_{i < j}^N \frac{1}{r_{ij}}, \quad (7)$$

$$\mathbf{h}(\mathbf{r}) = -\frac{1}{2}\nabla^2 - \sum_{A=1}^{N_{\text{atoms}}} \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}|}. \quad (8)$$

In Eqs. (7) and (8), the symbols have their standard meaning. The UHF wave function is in our spin orbital notation

$$\Psi^{\text{UHF}} = \frac{1}{\sqrt{N!}} \det\{\psi_1 \psi_2 \dots \psi_N\}, \quad (9)$$

and the corresponding UHF energy

$$E^{\text{UHF}} = \langle \Psi^{\text{UHF}} | \mathbf{H} \Psi^{\text{UHF}} \rangle. \quad (10)$$

In solving the Hartree-Fock problem, we also obtain a set of virtual spin orbitals $\{\psi_a; N+1 \leq a \leq M\}$, where M is twice the number of spatial basis functions. Let $\Phi_{(A;i_1)}^{a_1}$ denote a Slater determinant where the occupied spin orbital $\psi_{i_1}^A$ is replaced by the virtual orbital ψ_{a_1} , $\Phi_{(A;i_1 i_2)}^{a_1 a_2}$ a Slater determinant where the two occupied spin orbitals $\psi_{i_1}^A$ and $\psi_{i_2}^A$ are replaced by, respectively, the virtual orbitals ψ_{a_1} and ψ_{a_2} , and similar for higher order excitations associated with atom A . The FCI expansion for intra-atomic correlation for atom A has then the following form:

$$\begin{aligned} \Psi_A^{\text{FCI}} = & c_0 \Psi^{\text{UHF}} + \sum_{i_1=1}^{N^A} \sum_{a_1=N+1}^M c_{i_1}^{a_1} \Phi_{(A;i_1)}^{a_1} \\ & + \sum_{i_1 < i_2}^{N^A} \sum_{N+1 \leq a_1 < a_2}^M c_{i_1 i_2}^{a_1 a_2} \Phi_{(A;i_1 i_2)}^{a_1 a_2} \\ & + \dots + \sum_{N+1 \leq a_1 < a_2 < \dots < a_{N^A}}^M c_{1,2,\dots,N^A}^{a_1 a_2 \dots a_{N^A}} \Phi_{(A;1,2,\dots,N^A)}^{a_1 a_2 \dots a_{N^A}}. \end{aligned} \quad (11)$$

The coefficients of Eq. (11) is in principle determined by the FCI eigenvalue equation

$$\mathbf{H} \Psi_A^{\text{FCI}} = \lambda_A^{\text{FCI}} \Psi_A^{\text{FCI}}. \quad (12)$$

The intra-atomic correlation energy is defined as

$$E_A^{\text{FCI}} = \lambda_A^{\text{FCI}} - E^{\text{UHF}}. \quad (13)$$

It is well known that the solution of Eq. (12), from a computational point of view, is impossible except for systems with a small number of electrons and/or a small basis sets. Hence, approximative solutions of this equation are of paramount importance.

2. Numerical models

There are two problems to be attacked: to reduce the large number of virtual orbitals and to select a feasible correlation method. We will first address the basis set problem.

By construction, the occupied orbitals of atom A , i.e., $\{\psi_i^A; 1 \leq i \leq N^A\}$, are localized in the vicinity of the nucleus of atom A . Hence, intra-atomic correlation can to a very good approximation be described by a modified one center expansion.

In the following, we have to distinguish between α and β type spin orbitals. The spatial part of these orbitals are denoted, respectively, $\{\psi_i^\alpha; 1 \leq i \leq N_\alpha\}$ and $\{\psi_i^\beta; 1 \leq i \leq N_\beta\}$. We define a modified set of one-center functions for virtual α type orbitals

$$\hat{\chi}_\mu^A = \chi_\mu^A - P_{\text{occ}}^\alpha \chi_\mu^A, \quad 1 \leq \mu \leq m_A. \quad (14)$$

In Eq. (14), $\{\chi_\mu^A; 1 \leq \mu \leq m_A\}$ is the basis function centered on nucleus A , and P_{occ}^α is the projection operator defined by the occupied α type orbitals, i.e.,

$$P_{\text{occ}}^\alpha = \sum_{i=1}^{N_\alpha} |\psi_i^\alpha\rangle \langle \psi_i^\alpha|. \quad (15)$$

We then diagonalize the overlap matrix $(\langle \hat{\chi}_\mu^A | \hat{\chi}_\nu^A \rangle)$, and select the eigenfunctions corresponding to the $(m_A - N_\alpha^A)$ largest eigenvalues (N_α^A is the number of α type spin orbitals of atom A). By this procedure, we obtain the same number of virtual α type spin orbitals as we have for the isolated atom. The virtual orbital space for the atom in the complex is (slightly) distorted compared with the virtual space for the isolated atom. However, this distortion is in a sense a physical effect. It is due to the presence of the partner atoms. We can also notice that this procedure will eliminate a basis set superposition error (BSSE) at the correlation level. The virtual β type orbitals are obtained in a similar way.

As for an electron correlation method, there are several options: perturbation theory, conventional configuration interaction, coupled cluster methods, and the energy incremental scheme. In the first version of the computational implementation of the PATMOS model, we choose the energy incremental scheme. We shall use a modification of Nesbet's²¹ original formulation of the energy incremental scheme. In Nesbet's²¹ approach, the correlation energy is a sum of one-electron corrections, two-electron corrections, three-electron corrections, and so on. For an N -electron system with UHF wave function

$$\Psi^{\text{UHF}} = \det\{\psi_1 \psi_2 \dots \psi_N\}, \quad (16)$$

and where $\{\psi_i; 1 \leq i \leq N\}$ are the occupied spin orbitals, the correlation energy is simply written as

$$\epsilon^{\text{corr}} = \sum_{i=1}^N \epsilon_i + \sum_{i < j}^N \epsilon_{ij} + \sum_{i < j < k}^N \epsilon_{ijk} + \sum_{i < j < k < l}^N \epsilon_{ijkl} + \dots \quad (17)$$

The energy corrections are obtained by partial FCI calculations, i.e., one-electron FCI calculations, two-electron FCI calculations, and so on. For a fully optimized UHF root function, the orbital corrections $\{\epsilon_i; 1 \leq i \leq N\}$ are zero.

One problem with a straightforward application of this approach is the huge number of FCI calculations which are required. Furthermore, a large number of three- and four-electron corrections are very small. Hence, they can be neglected. In this work, we shall devise a computational strategy where we include the most important of the three- and four-electron corrections. A key element in this computational scheme is the introduction of UHF geminals. Our approach has strong similarity to the extended geminal models introduced by Røeggen^{22,26} and the form of the incremental scheme advocated by Stoll and co-workers²³⁻²⁵ for closed shell systems. The more recent work by Bytautas and Ruedenberg³⁰ is also relevant in this context.

For atom A , we have an effective Hamiltonian

$$\mathbf{H}_{\text{eff}}^A = \sum_{i=1}^{N^A} \mathbf{h}_{\text{eff}}^A(\mathbf{r}_i) + \sum_{i < j}^{N^A} \frac{1}{r_{ij}}, \quad (18)$$

where the effective one-electron Hamiltonian is given by

$$\mathbf{h}_{\text{eff}}^A(\mathbf{r}_i) = \mathbf{h}(\mathbf{r}_i) + \sum_{B \neq A}^{N_{\text{atoms}}} \sum_{j=1}^{N^B} (\mathbf{J}_j^B - \mathbf{K}_j^B). \quad (19)$$

In Eq. (19), $\mathbf{h}(\mathbf{r}_i)$ is the one-electron for the complex considered, i.e., Eq. (8), and \mathbf{J}_j^B and \mathbf{K}_j^B are, respectively, Coulomb

and exchange operators derived from the spin orbital ψ_j^B of atom **B**. The UHF wave function for atom **A** is written as

$$\begin{aligned} \Psi_A^{\text{UHF}} &= \det \left\{ \psi_1^{A,\alpha} \psi_1^{A,\beta} \cdots \psi_{N_\beta^A}^{A,\alpha} \psi_{N_\beta^A}^{A,\beta} \psi_{N_\beta^A+1}^{A,\alpha} \cdots \psi_{N_\alpha^A}^{A,\alpha} \right\} \\ &= \det \left\{ \left(\prod_{r=1}^{N_\beta^A} \Lambda_{A;r}^{\text{UHF}} \right) \psi_{N_\beta^A+1}^{A,\alpha} \cdots \psi_{N_\alpha^A}^{A,\alpha} \right\}. \end{aligned} \quad (20)$$

In Eq. (20), $\{\psi_i^{A,\alpha}; 1 \leq i \leq N_\alpha^A\}$ denote the α -type spin orbitals and $\{\psi_j^{A,\beta}; 1 \leq j \leq N_\beta^A\}$ the β -type spin orbitals of atom **A**. Furthermore, we assume $N_\alpha^A > N_\beta^A$. The UHF geminal $\Lambda_{A;r}^{\text{UHF}}$ is simply the product function

$$\Lambda_{A;r}^{\text{UHF}} = \psi_r^{A,\alpha} \psi_r^{A,\beta}. \quad (21)$$

The ordering of the orbitals is such that we have a maximum overlap between the spatial orbitals of the spin orbitals of a geminal.

One- and two-electron corrections are calculated as in the original Nesbet's scheme. By using the general spin orbital notation, i.e., not distinguishing between α and β spin orbitals, we have the following expansion for the correction of spin orbital ψ_j :

$$\Psi_{A;j}^{\text{FCI}} = c_0 \Psi_A^{\text{UHF}} + \sum_a^{\text{virt}} c_j^a \Phi_{(A;j)}^a, \quad (22)$$

where the sum runs over the orbitals in the modified virtual spaces, and $\Phi_{(A;j)}^a$ is a Slater determinant where the occupied orbital j is replaced by the virtual orbital a . The corresponding eigenvalue equation

$$\mathbf{H}_{\text{eff}}^A \Psi_{A;j}^{\text{FCI}} = \lambda_{A;j}^{\text{FCI}} \Psi_{A;j}^{\text{FCI}}. \quad (23)$$

The orbital correction

$$\epsilon_{A;j}^{\text{corr}} = \lambda_{A;j}^{\text{FCI}} - E_A^{\text{UHF}}, \quad (24)$$

where

$$E_A^{\text{UHF}} = \langle \Psi_A^{\text{UHF}} | \mathbf{H}_{\text{eff}}^A | \Psi_A^{\text{UHF}} \rangle. \quad (25)$$

The two-electron correction $\epsilon_{A;ij}^{\text{corr}}$ is derived from the expansion

$$\begin{aligned} \Psi_{A;ij}^{\text{FCI}} &= c_0 \Psi_A^{\text{UHF}} + \sum_a^{\text{virt}} c_i^a \Phi_{(A;i)}^a + \sum_a^{\text{virt}} c_j^a \Phi_{(A;j)}^a \\ &+ \sum_{a < b}^{\text{virt}} c_{ij}^{ab} \Phi_{(A;ij)}^{ab}. \end{aligned} \quad (26)$$

The corresponding eigenvalue equation

$$\mathbf{H}_{\text{eff}}^A \Psi_{A;ij}^{\text{FCI}} = \lambda_{A;ij}^{\text{FCI}} \Psi_{A;ij}^{\text{FCI}}. \quad (27)$$

Two-electron correction

$$\epsilon_{A;ij}^{\text{corr}} = \lambda_{A;ij}^{\text{FCI}} - E_A^{\text{UHF}} - \epsilon_{A;i}^{\text{corr}} - \epsilon_{A;j}^{\text{corr}}. \quad (28)$$

When $N_\alpha^A - N_\beta^A \geq 3$, there are two different sets of three-electron corrections: intra-“valence” and geminal-“valence” corrections. In a general spin orbital notation, we have in both case, a three-electron cluster associated with occupied orbitals $\{\psi_i^A, \psi_j^A, \psi_k^A\}$. In order to obtain a computational feasible scheme, we perform a truncation of the

virtual space. This can be done by expressing the wave functions $\Psi_{A;ij}^{\text{FCI}}$ in terms of natural orbitals (NOs) and select the most important NOs. For the three-electron cluster (i, j, k) , we construct different sets of NOs which we merge into one set of α -type virtual orbitals and of one set of β -type virtual orbitals. A detailed description of this procedure is given in Subsection II G. In this truncated virtual space, we recalculate one- and two-electron corrections: $\{\tilde{\epsilon}_{A;i}^{\text{corr}}, \tilde{\epsilon}_{A;j}^{\text{corr}}, \tilde{\epsilon}_{A;k}^{\text{corr}}\}$ and $\{\tilde{\epsilon}_{A;ij}^{\text{corr}}, \tilde{\epsilon}_{A;ik}^{\text{corr}}, \tilde{\epsilon}_{A;jk}^{\text{corr}}\}$. For the three-electron cluster (i, j, k) , we have the following FCI equation:

$$\mathbf{H}_{\text{eff}}^A \tilde{\Psi}_{A;ijk}^{\text{FCI}} = \tilde{\lambda}_{A;ijk}^{\text{FCI}} \tilde{\Psi}_{A;ijk}^{\text{FCI}}. \quad (29)$$

The wave function $\tilde{\Psi}_{A;ijk}^{\text{FCI}}$ is just an extension of Eq. (26) including triple excitations. The \sim sign implies that all terms refer to the truncated virtual space. The three-electron correlation correction for this particular cluster

$$\begin{aligned} \epsilon_{A;ijk}^{\text{corr}} &= \tilde{\lambda}_{A;ijk}^{\text{FCI}} - E_A^{\text{UHF}} - \tilde{\epsilon}_{A;i}^{\text{corr}} - \tilde{\epsilon}_{A;j}^{\text{corr}} - \tilde{\epsilon}_{A;k}^{\text{corr}} \\ &- \tilde{\epsilon}_{A;ij}^{\text{corr}} - \tilde{\epsilon}_{A;ik}^{\text{corr}} - \tilde{\epsilon}_{A;jk}^{\text{corr}}. \end{aligned} \quad (30)$$

For the inter-geminal corrections, we calculate corrections which also include the appropriate three-electron corrections. For the four-electron cluster associated with geminal product

$$\Lambda_{A;r}^{\text{UHF}} \Lambda_{A;s}^{\text{UHF}} = \psi_r^{A,\alpha} \psi_r^{A,\beta} \psi_s^{A,\alpha} \psi_s^{A,\beta} = \psi_i \psi_j \psi_k \psi_l, \quad (31)$$

we construct the relevant set of NOs and recalculate one- and two-electron corrections within the truncated virtual space: $\{\tilde{\epsilon}_{A;i}^{\text{corr}}, \tilde{\epsilon}_{A;j}^{\text{corr}}, \tilde{\epsilon}_{A;k}^{\text{corr}}, \tilde{\epsilon}_{A;l}^{\text{corr}}\}$ and $\{\tilde{\epsilon}_{A;ij}^{\text{corr}}, \tilde{\epsilon}_{A;ik}^{\text{corr}}, \tilde{\epsilon}_{A;il}^{\text{corr}}, \tilde{\epsilon}_{A;jk}^{\text{corr}}, \tilde{\epsilon}_{A;jl}^{\text{corr}}, \tilde{\epsilon}_{A;kl}^{\text{corr}}\}$. For the four-electron cluster (i, j, k, l) , we have the following FCI equation:

$$\mathbf{H}_{\text{eff}}^A \tilde{\Psi}_{A;ijkl}^{\text{FCI}} = \tilde{\lambda}_{A;ijkl}^{\text{FCI}} \tilde{\Psi}_{A;ijkl}^{\text{FCI}}. \quad (32)$$

The wave function $\tilde{\Psi}_{A;ijkl}^{\text{FCI}}$ is an expansion comprising Ψ_A^{UHF} and the determinants obtained by all single, double, triple, and quadruple excitations from (i, j, k, l) into the truncated virtual space. The inter-geminal correction, including all three-electron corrections for this cluster and the four-electron correction, is given by

$$\begin{aligned} \epsilon_{A;(\text{gem},r),(\text{gem},s)}^{\text{corr}} &= \tilde{\lambda}_{A;ijkl}^{\text{FCI}} - E_A^{\text{UHF}} - \tilde{\epsilon}_{A;i}^{\text{corr}} - \tilde{\epsilon}_{A;j}^{\text{corr}} - \tilde{\epsilon}_{A;k}^{\text{corr}} - \tilde{\epsilon}_{A;l}^{\text{corr}} \\ &- \tilde{\epsilon}_{A;ij}^{\text{corr}} - \tilde{\epsilon}_{A;ik}^{\text{corr}} - \tilde{\epsilon}_{A;il}^{\text{corr}} - \tilde{\epsilon}_{A;jk}^{\text{corr}} - \tilde{\epsilon}_{A;jl}^{\text{corr}} - \tilde{\epsilon}_{A;kl}^{\text{corr}}. \end{aligned} \quad (33)$$

The total correlation energy for atom **A** is then given by the following approximation:

$$\begin{aligned} E_A^{\text{corr}} &= \sum_{i=1}^{N_\alpha^A} \epsilon_{A;i}^{\text{corr}} + \sum_{i < j}^{N_\alpha^A} \epsilon_{A;ij}^{\text{corr}} + \sum_{N_\beta^A < i < j < k}^{N_\alpha^A} \epsilon_{A;ijk}^{\text{corr}} \\ &+ \sum_{r=1}^{N_\beta^A} \sum_{i=N_\beta^A+1}^{N_\alpha^A} \epsilon_{A;(\text{gem},r);i}^{\text{corr}} + \sum_{r < s}^{N_\beta^A} \epsilon_{A;(\text{gem},r),(\text{gem},s)}^{\text{corr}}. \end{aligned} \quad (34)$$

The third sum in Eq. (34) is included only when $N_\alpha^A - N_\beta^A \geq 3$, the fourth sum only when $N_\beta^A \geq 1$ and $N_\alpha^A > N_\beta^A$, and the last term only when $N_\beta^A \geq 2$. In the case when $N_\beta^A > N_\alpha^A$, we just interchange the role of α and β spin orbitals.

C. Diatomic correlation terms

1. The general solution

Let Ψ_{AB}^{FCI} represent the FCI expansion where we include all possible excitations from the occupied orbitals of both atoms **A** and **B**. The corresponding FCI equation

$$\mathbf{H}\Psi_{AB}^{\text{FCI}} = \lambda_{AB}^{\text{FCI}}\Psi_{AB}^{\text{FCI}}. \quad (35)$$

The interatomic correlation energy for the cluster (**A**, **B**) is then simply

$$E_{AB}^{\text{FCI}} = \lambda_{AB}^{\text{FCI}} - E^{\text{UHF}} - E_A^{\text{FCI}} - E_B^{\text{FCI}}. \quad (36)$$

In Eq. (36), E_A^{FCI} and E_B^{FCI} are the intra-atomic correlation energy defined in Sec. II B 1.

As in the intra-atomic case, we must in practical calculations reduce the number of virtual orbitals and choose an appropriate model for approximating the FCI energy.

2. Numerical models

The occupied orbitals of atoms **A** and **B** are localized in the vicinity of the nuclei **A** and **B**. Hence, it is sufficient to include virtual orbitals located in the same part of the physical space. Our procedure is a simple extension of the approach discussed in detail in Sec. II B 2. We have now two sets of modified one-center functions

$$\hat{\chi}_{\mu}^A = \chi_{\mu}^A - P_{\text{occ}}^{\alpha}\chi_{\mu}^A, \quad 1 \leq \mu \leq m_A, \quad (37)$$

$$\hat{\chi}_{\nu}^B = \chi_{\nu}^B - P_{\text{occ}}^{\alpha}\chi_{\nu}^B, \quad 1 \leq \nu \leq m_B. \quad (38)$$

We diagonalize the overlap matrix generated by the functions defined in Eqs. (37) and (38). Then we select the eigenfunctions corresponding to the $(m_A + m_B - N_{\alpha}^A - N_{\alpha}^B)$ largest eigenvalues. These functions are the spatial parts of the α type virtual orbitals for the diatomic cluster (**A**, **B**). A similar approach yields the β type virtual spin orbitals.

In order to calculate the diatomic correlation energy, we have to adopt a size-extensive correlation method. As in the intra-atomic case, we choose for the first implementation of the PATMOS model, the energy incremental scheme.

It is convenient to introduce an effective Hamiltonian for the diatomic cluster (**A**, **B**)

$$\mathbf{H}_{\text{eff}}^{AB} = \sum_{i=1}^{N_A+N_B} \mathbf{h}_{\text{eff}}^{AB}(\mathbf{r}_i) + \sum_{i<j}^{N_A+N_B} \frac{1}{r_{ij}}, \quad (39)$$

where the effective one-electron Hamiltonian is given by

$$\mathbf{h}_{\text{eff}}^{AB}(\mathbf{r}_i) = \mathbf{h}(\mathbf{r}_i) + \sum_{C \neq A, B} \sum_{j=1}^{N_C} (\mathbf{J}_j^C - \mathbf{K}_j^C). \quad (40)$$

In Eq. (40), $\mathbf{h}(\mathbf{r}_i)$ is the one-electron Hamiltonian defined in Eq. (8), and \mathbf{J}_j^C and \mathbf{K}_j^C are, respectively, Coulomb and exchange operators derived from the spin orbital ψ_j^C of atom **C**.

The non-paired spin orbitals of atoms **A** and **B**, Eq. (20), are either of equal spin type, or different spin types. We consider first the case when we have equal type of spin orbitals,

say α type spin orbitals. The UHF wave function for cluster (**A**, **B**) is then written as

$$\begin{aligned} \Psi_{AB}^{\text{UHF}} &= \det \{ \psi_1^A \dots \psi_{N_A}^A \psi_1^B \dots \psi_{N_B}^B \} \\ &= \det \left\{ \left(\prod_{r=1}^{N_{\beta}^A} \Lambda_{A;r}^{\text{UHF}} \right) \psi_{N_{\beta}^A+1}^{A,\alpha} \dots \psi_{N_{\alpha}^A}^{A,\alpha} \right. \\ &\quad \left. \times \left(\prod_{s=1}^{N_{\beta}^B} \Lambda_{B;s}^{\text{UHF}} \right) \psi_{N_{\beta}^B+1}^{B,\alpha} \dots \psi_{N_{\alpha}^B}^{B,\alpha} \right\}. \quad (41) \end{aligned}$$

The UHF geminals are defined in Eq. (21).

The interatom two-electron corrections are calculated as in the intra-atomic case with appropriate change of Hamiltonian and virtual space. The two-electron correction associated with the occupied orbitals ψ_i^A and ψ_j^B is denoted $\epsilon_{(A,i);(B,j)}^{\text{corr}}$.

There are four different groups of three-electron corrections to be considered. They are associated with the following sets of occupied orbitals:

$$\{ (\psi_i^{A,\alpha}, \psi_j^{A,\alpha}, \psi_k^{B,\alpha}); N_{\beta}^A < i < j \leq N_{\alpha}^A; N_{\beta}^B < k \leq N_{\alpha}^B \},$$

$$\{ (\psi_i^{A,\alpha}, \psi_j^{B,\alpha}, \psi_k^{B,\alpha}); N_{\beta}^A < i \leq N_{\alpha}^A; N_{\beta}^B < j < k \leq N_{\alpha}^B \},$$

$$\{ (\Lambda_{A,r}^{\text{UHF}}, \psi_k^{B,\alpha}); 1 \leq r \leq N_{\beta}^A; N_{\beta}^B < k \leq N_{\alpha}^B \},$$

$$\{ (\psi_i^{A,\alpha}, \Lambda_{B,s}^{\text{UHF}}); N_{\beta}^A < i \leq N_{\alpha}^A; 1 \leq s \leq N_{\beta}^B \}.$$

The procedure for calculating the three-electron correction is analogous to the one adopted for the intra-atomic case. We select appropriate NOs based on two-electron FCI wave functions, construct a linear independent set of orbitals, and calculate the relevant FCI eigenvalues. The corresponding correction terms are denoted, respectively, $\epsilon_{(A,i,j);(B,k)}^{\text{corr}}$, $\epsilon_{(A,i);(B,jk)}^{\text{corr}}$, $\epsilon_{(A;(gem,r));(B,k)}^{\text{corr}}$, $\epsilon_{(A,i);(B;(gem,s))}^{\text{corr}}$.

The interatomic inter-geminal correction are related to the geminal products $\{ \Lambda_{A,r}^{\text{UHF}} \Lambda_{B,s}^{\text{UHF}}; 1 \leq r \leq N_{\beta}^A; 1 < s \leq N_{\beta}^B \}$. The procedure for the calculation of these corrections is also in this case analogous to the one used in the intra-atomic case. The corrections are denoted $\{ \epsilon_{(A;(gem,r));(B;(gem,s))}^{\text{corr}}; 1 \leq r \leq N_{\beta}^A; 1 < s \leq N_{\beta}^B \}$.

When the non-paired spin orbitals are of different types, i.e., α type and β type or vice versa, the computational procedure for the three- and four-electron terms will be slightly different than the one described for the case of equal spin type orbitals. We assume for simplicity that $N_{\alpha}^A - N_{\beta}^A > N_{\beta}^B - N_{\alpha}^B > 0$. The UHF wave function for (**A**, **B**) can in this case be written as

$$\begin{aligned} \Psi_{AB}^{\text{UHF}} &= \det \left\{ \left(\prod_{r=1}^{N_{\beta}^A} \Lambda_{A;r}^{\text{UHF}} \right) \left(\prod_{s=1}^{N_{\alpha}^B} \Lambda_{B;s}^{\text{UHF}} \right) \right. \\ &\quad \left. \times \left(\prod_{t=1}^{N_{\text{gem}}^{AB}} \Lambda_{AB;t}^{\text{UHF}} \right) \psi_{N_{\beta}^A+N_{\text{gem}}^{AB}+1}^{A,\alpha} \dots \psi_{N_{\alpha}^A}^{A,\alpha} \right\}. \quad (42) \end{aligned}$$

The interatomic geminals are constructed for non-paired orbitals of atoms A and B , i.e.,

$$\Lambda_{AB;t}^{\text{UHF}} = \psi_{N_{\beta}^A+t}^{A,\alpha} \psi_{N_{\alpha}^B+t}^{B,\beta}. \quad (43)$$

The number of interatomic geminals is $N_{\text{gem}}^{AB} = N_{\beta}^B - N_{\alpha}^B$. The numbering of orbitals is such that we obtain maximum overlap between the spatial orbitals of a geminal. In a system like N_2 , we will then have the familiar picture of a core and a lone pair geminal on each atom and three bond pair geminals. On the other hand, if we consider the (C, C) group in C_2H_6 , we will have four diatomic geminals. However, only one corresponds to the familiar C–C bond geminal. The three other geminals include orbitals which are “bonded” to different hydrogen atoms. These geminals are introduced only with the purpose of simplifying the computational procedure, i.e., reducing the number of FCI calculations.

The interatomic three-electron corrections are related to the electron groups comprising a geminal and a spin orbital, i.e., $\Lambda_{B;s}^{\text{UHF}} \psi_j^{A,\alpha}$. The calculational procedure is similar to one used when all non-paired orbitals are of the same type, i.e., all are α type spin orbitals. The corresponding energy correction is denoted $\epsilon_{(A,j);(B;(\text{gem},s))}^{\text{corr}}$.

As for the inter-geminal correction there are two types. The type related to the group $\Lambda_{A;r}^{\text{UHF}} \Lambda_{B;s}^{\text{UHF}}$, which are already considered in the equal spin case, and the type related to groups $\Lambda_{A;r}^{\text{UHF}} \Lambda_{AB;t}^{\text{UHF}}$ and $\Lambda_{B;s}^{\text{UHF}} \Lambda_{AB;t}^{\text{UHF}}$. For this last group of corrections, we have to introduce a small modification when we compare with the intra-atomic case. Let us consider the four-electron group associated with the orbitals

$$\Lambda_{A;r}^{\text{UHF}} \Lambda_{AB;t}^{\text{UHF}} = \psi_r^{A,\alpha} \psi_r^{A,\beta} \psi_{N_{\beta}^A+t}^{A,\alpha} \psi_{N_{\alpha}^B+t}^{B,\beta}. \quad (44)$$

In Eq. (44), there are three spin orbitals linked with atom A . The corresponding three-electron correction is an intra-atomic correction which is already calculated as part of intra-atomic correlation energy E_A^{corr} . When we calculate the inter-geminal correction, we must also perform a calculation of $\tilde{\epsilon}_{(A;(\text{gem},r));N_{\beta}^A+t}$ and subtract this term in the appropriate equation which corresponds to Eq. (33) of the intra-atomic case.

In the case when the UHF function of the cluster (A, B) is given by Eq. (42), we finally arrive at the following approximation for the interatomic correlation energy:

$$\begin{aligned} E_{AB}^{\text{corr}} &= \sum_{i=1}^{N^A} \sum_{j=1}^{N^B} \epsilon_{(A,i);(B,j)}^{\text{corr}} + \sum_{j=N_{\beta}^A+N_{\text{gem}}^{AB}+1}^{N_{\alpha}^A} \sum_{s=1}^{N_{\alpha}^B} \epsilon_{(A,j);(B,(\text{gem},s))}^{\text{corr}} \\ &+ \sum_{r=1}^{N_{\beta}^A} \sum_{t=1}^{N_{\text{gem}}^{AB}} \epsilon_{(A,(\text{gem},r));(AB,(\text{gem},t))}^{\text{corr}} \\ &+ \sum_{s=1}^{N_{\alpha}^B} \sum_{t=1}^{N_{\text{gem}}^{AB}} \epsilon_{(B,(\text{gem},s));(AB,(\text{gem},t))}^{\text{corr}} \\ &+ \sum_{r=1}^{N_{\beta}^A} \sum_{s=1}^{N_{\alpha}^B} \epsilon_{(A,(\text{gem},r));(B,(\text{gem},s))}^{\text{corr}}. \end{aligned} \quad (45)$$

The terms in Eq. (45) represent the dominant contributions to the interatomic correlation energy for the (A, B) cluster.

Increased accuracy can be obtained by including some of the neglected three- and four-electron terms.

It has for a long time been recognized that a key to circumvent the strong basis set dependency in electron correlation models is to use localized orbitals.^{31–45} Such an approach was introduced as early as in 1964 by Sinanoğlu³¹ and Nesbet.³² However, the real breakthrough for local correlation models can be attributed to an idea proposed by Pulay,³³ and later implemented by Sæbø and Pulay^{34–37} for Møller–Plessett perturbation theory. Their approach implies that excitations from localized molecular orbitals are restricted to subspaces of projected atomic orbitals which are spatially close to the localized orbitals. The dimension for the local basis set for interorbital correlation is chosen equal to 70,³⁴ independent of the size of the system.³⁶ This is to be contrasted with our approach where all orbitals pertaining to an atom or a pair of atoms are attributed a common local basis. With large basis sets, we will then have a somewhat higher accuracy, but at a higher computational cost. However, these particular local basis sets will only be used for two-electron corrections. The truncated virtual basis for three- or four-electron FCI calculations will be based on natural orbitals, Sec. II G.

D. Triatomic correlation terms

The general approach in this section would be a straightforward generalization of the procedure in Sec. II C 1. We therefore embark directly on the numerical model. The most important terms are related to the three-electron components. As in the diatomic case, we have to truncate the virtual space. Based on two-electron FCI calculations and natural orbital expansions, we can construct a truncated set of virtual orbitals. Let A, B , and C denote three atoms. We then consider the occupied orbitals $\{\psi_i^A, \psi_j^B, \psi_k^C\}$. We recalculate the relevant one- and two-electron corrections based on this truncated set of orbitals: $\tilde{\epsilon}_{A,i}^{\text{corr}}$, $\tilde{\epsilon}_{B,j}^{\text{corr}}$, $\tilde{\epsilon}_{C,k}^{\text{corr}}$, $\tilde{\epsilon}_{A,i;B,j}^{\text{corr}}$, $\tilde{\epsilon}_{A,i;C,k}^{\text{corr}}$, and $\tilde{\epsilon}_{B,j;C,k}^{\text{corr}}$. By solving the three-electron FCI problem for this particular electron group, we obtain an energy eigenvalue $\tilde{\lambda}_{A,i;B,j;C,k}^{\text{FCI}}$. Accordingly, the three-electron correction for this group of electrons, is simply

$$\begin{aligned} \epsilon_{A,i;B,j;C,k}^{\text{corr}} &= \tilde{\lambda}_{A,i;B,j;C,k}^{\text{FCI}} - E^{\text{UHF}} - \tilde{\epsilon}_{A,i}^{\text{corr}} - \tilde{\epsilon}_{B,j}^{\text{corr}} - \tilde{\epsilon}_{C,k}^{\text{corr}} \\ &- \tilde{\epsilon}_{A,i;B,j}^{\text{corr}} - \tilde{\epsilon}_{A,i;C,k}^{\text{corr}} - \tilde{\epsilon}_{B,j;C,k}^{\text{corr}}. \end{aligned} \quad (46)$$

By summing up the contributions from different electron groups, we obtain

$$E_{ABC}^{\text{corr}} = \sum_{i=1}^{N^A} \sum_{j=1}^{N^B} \sum_{k=1}^{N^C} \epsilon_{A,i;B,j;C,k} + \dots \quad (47)$$

E. Symmetry simplification

Since the PATMOS model has an additive structure, we need to calculate only symmetry independent terms. For example, when considering the equilibrium structure of benzene, C_6H_6 , there are 12 intra-atomic terms, but only 2 terms are symmetry independent. For the same system, there are 66 diatomic terms, but only 10 terms are symmetry

independent. The computation time can be substantially reduced by taking this type of symmetry into account.

As for intra-atomic correlation energies, we start by ordering the intra-atomic energies $\{E_A^{\text{UHF}}; 1 \leq A \leq N_{\text{atoms}}\}$ in different groups. The definition of the symbol E_A^{UHF} should be self-explained, but is explicitly defined in Sec. II F. Atoms belonging to the same group have UHF energies which differ by less than a certain threshold value. Then we need only to calculate the intra-atomic correlation energy for one atom of each group.

For the diatomic terms, we order the Coulomb energies between pairs of atoms, i.e., an ordering of the set $\{E_{AB}^{\text{Coul}}; 1 \leq A < B \leq N_{\text{atoms}}\}$, in different groups according to the value of the Coulomb energy. As in the intra-atomic case, we calculate the correlation energy only for one element in each group.

F. The PATMOS energy

If we write the one-electron Hamiltonian, Eq. (8), in the following way:

$$\mathbf{h}(\mathbf{r}) = -\frac{1}{2}\nabla^2 + \sum_{A=1}^{N_{\text{atoms}}} V_A(\mathbf{r}), \quad (48)$$

$$V_A(\mathbf{r}) = -\frac{Z_A}{|\mathbf{R}_A - \mathbf{r}|}, \quad (49)$$

$$\mathbf{h}_A(\mathbf{r}) = -\frac{1}{2}\nabla^2 + V_A(\mathbf{r}), \quad (50)$$

and add the electrostatic energy between the nuclei, i.e.,

$$V_{AB} = \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|}, \quad (51)$$

we obtain the following partitioning of the UHF energy including the nuclear electrostatic energy:

$$E^{\text{UHF}} + \sum_{A<B} V_{AB} = \sum_{A=1}^{N_{\text{atoms}}} E_A^{\text{UHF}} + \sum_{A<B} (E_{AB}^{\text{Coul}} + E_{AB}^{\text{exch}}), \quad (52)$$

where

$$E_A^{\text{UHF}} = \sum_{i=1}^{N^A} \langle \psi_i^A | \mathbf{h}_A \psi_i^A \rangle + \sum_{i<j}^{N^A} \{ [\psi_i^A \psi_i^A | \psi_j^A \psi_j^A] - [\psi_i^A \psi_j^A | \psi_i^A \psi_j^A] \}. \quad (53)$$

In Eq. (53), we have adopted the Mulliken notation for the two-electron integrals. The diatomic terms are given as

$$E_{AB}^{\text{Coul}} = V_{AB} + \sum_{i=1}^{N^A} \langle \psi_i^A | V_B \psi_i^A \rangle + \sum_{j=1}^{N^B} \langle \psi_j^B | V_A \psi_j^B \rangle + \sum_{i=1}^{N^A} \sum_{j=1}^{N^B} [\psi_i^A \psi_i^A | \psi_j^B \psi_j^B], \quad (54)$$

and

$$E_{AB}^{\text{exch}} = - \sum_{i=1}^{N^A} \sum_{j=1}^{N^B} [\psi_i^A \psi_j^B | \psi_i^A \psi_j^B]. \quad (55)$$

By including electron correlation, we obtain the following expression for the total energy:

$$E^{\text{PATMOS}} = E^{\text{UHF}} + \sum_{A<B} V_{AB} + E^{\text{corr}} = \sum_{A=1}^{N_{\text{atoms}}} \{ E_A^{\text{UHF}} + E_A^{\text{corr}} \} + \sum_{A<B}^{N_{\text{atoms}}} \{ E_{AB}^{\text{Coul}} + E_{AB}^{\text{exch}} + E_{AB}^{\text{corr}} \} + \sum_{A<B<C}^{N_{\text{atoms}}} E_{ABC}^{\text{corr}} + \dots \quad (56)$$

If we partition the interaction energies equally among the atoms involved, we can express the total energy as a sum of effective atomic energies

$$E^{\text{PATMOS}} = \sum_{A=1}^{N_{\text{atoms}}} E_A^{\text{eff,add}}, \quad (57)$$

where

$$E_A^{\text{eff,add}} = E_A^{\text{UHF}} + E_A^{\text{corr}} + \frac{1}{2} \sum_{B \neq A}^{N_{\text{atoms}}} (E_{AB}^{\text{Coul}} + E_{AB}^{\text{exch}} + E_{AB}^{\text{corr}}) + \frac{1}{3} \left\{ \sum_{\substack{B,C \\ A<B<C}}^{N_{\text{atoms}}} E_{ABC}^{\text{corr}} + \sum_{\substack{B,C \\ B<A<C}}^{N_{\text{atoms}}} E_{BAC}^{\text{corr}} + \sum_{\substack{B,C \\ B<C<A}}^{N_{\text{atoms}}} E_{BCA}^{\text{corr}} \right\} + \dots \quad (58)$$

In a calculation on an extended system, only the effective energies for the atoms in the unit cell need to be calculated. This is analogous to the approach adopted by Røeggen⁴⁶ in a study of the face centered cubic and hexagonal close packed structure of helium.

G. Natural orbitals for poly-electron FCI

The computation time for a FCI calculation depends crucially on the dimension of the orbital space. If m is the number of orbitals, the computation time for a two-electron FCI calculation is proportional to m^4 , and for three- and four-electron FCI calculations, it is proportional to m^5 and m^6 , respectively. The two-electron calculations can be performed with large basis sets. But similar calculations on three- and four-electron systems are prohibitively expensive. Hence, it is of paramount importance to reduce the virtual space for three- and four-electron FCI calculations. This can be achieved by using natural orbitals derived from two-electron FCI calculations.

The two-electron FCI calculations are based on either intra-atomic virtual space, Subsection II B 2, or diatomic virtual space, Subsection II C 2. In both cases, the orbital space

comprises two occupied orbitals and the virtual space. The spatial orbitals for α and β spin orbitals are denoted, respectively, $\{\psi_i^\alpha; 1 \leq i \leq m^\alpha\}$ and $\{\psi_j^\beta; 1 \leq j \leq m^\beta\}$. Two different cases must be considered: equal spin for the two electrons involved or different spin.

First, two electrons with equal spin, say α spin. The two-electron FCI wave function

$$\Phi^{\text{FCI};\alpha\alpha} = \sum_{i < j} \tilde{c}_{ij}^{\alpha\alpha} \det \{ \psi_i^\alpha \alpha \psi_j^\alpha \alpha \} = \sum_{i,j=1}^{m^\alpha} c_{ij}^{\alpha\alpha} \det \{ \psi_i^\alpha \alpha \psi_j^\alpha \alpha \}, \quad (59)$$

where

$$c_{ii}^{\alpha\alpha} = 0, \quad (60)$$

and

$$c_{ji}^{\alpha\alpha} = -c_{ij}^{\alpha\alpha} = -\frac{1}{2} \tilde{c}_{ij}^{\alpha\alpha}, \quad i < j. \quad (61)$$

Standard definition of the corresponding one-electron density

$$\rho^{\alpha\alpha}(\mathbf{x}'_1, \mathbf{x}_1) = 2 \int \Phi^{\text{FCI};\alpha\alpha}(\mathbf{x}'_1, \mathbf{x}_2)^* \Phi^{\text{FCI};\alpha\alpha}(\mathbf{x}_1, \mathbf{x}_2) d\mathbf{x}_2, \quad (62)$$

where the coordinate \mathbf{x} denotes the three spatial coordinates and the spin coordinate, i.e., $\mathbf{x} = (\mathbf{r}, \sigma)$. A simple derivation yields

$$\rho^{\alpha\alpha}(\mathbf{x}'_1, \mathbf{x}_1) = P^{\alpha\alpha}(\mathbf{r}'_1, \mathbf{r}_1) \alpha^*(\sigma'_1) \alpha(\sigma_1), \quad (63)$$

where

$$P^{\alpha\alpha} = \sum_{i,k=1}^{m^\alpha} \psi_i^\alpha(\mathbf{r}'_1)^* \psi_k^\alpha(\mathbf{r}_1) \Omega_{ik}^{\alpha\alpha}, \quad (64)$$

and

$$\Omega_{ik}^{\alpha\alpha} = 4 \sum_{j=1}^{m^\alpha} (c_{ij}^{\alpha\alpha})^* c_{kj}^{\alpha\alpha}. \quad (65)$$

The two-electron FCI wave function in the case of different spin

$$\Phi^{\text{FCI};\alpha\beta} = \sum_{i=1}^{m^\alpha} \sum_{j=1}^{m^\beta} c_{ij}^{\alpha\beta} \det \{ \psi_i^\alpha \alpha \psi_j^\beta \beta \}. \quad (66)$$

The corresponding one-electron density

$$\rho(\mathbf{x}'_1, \mathbf{x}_1) = P^{\alpha\alpha}(\mathbf{r}'_1, \mathbf{r}_1) \alpha^*(\sigma'_1) \alpha(\sigma_1) + P^{\beta\beta}(\mathbf{r}'_1, \mathbf{r}_1) \beta^*(\sigma'_1) \beta(\sigma_1), \quad (67)$$

where

$$P^{\alpha\alpha} = \sum_{i,k=1}^{m^\alpha} \psi_i^\alpha(\mathbf{r}'_1)^* \psi_k^\alpha(\mathbf{r}_1) \Omega_{ik}^{\alpha\alpha}, \quad (68)$$

$$\Omega_{ik}^{\alpha\alpha} = \sum_{j=1}^{m^\beta} (c_{ij}^{\alpha\beta})^* c_{kj}^{\alpha\beta}, \quad (69)$$

$$P^{\beta\beta} = \sum_{j,l=1}^{m^\beta} \psi_j^\beta(\mathbf{r}'_1)^* \psi_l^\beta(\mathbf{r}_1) \Omega_{jl}^{\beta\beta}, \quad (70)$$

$$\Omega_{jl}^{\beta\beta} = \sum_{i=1}^{m^\alpha} (c_{ij}^{\alpha\beta})^* c_{il}^{\alpha\beta}. \quad (71)$$

In order to construct the natural orbitals for the virtual space, we pick out the part of the density matrix referring to the virtual space. For both the $P^{\alpha\alpha}$ and $P^{\beta\beta}$ matrices, we have the same formula (when the orbitals are real)

$$P = \sum_{i,j=1}^{m^{\text{virt}}} \Omega_{ij}^{\text{virt}} \psi_i^{\text{virt}} \psi_j^{\text{virt}}. \quad (72)$$

In matrix notation,

$$P = (\boldsymbol{\psi}^{\text{virt}})^T \boldsymbol{\Omega}^{\text{virt}} \boldsymbol{\psi}^{\text{virt}}, \quad (73)$$

$$\boldsymbol{\psi}^{\text{virt}} = \begin{pmatrix} \psi_1^{\text{virt}} \\ \psi_2^{\text{virt}} \\ \vdots \\ \psi_{m^{\text{virt}}}^{\text{virt}} \end{pmatrix}. \quad (74)$$

Diagonalization of $\boldsymbol{\Omega}^{\text{virt}}$ gives

$$\boldsymbol{\Omega}^{\text{virt}} \mathbf{u}_i = \lambda_i \mathbf{u}_i. \quad (75)$$

By introducing the transformation matrix

$$\mathbf{U} = (\mathbf{u}_1 \mathbf{u}_2 \dots \mathbf{u}_{m^{\text{virt}}}) \quad (76)$$

and the orbitals

$$\boldsymbol{\phi}^{\text{virt}} = \mathbf{U}^T \boldsymbol{\psi}^{\text{virt}}, \quad (77)$$

we obtain

$$\begin{aligned} P &= (\boldsymbol{\psi}^{\text{virt}})^T \boldsymbol{\Omega}^{\text{virt}} \boldsymbol{\psi}^{\text{virt}} = (\mathbf{U} \boldsymbol{\phi}^{\text{virt}})^T \boldsymbol{\Omega}^{\text{virt}} (\mathbf{U} \boldsymbol{\phi}^{\text{virt}}) \\ &= (\boldsymbol{\phi}^{\text{virt}})^T \mathbf{U}^T \boldsymbol{\Omega}^{\text{virt}} \mathbf{U} \boldsymbol{\phi}^{\text{virt}} \\ &= \sum_i \phi_i^{\text{virt}} \lambda_i \phi_i^{\text{virt}}. \end{aligned} \quad (78)$$

The set $\{\phi_i^{\text{virt}}; 1 \leq i \leq m^{\text{virt}}\}$ is the NOs of the type considered (either α or β type). The truncation of the virtual space to a lower dimension, say, $m_{\text{trunc}}^{\text{virt}} < m^{\text{virt}}$, is simply obtained by selecting the $m_{\text{trunc}}^{\text{virt}}$ orbitals with the largest occupation numbers $\{\lambda_i\}$.

For a three- or four-electron FCI calculation, we select the NOs from the appropriate two-electron FCI calculations, and construct a linear independent set of α and/or β type virtual orbitals.

Within the PATMOS framework, there is a considerable manipulation of two-electron integrals: different virtual space for each atom and for each pair of atoms, different truncated virtual spaces for each calculation of a poly-electron FCI calculation. The computation of the relevant two-electron integrals can be simplified if one utilizes a Cholesky decomposition of the two-electron integral matrix. The advantage has been discussed in detail in a work by Røeggen and Johansen.⁴⁷

H. Localization measures

For describing the localized character of orbitals, we use charge centroids and charge ellipsoids.⁴⁸ For any spatial

orbital ϕ , we define the charge centroid by the following relation:

$$\mathbf{r}^C = \langle \phi | \mathbf{r} \phi \rangle. \quad (79)$$

The extension or spread of an orbital can be described by the second-order variance matrix

$$M_{rs} = \langle \phi | (x_r - x_r^C)(x_s - x_s^C) \phi \rangle, \quad r, s \in \{1, 2, 3\}. \quad (80)$$

In Eq. (80), x_r^C is the r th component of the charge centroid vector \mathbf{r}^C . Diagonalization of the variance matrix yields the charge ellipsoid. The eigenvalues $\{a_1, a_2, a_3\}$ of the matrix (M_{rs}) correspond to the squares of the half-axes of the ellipsoid. The standard deviations in three orthogonal directions, i.e., the directions of the half-axes, are therefore given by

$$\Delta l_i = a_i^{1/2}, \quad i \in \{1, 2, 3\}. \quad (81)$$

The quantities $\{\Delta l_i\}$ can then be used as a measure of the extension of the orbital with respect to the charge centroid position. We may also use the volume of the ellipsoid as a single number for the extension

$$V = \frac{4}{3} \pi \Delta l_1 \Delta l_2 \Delta l_3. \quad (82)$$

The half-axes also define a lower bound for the kinetic energy associated with a given orbital. From the Heisenberg uncertainty principle, we derive

$$E_{\text{kin}} = \langle \phi | -\frac{1}{2} \nabla^2 \phi \rangle \geq \frac{1}{8} \left[\frac{1}{(\Delta l_1)^2} + \frac{1}{(\Delta l_2)^2} + \frac{1}{(\Delta l_3)^2} \right]. \quad (83)$$

This inequality expresses neatly that the kinetic energy increases when an orbital becomes more compact.

III. UHF FOR EXTENDED SYSTEMS

An extended system is in this work modelled by a finite cluster. However, a ‘‘brute force’’ cutting of the extended system yields a cluster with atoms on the surface, which have surroundings that are different from the surroundings of the atoms in the interior of the cluster. To avoid this problem, we construct a specific UHF state for the atoms on the surface, i.e., the boundary atoms. The orbitals of the boundary atoms are frozen and closely related to orbitals of an atom in the interior of the cluster.

Our approach starts with a reference fragment which corresponds to the unit cell of the extended systems. This reference fragment is surrounded by a certain number of partner fragments. The reference fragment and its partner fragments are enclosed by a set of boundary fragments. This particular embedding is illustrated in Fig. 1. The numbers of embedded atoms and boundary atoms are denoted, respectively, $N_{\text{atoms}}^{\text{emb}}$ and $N_{\text{atoms}}^{\text{bound}}$. A localized spin orbital of an embedded atom A is denoted $\psi_i^{A;\text{emb}}$, and the corresponding spatial $\phi_i^{A;\text{emb}}$. Similarly, we have for a boundary atom A the orbitals $\psi_i^{A;\text{bound}}$

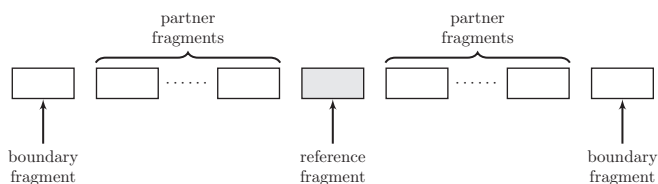


FIG. 1. One-dimensional model system for an extended system of identical fragments. The model system comprises a reference fragment, primary fragments, and boundary fragments.

and $\phi_i^{A;\text{bound}}$. The spatial orbital basis set for the cluster

$$\begin{aligned} & \{\chi_{\mu}; 1 \leq \mu \leq m_{\text{model}}\} \\ &= \bigcup_{A=1}^{N_{\text{atoms}}^{\text{emb}}} \{\chi_{\mu}^{A;\text{emb}}; 1 \leq \mu \leq m^{A;\text{emb}}\} \\ & \cup \bigcup_{A=1}^{N_{\text{atoms}}^{\text{bound}}} \{\chi_{\mu}^{A;\text{bound}}; 1 \leq \mu \leq m^{A;\text{bound}}\}. \quad (84) \end{aligned}$$

The spatial part of a spin orbital associated with a boundary atom is restricted to a one-center set

$$\phi_i^{A;\text{bound}} = \sum_{\mu=1}^{m^{A;\text{bound}}} \chi_{\mu}^{A;\text{bound}} \mathbf{U}_{\mu i}^{A;\text{bound}}. \quad (85)$$

As for the orbital of an embedded atom, the complete orbital space is used

$$\phi_i^{A;\text{emb}} = \sum_{\mu=1}^{m_{\text{model}}} \chi_{\mu} \mathbf{U}_{\mu i}^{A;\text{emb}}. \quad (86)$$

The determination of the spin orbitals of the model cluster is a double iterative procedure. Each cycle of the procedure consists of the following steps:

- Step 1. For fixed boundary orbitals, the orbitals of the embedded atoms are determined. These orbitals are orthogonal to the boundary orbitals.
- Step 2. Edmiston-Ruedenberg localization of the spin orbitals and definition of the perturbed atoms.
- Step 3. Construction of one-center approximations for the orbitals of the atoms in the reference fragment. Let $P^{A;\text{emb}}$ denote the orbital projection operator associated with the one-center set $\{\chi_{\mu}^{A;\text{emb}}; 1 \leq \mu \leq m^{A;\text{emb}}\}$. The one-center approximation is defined by the equation

$$\phi_i^{A;\text{one-center}} = P^{A;\text{emb}} \phi_i^{A;\text{emb}}. \quad (87)$$

The projected orbitals associated with an atom is then orthonormalized by a symmetric orthonormalization procedure, i.e., Löwdin orthonormalization.

- Step 4. New boundary orbitals are obtained by translations of the one-center orbitals of the atoms of reference fragment, to the boundary fragments.

The steps 1-4 are repeated until the UHF energy of the embedded atoms has converged.

TABLE I. Sum of three- and four-electron FCI correlation corrections for N_2 in the ground state as a function of the number of pair natural orbitals. Internuclear distance: $R = 2.1$ bohrs. GTF basis: (17s9p4d3f1g).

$n_{\text{pair}}^{\text{NO}}$	$\epsilon_{\text{poly-elec}}^{\text{intra-atom}}$ (E_h)	$\epsilon_{\text{poly-elec}}^{\text{interatom}}$ (E_h)
10	0.005868	0.081709
15	0.005730	0.080602
20	0.005665	0.080229
	0.005608 ^a	0.080044 ^a

^aGeometrical extrapolation.

IV. TEST CALCULATIONS

In order to evaluate the accuracy of the PATMOS model, we present test calculations on N_2 and Li_2 . Some preliminary test calculations on parallel arrays of hydrogen atoms demonstrate the convergence property of the UHF procedure for large systems.

Our integral code requires family type basis sets. The way we construct these family basis sets is explained in the Appendix. Only the spherical component subsets of the Cartesian functions are used. The parameters defining the basis sets, are given in Table XV.

In this work, we use an integral threshold of $10^{-8} E_h$ or less, for the Cholesky decomposition of the two-electron matrix.

A. Nitrogen molecule

The localization of the UHF orbitals by the Edmiston-Ruedenberg procedure yields three equivalent ‘‘banana shaped’’ bond pair geminals. We could of course use these geminals in the calculations. However, we discovered that this bonding picture changed for longer internuclear distances mixing the lone pair and bond pair orbitals. To obtain the conventional picture of bonding, for each set of localized atom type orbitals, we diagonalized the Fock operator with the localized intra-atomic orbitals as basis. Hence, we obtained an intra-atomic canonical picture. The bonding picture is then a σ -type bond pair geminal, two equivalent π -type bond pair geminals, and a lone pair and a core pair geminal for each atom.

As for the three basis sets adopted for the N_2 calculations, the smallest set, i.e., Gaussian-type function (GTF) (17s9p4d3f1g), is constructed according to the recipe described in the Appendix. However, when we tried to use the same procedure for the two largest sets, we ran into numerical difficulties due to near linear dependency. These two sets

TABLE III. Equilibrium distance and electronic dissociation energy for the ground state of N_2 calculated by the PATMOS model with different basis sets.

GTF basis	R_e (\AA)	D_e (eV)
17s9p4d3f1g	1.1033	9.576
17s9p5d4f3g2h1i	1.1014	9.746
17s9p7d6f5g4h3i2j1k	1.1010	9.755
Expt. ^a	1.0977	9.910

^aReference 56.

are therefore derived from the smallest one by even-tempered expansions based on the same β -parameter as for the smallest set. Hence, our basis sets are not fully optimized.

The accuracy of the PATMOS model depends on the dimension of the truncated virtual subspaces used for calculating three- and four-electron corrections. In Table I, we present the sum of three- and four-electron corrections for N_2 in the ground state, as a function of the number, $n_{\text{pair}}^{\text{NO}}$, of pair natural orbitals. For a three-electron correction, the dimension of the virtual space is $3n_{\text{pair}}^{\text{NO}}$, and for a four-electron correction $5n_{\text{pair}}^{\text{NO}}$. These numbers might be slightly reduced due to a linear dependency. We notice that both the intra- and interatomic corrections decrease in magnitude with increasing dimension of the truncated virtual space. A geometrical extrapolation suggests that with $n_{\text{pair}}^{\text{NO}} = 20$, the poly-electron corrections are in error with approximately 1% and 0.3% for, respectively, the intra- and interatomic corrections. An error of $0.2 mE_h$, interatomic term, is well below the conventionally requirement of chemical accuracy of $1 \text{ kcal/mol} = 1.59 mE_h$.

In Table II, we display total energies and correlation corrections for different basis sets. The number of pair natural orbitals for the poly-electron corrections is $n_{\text{pair}}^{\text{NO}} = 20$. As expected, the two-electron corrections have the largest variation with respect to the quality of the basis set. This result is in accordance with the fact that the ‘‘cusp’’ – feature in an exact wave function is essentially a two-electron correlation effect. As for the poly-electron corrections, we notice a considerably smaller variation with respect to the basis sets. Our results suggest that we can obtain very accurate results if we perform calculations of two-electron calculations with huge basis sets while include poly-electron corrections from calculations with more modest basis sets. For the two-electron corrections, one could also consider R12 methods^{49–51} to approach the basis set limit.

In Table III, we display the equilibrium distance and the electronic dissociation energy for the ground state of N_2 , calculated with the three chosen basis sets. Since our model is

TABLE II. Total energies and correlation corrections for N_2 in the ground state for different basis sets. Internuclear distance: $R = 2.1$ bohrs. Number of pair natural orbitals: $n_{\text{pair}}^{\text{NO}} = 20$.

GTF basis	E^{UHF} (E_h)	$\sum \epsilon_{A:ij}^{\text{corr}}$ (E_h)	$\epsilon_{\text{poly-elec}}^{\text{intra-atom}}$ (E_h)	$\sum \epsilon_{ij}^{\text{interatom}}$ (E_h)	$\epsilon_{\text{poly-elec}}^{\text{interatom}}$ (E_h)	E^{PATMOS} (E_h)
17s9p4d3f1g	−108.989398	−0.156531	0.005665	−0.316353	0.080229	−109.527254
17s9p5d4f3g2h1i	−108.989561	−0.160668	0.005652	−0.321257	0.080210	−109.540641
17s9p7d6f5g4h3i2j1k	−108.989561	−0.164306	0.005645	−0.323057	0.080187	−109.548857

TABLE IV. Calculated equilibrium distance and electronic dissociation energy for the ground state of N₂.

Model/basis	R_e (Å)	D_e (eV)
CCSD(T)/aug-cc-pV6Z/Composite ^a	1.0979	9.865
CCSD(T)/aug-cc-pV7Z/Composite ^a	1.0977	9.872
PATMOS/17s9p7d6f5g4h3i2j1k	1.1010	9.755
Expt. ^b	1.0977	9.910

^aReference 52.^bReference 56.

both size consistent and size extensive, the dissociation energy can be calculated by either performing separated calculations on the isolated atoms or performing a molecular calculation at a very large internuclear distance. In our case, the latter approach is used. We notice increased accuracy with increasing quality of the basis set. However, even with our largest basis set there is a discrepancy between the calculated values and the experimental ones, in particular for the dissociation energy. In Table IV, we include theoretical results from the very extensive work by Feller and co-workers.⁵² The coupled-cluster singles, doubles (and triples) (CCSD(T)) calculation with the largest basis set is in very good agreement with the experimental results. The quoted work from Feller and co-workers⁵² also includes corrections from CCSDTQ/cc-pVTZ and FCI/cc-pVTZ calculations. But when we compare the PATMOS model with the coupled-cluster composite model, one must have in mind that the quoted results of Table IV are not corrected for the BSSE. As for the PATMOS model, there is no BSSE at the correlation level. There is a small BSSE at the UHF level, but this error is negligible with large basis sets. Even though our largest basis set is not fully optimized, it can be considered to be of a similar quality as the basis of the CCSD(T) calculation. Hence, we can estimate the magnitude of the BSSE in the CCSD(T) calculation by quantities obtained in the PATMOS calculation, i.e., the intra-atomic FCI two-electron corrections calculated in this work. In Table V, we consider intra-atomic two-electron pairs related to the triple bond. They are the electron pairs mostly affected by the BSSE. By construction, there is no BSSE in the terms $\{\epsilon_{A:ij}^{\text{corr}}\}$. On the other hand, the terms $\{\tilde{\epsilon}_{A:ij}^{\text{corr}}\}$ which are required for four-electron corrections, are calculated with a molecular basis derived from natural orbitals. Hence, they are prone to BSSE. We notice that the BSSE is not negligible. For the triple bond pairs, it amounts to 0.12 eV for the inter-

TABLE VI. Calculated equilibrium distance and electronic dissociation energy for the ground state of Li₂.

Model/basis	R_e (Å)	D_e (eV)
MCSCFCI/9s5p1d ^a	2.676	1.00
EXGEM1/9s5p1d ^b	2.690	1.024
OVC/STO(13σ8π3δ) ^c	2.692	1.03
CCSD(T)/cc-pwCVQZ ^d	2.674	1.051
PATMOS/19s8p7d5f2g1h	2.674	1.052
Expt. ^e	2.673	1.048±0.006

^aReference 57.^bReference 58.^cReference 59.^dReference 60.^eReference 61.

nuclear distance $R = 2.1$ bohrs. This is of course an estimate of the BSSE in the CCSD(T) calculation. But we do believe the estimate gives the correct order of magnitude of the BSSE. If we subtract this estimate from $D_e = 9.872$ eV, we arrive at a modified value equal to $D_e = 9.752$ eV, a value very close to the PATMOS result of $D_e = 9.755$ eV.

The BSSE also affects the calculated geometrical structure. Without correcting for the BSSE, calculations yield equilibrium distances shorter than the BSSE corrected distances. We estimate that correction for BSSE might increase the equilibrium distance by roughly 0.003 Å. By adding this estimate to $R_e = 1.0977$ Å, we obtain the value 1.1007 Å. The PATMOS value and the experimental value are, respectively, 1.1010 Å and 1.0977 Å.

To conclude this subsection, the calculations on N₂ suggest that the PATMOS model and the coupled-cluster composite model, when correcting for BSSE, yield results which are very similar.

B. Lithium molecule

When comparing different computational models, the lithium molecule is an interesting test case since for this molecule there is no intra-atomic valence correlation energy. Hence, the intrinsic BSSE in standard models is negligible. In Table VI, we display the internuclear distance and electronic dissociation energy for the ground state of Li₂ obtained by different models. We notice in particular the almost identical results for CCSD(T) and the PATMOS model, and the very good agreement with the experimental values. Hence, the Li₂

TABLE V. Intra-atomic two-electron corrections for N₂ in the ground state calculated with different basis sets: intra-atomic basis and truncated virtual space used for four-electron corrections.^a

Internuclear distance (bohr)	Orbital pairs	$\epsilon_{A:ij}^{\text{corr}}$ ^b (E_h)	$\tilde{\epsilon}_{A:ij}^{\text{corr}}$ ^c (E_h)	$\Delta_{ij} = \tilde{\epsilon}_{A:ij}^{\text{corr}} - \epsilon_{A:ij}^{\text{corr}}$ (E_h)
2.1	$(2p\pi_x^A, 2p\sigma^A)$	-0.005672	-0.006303	-0.000631
	$(2p\pi_y^A, 2p\sigma^A)$	-0.005672	-0.006303	-0.000631
	$(2p\pi_x^A, 2p\pi_y^A)$	-0.020157	-0.021065	-0.000908

^aBasis: (17s9p7d6f5g4h3i2j1k).^bDimension of virtual space: 294.^cDimension of virtual space: 100.

TABLE VII. Total energies and correlation corrections for Li_2 in the ground state as a function of selected internuclear distances. GTF basis: $(19s8p7d5f2g1h)$. Number of pair natural orbitals: $n_{\text{pair}}^{\text{NO}} = 20$.

R (bohr)	E^{UHF} (E_h)	$\sum \epsilon_{A;ij}^{\text{corr}}$ (E_h)	$\epsilon_{\text{poly-elec}}^{\text{intra-atom}}$ (E_h)	$\sum \epsilon_{ij}^{\text{interatom}}$ (E_h)	$\epsilon_{\text{poly-elec}}^{\text{interatom}}$ (E_h)	E^{PATMOS} (E_h)
5.00	-14.871638	-0.044026	0.000060	-0.032420	-0.000836	-14.992827
5.05	-14.871914	-0.044037	0.000061	-0.032158	-0.000819	-14.992843
5.10	-14.872152	-0.044046	0.000061	-0.031901	-0.000807	-14.992830
50.0	-14.865489	-0.044420	0.000071	0.0	0.0	-14.954187

calculations give support to the conjecture presented in Subsection IV A concerning the relation between the PATMOS model and the coupled-cluster composite model.

In Table VII, we present total energies and correlation corrections for selected internuclear distances. The electronic dissociation energy is $D_e = 0.038656 E_h$. The contributions to D_e from the UHF model, two-electron FCI corrections, and poly-electron correction are, respectively, $0.006425 E_h$, $0.031392 E_h$, and $0.000839 E_h$. The two-electron corrections dominate. The contribution from the UHF term is approximately 17%, and the poly-electron correction is approximately 2% of the dissociation energy.

We have calculated the potential energy curve for the ground state of the lithium molecule. A detailed account of the work will be presented elsewhere.

C. Arrays of hydrogen atoms

Arrays of hydrogen atoms are used as test systems for the convergence properties with respect to size, of the advocated UHF model for extended systems. The basis used is an uncontracted $(10s2p1d)$ set of Gaussian type functions. A linear array is defined as $(\text{H}_2)_{2n+1}$, and where the neighbouring nuclei are separated by a distance $r = 2.63$ bohrs. The chosen distance is obtained by minimizing the UHF energy of the primary cluster for three parallel arrays, $(\text{H}_2)_6$, in xy -plane. Boundary fragments are added only at the end of the parallel arrays. Localization of the UHF orbitals yields orbitals with charge centroids close to the positions of the nuclei. In Table VIII, we consider charge centroid and half-axes of the α -type orbital of the central H_2 fragment of $(\text{H}_2)_{2n+1}$. We notice a nice convergence as the number of H_2 fragments in-

TABLE VIII. Charge centroid x_c , half axes, and volume of the α -type orbital of the central H_2 fragment of $(\text{H}_2)_{2n+1}$. The cluster $(\text{H}_2)_{2n+1}$ is a one-dimensional array along the x -axis with equal distance, $r = 2.63$ bohrs, between nearest neighbour atoms. Position of the nucleus associated with the orbital: $(0.0, 0.0, 0.0)$ bohr.

n	x_c (bohr)	Δl_x (bohr)	Δl_y (bohr)	Δl_z (bohr)	V (bohr) ³
1	-0.000493	1.3574	0.9625	0.9625	5.2680
2	-0.001753	1.3752	0.9636	0.9636	5.3496
3	-0.001077	1.3786	0.9640	0.9640	5.3661
4	-0.000592	1.3791	0.9641	0.9641	5.3695
5	-0.000327	1.3791	0.9642	0.9642	5.3703
6	-0.000187	1.3791	0.9642	0.9642	5.3705

creases. The bonding in such an array is characterized by atomic like orbitals which elongate along the direction of the array and contracts slightly in the two orthogonal directions to the array axis. The length of the half-axis in isolated hydrogen is 0.9716 bohr calculated with the adopted basis set (a large basis set yields 1.0000 bohr). By considering parallel arrays, Table IX, we get a qualitative picture of the bonding in an infinite two-dimensional lattice of equidistant hydrogen atoms. The charge centroid of a localized orbital coincides with the position of a hydrogen nucleus and there will be two equal half-axes in the lattice plane. These two half-axes are lengthened compared with a half-axis of the orbital of an isolated atom. There will be a contraction of the half-axis orthogonal to the lattice plane. No delocalization, i.e., metallic character is to be expected.

To our knowledge, the bonding character of hydrogen under high pressure, is an unresolved problem. In the future, when we have optimized our code, we shall use the PATMOS model to address this problem.

V. THE CHARACTER OF THE PERTURBED ATOMS

In this section, we will illustrate the character of perturbed atoms by considering the bonding in three molecules: hydrogen molecule, methane, and benzene.

A. Hydrogen molecule

The hydrogen molecule is undoubtedly the most thoroughly studied molecule in quantum chemistry. Even so, we believe that our approach can throw some new light on the question of bonding in the hydrogen molecule. This molecule is particularly simple in our approach. No localization is required. The α spin orbital is associated with one nucleus, and

TABLE IX. Charge centroid x_c , half axes, and volumes of the α -type orbital of the central H_2 fragment of $(\text{H}_2)_5$ parallel arrays. The central $(\text{H}_2)_5$ cluster is a one-dimensional array along the x -axis with equal distance, $r = 2.63$ bohrs, between nearest atoms. The distance between neighbouring parallel arrays: 2.63 bohrs. Position of the nucleus associated with the orbital: $(0.0, 0.0, 0.0)$ bohr.

	x_c (bohr)	Δl_x (bohr)	Δl_y (bohr)	Δl_z (bohr)	V (bohr) ³
One array	-0.001753	1.3752	0.9636	0.9636	5.3496
Three parallel arrays	-0.001134	1.2628	1.1334	0.9025	5.4113
Five parallel arrays	-0.000901	1.2251	1.1868	0.9024	5.4962

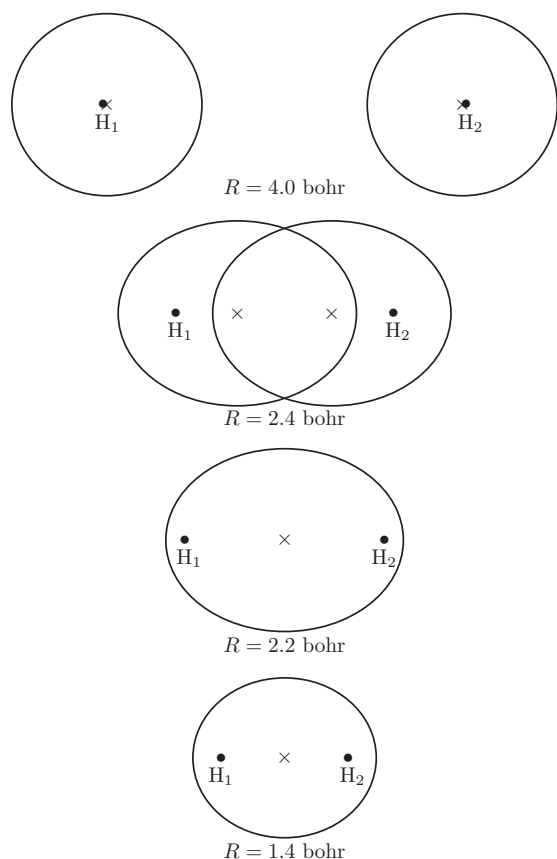


FIG. 2. Intersection between the xy -plane and the charge ellipsoid of the hydrogen molecule for four different internuclear distances. Half-axes and distances in scale.

the β spin orbital with the other. Hence, a perturbed atom comprises a spin orbital and a nucleus. This definition of perturbed atoms is valid for any internuclear distance.

The basis set adopted in this study is an uncontracted ($12s, 4p, 3d, 1f$) set of GTFs. In this case, the PATMOS energy is identical to the FCI energy.

The results of our calculations are displayed in Figures 2 and 3, and Tables X and XI. Figure 2 and Table X give a picture of bonding based on charge centroids and charge ellipsoids of orbitals. In Figure 2, we present charge centroids and

TABLE X. Charge centroid, distance between charge centroids, half-axes of charge ellipsoid, volume of charge ellipsoid, and the overlap matrix element between the spatial orbitals of the hydrogen molecule as a function of the internuclear distance R . Nuclear position: $(0, 0, 0)$ and $(R, 0, 0)$.

R (bohr)	x^{H_1} (bohr)	$ x^{\text{H}_1} - x^{\text{H}_2} $ (bohr)	$\Delta l_x^{\text{H}_1}$ (bohr)	$\Delta l_y^{\text{H}_1}, \Delta l_z^{\text{H}_1}$ (bohr)	V (bohr ³)	$ \langle \psi_1^{\text{H}_1} \psi_1^{\text{H}_2} \rangle $
0.5	0.2500	0.0000	0.7217	0.7032	1.4946	1.0000
1.4	0.7000	0.0000	1.0108	0.8811	3.2870	1.0000
2.2	1.1000	0.0000	1.3095	1.0076	5.5688	1.0000
2.3	1.0397	0.2206	1.3457	1.0204	5.8694	0.9964
2.4	0.6791	1.0417	1.3135	1.0194	5.7178	0.9211
2.5	0.5243	1.4513	1.2820	1.0179	5.5643	0.8498
3.0	0.2019	2.5961	1.1617	1.0111	4.9752	0.5714
4.0	0.0412	3.9175	1.0477	1.0036	4.4203	0.2628
10.0	0.0000	10.0000	1.0000	1.0000	4.1895	0.0022

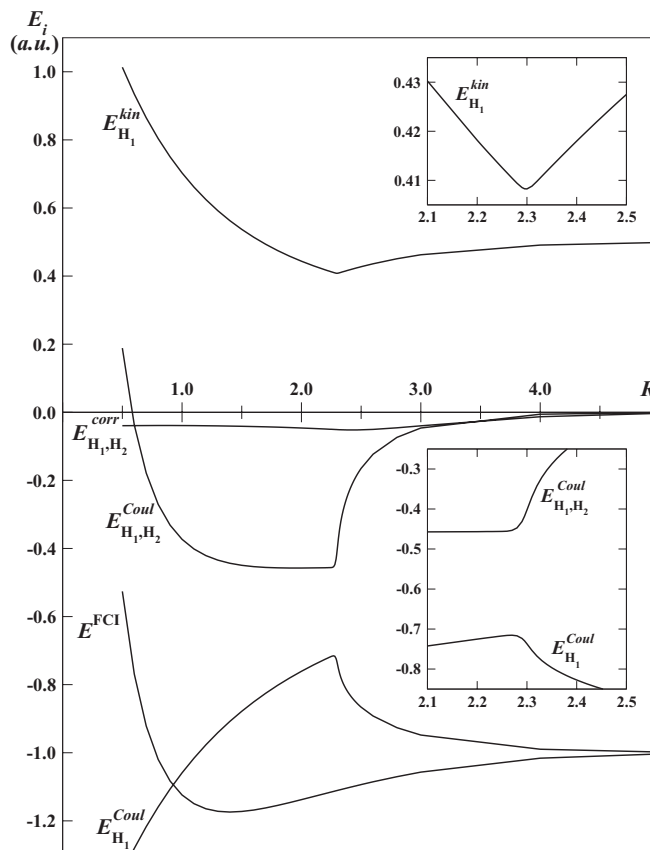


FIG. 3. Energy components of the PATMOS energy of the hydrogen molecule as a function of the internuclear distance.

charge ellipsoids for four different distances. For the largest internuclear distance, $R = 4.0$ bohrs, the charge ellipsoids are slightly distorted compared with the case of isolated atoms. There is a small elongation along the internuclear axis, and the volume of the ellipsoid has increased. If we look at the overlap matrix element of the two orbitals involved, Table X, we notice a value of 0.262796. Hence, the bonding process is well underway. There is a smooth shift and expansion of the ellipsoids until the internuclear distance is reduced to approximately 2.4 bohrs. Then there is a drastic change in bonding character with reduced internuclear distance. At the distance, $R = 2.2$ bohrs, the two spatial UHF orbitals overlap completely. A further reduction of the internuclear distance implies an overall contraction of the orbital density. At the distance, $R = 1.4$ bohrs, close to the calculated equilibrium distance of $R_{\text{eq}} = 1.407$ bohrs, the volume of the charge ellipsoid is 3.287 bohrs³. This can be compared with 5.7178 bohrs³ and 5.5688 bohrs³ for the volumes at the distances $R = 2.4$ bohrs and $R = 2.2$ bohrs, respectively.

In Figure 3, we display the energy components of PATMOS energy as functions of the internuclear distance. We notice immediately that the Coulombic interaction between the perturbed atoms, is the driving force of the bonding. But there is more to come. As the orbitals are shifted into the internuclear region, with accompanying increase of volume of the charge ellipsoids, the kinetic energy is reduced. Furthermore, the intra-atomic Coulomb energy increases at first due to this orbital shift into the internuclear region. As noticed in

TABLE XI. Decomposition of the PATMOS energy of the hydrogen molecule as a function of the internuclear distance.

R (bohr)	$E_{H_1}^{\text{kin}}$ (E_h)	$E_{H_1}^{\text{Coul}}$ (E_h)	$E_{H_1,H_2}^{\text{Coul}}$ (E_h)	$E_{H_1,H_2}^{\text{corr}}$ (E_h)	E^{PATMOS} (E_h)
0.5	1.0132	-1.3508	0.1886	-0.0394	-0.5260
0.7	0.8650	-1.2170	-0.1785	-0.0390	-0.9215
0.9	0.7502	-1.1066	-0.3316	-0.0389	-1.0832
1.1	0.6614	-1.0159	-0.4015	-0.0392	-1.1497
1.3	0.5920	-0.9408	-0.4343	-0.0400	-1.1720
1.4	0.5630	-0.9081	-0.4434	-0.0405	-1.1741
1.5	0.5372	-0.8782	-0.4494	-0.0411	-1.1725
1.7	0.4936	-0.8254	-0.4556	-0.0428	-1.1621
1.9	0.4585	-0.7806	-0.4574	-0.0449	-1.1465
2.1	0.4302	-0.7423	-0.4571	-0.0476	-1.1288
2.2	0.4182	-0.7252	-0.4566	-0.0492	-1.1198
2.3	0.4083	-0.7370	-0.4025	-0.0509	-1.1108
2.35	0.4129	-0.7966	-0.2876	-0.0516	-1.1064
2.40	0.4180	-0.8273	-0.2316	-0.0519	-1.1021
2.45	0.4229	-0.8491	-0.1935	-0.0519	-1.0978
2.50	0.4275	-0.8662	-0.1647	-0.0516	-1.0936
2.6	0.4361	-0.8921	-0.1232	-0.0503	-1.0854
2.8	0.4509	-0.9263	-0.0737	-0.0457	-1.0703
3.0	0.4625	-0.9480	-0.0462	-0.0397	-1.0569
4.0	0.4911	-0.9897	-0.0057	-0.0132	-1.0160
5.0	0.4981	-0.9979	-0.0008	-0.0031	-1.0036
6.0	0.499600	-0.999577	-0.000122	-0.000672	-1.000749
7.0	0.499917	-0.999913	-0.000019	-0.000150	-1.000160
8.0	0.499980	-0.999978	-0.000003	-0.000032	-1.000031
9.0	0.499992	-0.999990	0.000000	-0.000010	-1.000006
10.0	0.499994	-0.999992	0.000000	-0.000005	-1.000001

Figure 2, there is an overall contraction of the charge density when the internuclear distance is reduced from 2.2 bohrs to 1.4 bohrs. With this contraction there is a small change in interatomic Coulomb energy, an increase in kinetic energy, and an increase in magnitude of the intra-atomic Coulomb energy due to the orbitals coming closer to the nuclei. A further reduction of the internuclear distance implies a large increase in the nuclear-nuclear interaction and an increase in kinetic energy. These two terms are the main origin of the repulsive part of the interatomic potential.

In Table XI, we have displayed a partitioning of the total electronic energy into intra-atomic components $E_{H_1}^{\text{kin}}$ (kinetic energy component) and $E_{H_1}^{\text{Coul}}$ (intra-atomic Coulombic), and the interatomic components $E_{H_1,H_2}^{\text{Coul}}$ and $E_{H_1,H_2}^{\text{corr}}$, for internuclear distances up to $R = 10.0$ bohrs. We notice that for distances larger than $R = 7.0$ bohrs, the correlation energy is the dominant attractive term. It is also important in the intermediate region even if the interatomic Coulomb energy is in magnitude approximately 10 times larger. Due to different signs of attractive and repulsive terms, the correlation energy contributes approximately 24% to the binding energy.

Pertaining to Table X, we would also like to emphasize the overlap property of the spatial parts of the two spin orbitals. For the inter-nuclear distance $R \approx 2.2$ bohrs and smaller distances, the overlap matrix element is equal to one. In this region, restricted Hartree-Fock (RHF) and UHF give equal results. Let us ask the question: at what distance is the

chemical bond between the two atoms being fully formed? Or equivalently: at what distance starts the two-electron bond to break? If we look at the potential energy curve, we can hardly answer. However, the UHF suggests an answer: the distance for which the RHF method and UHF model starts to give different results, is the bond breaking point.

Our interpretation of the bonding in H_2 differs from the common notion of bonding derived from valence bond theory. In that case, bonding in H_2 is essentially due to an ‘‘exchange’’ interaction between different atomic like orbitals. However, our approach has some resemblance with the more recent analysis of Ruedenberg and Schmidt.⁵³ They analyse bonding in terms of two components: the kinetic energy and the total Coulombic energy. The difference between their work and the present one is essentially that we partition the Coulombic energy into intra- and interatomic terms.

It is important to have in mind that different computational models with different concepts lead to different interpretations. The interpretation based on the PATMOS model accords closely with the intuitive idea that a chemical bond is due to the attraction between the electrons of one atom with the nucleus of a second atom and vice versa.

To conclude, an analysis based on perturbed atoms yields a simple, intuitive, consistent, and accurate description of the bonding in the hydrogen molecule.

B. Methane

The basis for the calculation on methane is uncontracted ($18s, 10p, 3d, 1f/10s, 3p, 1d$) set of Gaussian type functions. The C–H bond length is 2.05 bohrs, and hydrogen nuclei are in tetrahedral arrangement around the carbon nucleus. In the three-electron FCI calculations of the intra-atomic correlation energy of the carbon atom, 20 NOs from each pair are adopted.

The calculation yields identical sets of spatial orbitals for the α and β type spin orbitals. Localization of the spatial orbitals gives four equivalent valence orbitals. Each orbital is located in the region between the carbon nucleus and a hydrogen nucleus. The longest half-axis of the corresponding charge ellipsoid is directed along the C–H axis. The values of the half-axes are, respectively, 1.0605 bohrs, 0.8238 bohr, and 0.8238 bohr. The spatial position of the charge centroid is on the C–H axis, and 0.6608 bohr away from the hydrogen nucleus.

In Table XII, we present intra- and interatomic energy components of the PATMOS energy of methane. We notice that for the C–H interaction energy, the Coulomb energy

TABLE XII. Intra- and interatomic energy components of the PATMOS energy of methane. Bond length is 2.05 bohrs.

Fragment(s)	Kinetic (E_h)	Coulombic (E_h)	Exchange (E_h)	Correlation (E_h)	Total (E_h)
C	36.7004	-73.3928	-0.1353	-0.0672	-36.8948
H	0.8740	-0.9494			-0.0754
(C,H)		-0.7550	-0.0084	-0.0492	-0.8126
(H_1,H_2)		0.0113	-0.0170	-0.0026	-0.0083

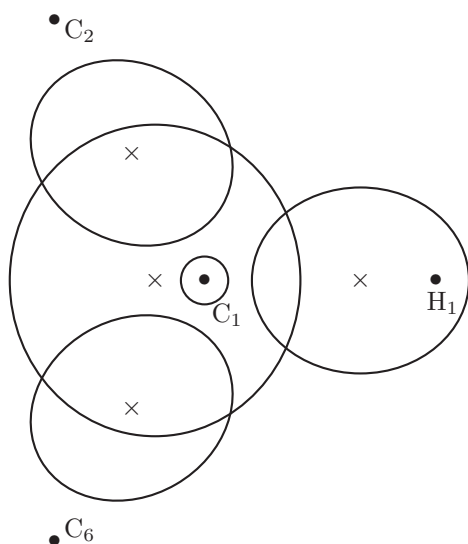


FIG. 4. Intersection between the xy -plane and selected charge ellipsoids of a fragment of benzene. Bond lengths: $l(\text{C}-\text{C}) = 2.65$ bohrs, $l(\text{C}-\text{H}) = 2.04$ bohrs.

dominates. It is in magnitude, respectively, 90 and 15 times larger than the exchange and correlation terms. As for the interaction energy between two hydrogen atoms, the Coulomb, exchange, and correlation terms are of the same magnitude. The interaction energy between the carbon atom and a hydrogen atom is two orders of magnitude larger than the interaction energy between two hydrogen atoms. Hence, our energy component analysis yields strong support for the familiar bond structure in methane.

C. Benzene

Conjugated or aromatic molecules are usually considered to be systems where an effective localization of the orbitals is impossible. The benzene molecule is an important member of this class of molecules. It will undoubtedly represent a challenge to define a perturbed atom in this molecule.

The basis set adopted for the study is the same as the one used in the methane calculation. The geometry of the molecule is the experimental one, i.e., 2.65 bohrs and 2.04 bohrs for the C–C and C–H bond lengths, respectively. The correlation energy is approximated by two-electron FCI calculations.

In Figure 4, we display charge centroids and charge ellipsoids for a fragment of the molecule. We notice the conven-

TABLE XIII. Half axes and volumes of charge ellipsoids of orbitals of benzene. Bond lengths: $l(\text{C}-\text{C}) = 2.65$ bohrs, $l(\text{C}-\text{H}) = 2.04$ bohrs.

Orbital	Δl_1 (bohr)	Δl_2 (bohr)	Δl_3 (bohr)	V (bohr ³)
1s	0.2109	0.2086	0.1910	0.0352
$\sigma_{\text{C}_1-\text{C}_2}$	0.9245	0.7787	0.7743	2.3349
$\sigma_{\text{C}_1-\text{H}_1}$	0.9572	0.8201	0.8137	2.6753
π_{C}	1.5611	1.3739	1.2817	11.5155

TABLE XIV. Intra- and interatomic energy components of the PATMOS energy of benzene. Bond lengths: $l(\text{C}-\text{C}) = 2.65$ bohrs, $l(\text{C}-\text{H}) = 2.04$ bohrs. Numbering of atoms given in Fig. 4.

Fragment(s)	Kinetic (E_h)	Coulombic (E_h)	Exchange (E_h)	Correlation (E_h)	Total (E_h)
C	37.5393	-73.5901	-0.1978	-0.0691	-36.3177
H	0.8868	-0.9637			-0.0769
(C ₁ ,C ₂)		-1.2502	-0.0208	-0.0901	-1.3611
(C ₁ ,C ₃)		-0.0154	-0.0364	-0.0066	-0.0585
(C ₁ ,C ₄)		-0.0361	-0.0006	-0.0051	-0.0418
(C ₁ ,H ₁)		-0.7560	-0.0084	-0.0518	-0.8162
(C ₁ ,H ₂)		-0.0015	-0.0193	-0.0032	-0.0240
(C ₁ ,H ₃)		-0.0002	0.0000	-0.0003	-0.0006
(C ₁ ,H ₄)		0.0010	-0.0010	-0.0001	-0.0001
(H ₁ ,H ₂)		0.0048	0.0	-0.0002	0.0046
(H ₁ ,H ₃)		0.0018	-0.0001	0.0000	0.0016
(H ₁ ,H ₄)		0.0014	0.0	0.0000	0.0014

tional picture of three σ -type orbitals and a localized π -type orbital. The spatial part of an α -type spin orbital can in principle have a non-zero overlap matrix element with the spatial part of any β -type orbital. As for the σ -type geminal, the overlap matrix element for the orbitals of the $\sigma_{\text{C}_1-\text{C}_2}$ geminal and $\sigma_{\text{C}_1-\text{H}_1}$ geminal are, respectively, 0.9993 and 0.9996. Overlap matrix elements between orbitals belonging to different σ -type geminals are typically of the order 0.001 or smaller. Regarding the π orbitals, the π -orbitals have alternate spin along the ring structure, the overlap matrix elements are 0.6216 and 0.2674, respectively, for the nearest neighbour carbon atoms and the third nearest neighbour atoms. As is evident from Figure 4 and Table XIII, the volume of a π -ellipsoid is considerably larger, i.e., 4.3 times larger than the volume of $\sigma_{\text{C}-\text{H}}$ -ellipsoid.

In Table XIV, we present intra- and interatomic energy components of the PATMOS energy of benzene. The numbering of atoms is a sequential numbering along the ring of the molecule in such a way that there is σ bond between carbon atom C_p and hydrogen atom H_p . In the table, we display only symmetry independent terms. We observe that a carbon atom is strongly bounded to its two nearest neighbour carbon atoms, and also strongly bonded to its partner hydrogen atoms. As in the methane case, the dominant component of the interaction energy is the Coulombic term. A carbon atom is weakly bounded to its second and third nearest neighbour atoms. There are weak interactions between a carbon atom and the hydrogen atoms which are not a nearest neighbour atom. As for interactions between the hydrogen atoms, they are all repulsive and dominated by the Coulombic term.

To conclude, it is possible to define perturbed atoms in a system like benzene and obtain a meaningful energy component analysis.

VI. CONCLUDING REMARKS

A well-known disadvantage of the UHF wave function, when UHF is different from restricted Hartree-Fock, is spin contamination. However, this contamination is reduced or perhaps eliminated by the partial FCI corrections in the PATMOS model. In our calculations, we monitor this

TABLE XV. Lowest exponents of the GTF-family basis sets used in this work.

Atom/basis	Reference basis ⁵⁵	<i>s</i>	<i>p</i>	<i>d</i>	<i>f</i>	<i>g</i>	<i>h</i>	<i>i</i>	<i>j</i>	<i>k</i>
H/(10 <i>s</i> ,2 <i>p</i> ,1 <i>d</i>)	cc-pVTZ	0.144717	0.391109	1.057000						
H/(10 <i>s</i> ,3 <i>p</i> ,1 <i>d</i>)	cc-pVTZ	0.067640	0.422800	1.057000						
H/(12 <i>s</i> ,5 <i>p</i> ,3 <i>d</i>)	cc-pVTZ	0.010823	0.067640	0.169100						
H/(12 <i>s</i> ,4 <i>p</i> ,3 <i>d</i> ,1 <i>f</i>)	cc-pVQZ	0.035763	0.223520	0.558800	1.397000					
C/(18 <i>s</i> ,9 <i>p</i> ,3 <i>d</i> ,1 <i>f</i>)	cc-pVTZ	0.044364	0.044364	0.295076	0.761000					
N/(17 <i>s</i> ,9 <i>p</i> ,4 <i>d</i> ,3 <i>f</i> ,1 <i>g</i>)	cc-pVQZ	0.032554	0.083764	0.215533	0.554586	1.427				
N/(17 <i>s</i> ,9 <i>p</i> ,5 <i>d</i> ,4 <i>f</i> ,3 <i>g</i> ,2 <i>h</i> ,1 <i>i</i>)	cc-pVQZ	0.032554	0.083764	0.215533	0.215533	0.554586	1.427	3.671799		
N/(17 <i>s</i> ,9 <i>p</i> ,7 <i>d</i> ,6 <i>f</i> ,5 <i>g</i> ,4 <i>h</i> ,3 <i>i</i> ,2 <i>j</i> ,1 <i>k</i>)	cc-pVQZ	0.032554	0.083764	0.215533	0.215533	0.554586	1.427	3.671799	9.447870	24.310220
Li/(19 <i>s</i> ,8 <i>p</i> ,7 <i>d</i> ,5 <i>f</i> ,2 <i>g</i> ,1 <i>h</i>)	cc-pV5Z	0.003801	0.054000	0.054000	0.131000	0.131000	0.320000			

contamination in the FCI two-electron wave functions for the (α , β) clusters. The final wave function is expressed in terms of a singlet ($S = 0$) and a triplet ($S = 1$) component. Consider the H₂ molecule which has a singlet ground state. For internuclear distances greater than $R \approx 2.2$ bohrs, there is a spin contamination in the UHF wave function. At infinite internuclear distance, there is a complete mixture of the singlet and triplet states in UHF wave function. On the other hand, this is not the case for the FCI wave function. Fully converged, the FCI wave function is a singlet state for any internuclear distance. But in approaching the infinite distance, great care must be taken to ensure that the wave function is converged. Due to the very small energy difference between the singlet and triplet states, the energy converges much faster than the wave function.

In the Introduction, we have suggested that the PATMOS model has both conceptual and computation advantages. The decomposition of the total electronic energy into intra- and interatomic components opens up for a new type of analysis of chemical bonding. This interpretative feature has been demonstrated in this work by a study on different groups of molecules. The computational advantage of the model is partly due to a well-defined procedure for reducing the virtual space in electron correlation calculations, and the use of symmetry in reducing the number of calculated intra- and interatomic correlation terms. The structure of the model, i.e., independent calculations of intra- and interatomic terms, makes it ideally suited for parallel processor computers. The model should be particularly useful for large systems since it allows for focusing on the relevant part of the correlation energy.

One of the most difficult problem in electronic structure theory, is the basis set problem: how to approach the basis set limit. The usual procedure is one or another type of extrapolation technique. However, mathematically, extrapolation is subjected to uncertainties. An alternative procedure can be based on R12 methods. As argued in Sec. IV A, the PATMOS model is ideally suited for such an approach. In the future, we shall contemplate to include R12 procedure in the PATMOS model.

ACKNOWLEDGMENTS

This work was supported by the Norwegian HPC project NOTUR that granted access to the supercomputer facilities in Norway.

APPENDIX: A CONSTRUCTION OF FAMILY TYPE BASIS SET

Our computer code is based on family type basis sets. The arguments for using this type of basis set are given by Røeggen and Johansen⁴⁷ in an article on the Cholesky decomposition of two-electron integral matrix. A family type basis is a basis of GTFs constructed as a union of families

$$\{\chi_{\mu}; 1 \leq \mu \leq m\} = \bigcup_{\underline{A}, a, L} \mathcal{F}_L(a, \underline{A}), \quad (\text{A1})$$

where the L -family, $\mathcal{F}_L(a, \underline{A})$, comprises a set of GTFs located on the same center \underline{A} , all functions have the same exponent a , and with angular momenta ranging from 0 to L . The exponent are chosen from an even-tempered sequence

$$\eta_k = \alpha\beta^{k-1}, \quad k = 1, 2, \dots \quad (\text{A2})$$

The parameter β is defined by a set of s -type GTFs, say by the optimized set of exponents in the work of Schmidt and Ruedenberg,⁵⁴ or as the ratio of the exponents of the two most diffuse s -type functions in a conventional set of basis functions. The exponent of the family with the highest angular momentum quantum number, i.e., L_{\max} , is chosen from a reference basis set. The upper and lower exponents of L -families, $L < L_{\max}$, are such that these two values bracket the exponents in the reference set. Since the β -parameter is deduced from a set of s -type exponents, the family basis will usually contain a larger set of functions than the reference set. Our reference sets are all from the EMSL database.⁵⁵

In Table XV, we give the lowest exponents of the GTF-family basis sets used in this work. The β -parameter can be derived from the exponents included in the table. Hence, the basis sets are completely specified.

¹R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics* (Addison-Wesley, Massachusetts, 1966).

²A. Einstein, B. Podolsky, and N. Rosen, *Phys. Rev.* **47**, 777 (1935).

³W. Moffitt, *Proc. R. Soc. London, Ser. A* **210**, 245 (1951).

⁴R. G. Parr, R. A. Donnelly, M. Levy, and W. E. Palke, *J. Chem. Phys.* **68**, 3801 (1978).

⁵W. E. Palke, *J. Chem. Phys.* **72**, 2511 (1980).

⁶M. P. Guse, *J. Chem. Phys.* **75**, 828 (1981).

⁷R. G. Parr, *Int. J. Quantum Chem.* **26**, 687 (1984).

⁸L. Li and R. G. Parr, *J. Chem. Phys.* **84**, 1704 (1986).

⁹J. Rychlewski and R. G. Parr, *J. Chem. Phys.* **84**, 1696 (1986).

¹⁰R. F. W. Bader, *Acc. Chem. Res.* **8**, 34 (1975).

¹¹R. F. W. Bader and H. J. T. Preston, *Int. J. Quantum Chem.* **3**, 327 (1969).

¹²R. F. W. Bader, S. G. Anderson, and A. J. Duke, *J. Am. Chem. Soc.* **101**, 1389 (1979).

- ¹³R. F. W. Bader, T. T. Nguyen-Dang, and Y. Tal, *J. Chem. Phys.* **70**, 4316 (1979).
- ¹⁴R. F. W. Bader, T. T. Nguyen-Dang, and Y. Tal, *Rep. Prog. Phys.* **44**, 893 (1981).
- ¹⁵R. F. W. Bader, *J. Chem. Phys.* **85**, 3133 (1986).
- ¹⁶R. F. W. Bader, A. Larouche, C. Gatti, M. T. Carroll, P. J. MacDougall, and K. B. Wiberg, *J. Chem. Phys.* **87**, 1142 (1987).
- ¹⁷R. F. W. Bader, *Pure Appl. Chem.* **60**, 145 (1988).
- ¹⁸J. Gerratt, *Adv. At. Mol. Phys.* **7**, 141 (1971).
- ¹⁹J. Gerratt and M. Raimondi, *Proc. R. Soc. London, Ser. A* **371**, 525 (1980).
- ²⁰M. Raimondi, and D. L. Cooper, in *Correlation and Localization*, Topics in Current Chemistry Vol. 203, edited by P. R. Surján (Springer, Berlin, 1999), pp. 105–120.
- ²¹R. K. Nesbet, *Adv. Chem. Phys.* **14**, 1 (1969).
- ²²I. Røeggen, *J. Chem. Phys.* **79**, 5520 (1983).
- ²³H. Stoll, *Phys. Rev. B* **46**, 6700 (1992).
- ²⁴H. Stoll, *J. Chem. Phys.* **97**, 8449 (1992).
- ²⁵A. Shukla, M. Dolg, P. Fulde, and H. Stoll, *Phys. Rev. B* **57**, 1471 (1998).
- ²⁶I. Røeggen, in *Correlation and Localization*, Topics in Current Chemistry Vol. 203, edited by P. R. Surján (Springer, Berlin, 1999), pp. 89–103.
- ²⁷S. F. Boys, *Rev. Mod. Phys.* **32**, 296 (1960).
- ²⁸J. Pipek and P. G. Mezey, *J. Chem. Phys.* **90**, 4916 (1989).
- ²⁹C. Edmiston and K. Ruedenberg, *Rev. Mod. Phys.* **35**, 457 (1963).
- ³⁰L. Bytautas and K. Ruedenberg, *J. Phys. Chem. A* **114**, 8601 (2010).
- ³¹O. Sinanoğlu, *Many-Electron Theory of Atoms, Molecules and Their Interactions*, Advances in Chemical Physics Vol. 6 (John Wiley & Sons, Inc., 1964), pp. 315–412.
- ³²R. K. Nesbet, *Electronic Correlation in Atoms and Molecules*, Advances in Chemical Physics Vol. 9 (John Wiley & Sons, Inc., 1965), pp. 321–363.
- ³³P. Pulay, *Chem. Phys. Lett.* **100**, 151 (1983).
- ³⁴S. Sæbø and P. Pulay, *Chem. Phys. Lett.* **113**, 13 (1985).
- ³⁵S. Sæbø and P. Pulay, *J. Chem. Phys.* **86**, 914 (1987).
- ³⁶S. Sæbø and P. Pulay, *J. Chem. Phys.* **88**, 1884 (1988).
- ³⁷S. Sæbø and P. Pulay, *Annu. Rev. Phys. Chem.* **44**, 213 (1993).
- ³⁸C. Hampel and H.-J. Werner, *J. Chem. Phys.* **104**, 6286 (1996).
- ³⁹M. Schütz and H.-J. Werner, *J. Chem. Phys.* **114**, 661 (2001).
- ⁴⁰M. Schütz and H.-J. Werner, *Chem. Phys. Lett.* **318**, 370 (2000).
- ⁴¹M. Schütz, *J. Chem. Phys.* **113**, 9986 (2000).
- ⁴²M. Schütz, *J. Chem. Phys.* **116**, 8772 (2002).
- ⁴³P. Y. Ayala and G. E. Scuseria, *J. Chem. Phys.* **110**, 3660 (1999).
- ⁴⁴M. S. Lee, P. E. Maslen, and M. Head-Gordon, *J. Chem. Phys.* **112**, 3592 (2000).
- ⁴⁵R. A. Friesner, R. B. Murphy, M. D. Beachy, M. N. Ringnald, W. T. Pollard, B. D. Dunietz, and Y. Cao, *J. Phys. Chem. A* **103**, 1913 (1999).
- ⁴⁶I. Røeggen, *J. Chem. Phys.* **124**, 184502 (2006).
- ⁴⁷I. Røeggen and T. Johansen, *J. Chem. Phys.* **128**, 194107 (2008).
- ⁴⁸M. A. Robb, W. J. Haines, and I. G. Csizmadia, *J. Am. Chem. Soc.* **95**, 42 (1973).
- ⁴⁹W. Klopper and W. Kutzelnigg, *Chem. Phys. Lett.* **134**, 17 (1987).
- ⁵⁰W. Klopper, R. Röhse, and W. Kutzelnigg, *Chem. Phys. Lett.* **178**, 455 (1991).
- ⁵¹J. Noga, W. Kutzelnigg, and W. Klopper, *Chem. Phys. Lett.* **199**, 497 (1992).
- ⁵²D. Feller, K. A. Peterson, and T. D. Crawford, *J. Chem. Phys.* **124**, 054107 (2006).
- ⁵³K. Ruedenberg and M. W. Schmidt, *J. Phys. Chem. A* **113**, 1954 (2009).
- ⁵⁴M. W. Schmidt and K. Ruedenberg, *J. Chem. Phys.* **71**, 3951 (1979).
- ⁵⁵See <https://bse.pnl.gov/bse/portal> for EMSL basis set exchange library.
- ⁵⁶K. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure. IV. Constants of Diatomic Molecules* (Van Nostrand, Princeton, 1979).
- ⁵⁷B. Jönsson, B. O. Roos, P. R. Taylor, and P. E. M. Siegbahn, *J. Chem. Phys.* **74**, 4566 (1981).
- ⁵⁸I. Røeggen, *Chem. Phys. Lett.* **92**, 398 (1982).
- ⁵⁹D. D. Konowalow and M. L. Olson, *J. Chem. Phys.* **71**, 450 (1979).
- ⁶⁰S. E. Wheeler, K. W. Sattelmeyer, P. v. R. Schleyer, and H. F. Schaefer III, *J. Chem. Phys.* **120**, 4683 (2004).
- ⁶¹R. Velasco, C. Ottinger, and R. N. Zare, *J. Chem. Phys.* **51**, 5522 (1969).