Theor Chem Acc (1998) 99:329–343
DOI 10.1007/s002149800m18

Theoretical
Chemistry Accounts

Regular article

One - determinantal pure state versus ensemble Kohn-Sham solutions in the case of strong electron correlation: CH_2 and C_2

P.R.T. Schipper, O.V. Gritsenko, E.J. Baerends

Scheikundig Laboratorium der Vrije Universiteit De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands

Received: 10 February 1998 / Accepted: 10 June 1998 /

Abstract. The possibility that the Kohn-Sham (KS) solution for a noninteracting auxiliary electron system is not the conventional one-determinantal pure state but a few-determinantal ensemble has been investigated. The KS solutions (the exchange-correlation potential v_{xc} and the orbitals) have not been approximated by local-density or density-gradient approximations but have been constructed from an accurate ab initio electron density. The lowest singlet states of the CH_2 and C_2 molecules have been selected for this investigation since for these cases the ground-state wave function Ψ is nondegenerate but has an essentially multideterminantal character (electron correlation is strong). For C_2 the dependence of the type of KS solution on the bond distance R(C–C) has been studied at the QZ level. For the shortest distance considered, $R(\text{C–C}) = 1.8$ a.u., a pure-state KS solution has been obtained. For the equilibrium distance $R_e(\text{C–C}) = 2.348$ a.u. and at larger distances ensemble solutions have been obtained with widely varying weights of the individual determinants, depending on the bond distance. For CH_2 the dependence of the type of KS solution on the basis has been studied: calculation in the triple zeta (TZ) basis for the KS orbitals yields an ensemble solution, while the pure-state KS solution has been obtained in the quadruple zeta (QZ) basis. The form of the KS orbitals has been compared with that of the natural orbitals (NOs). It has been shown for the model example of the stretched H_2 molecule as well as for CH_2 and C_2, that the KS orbitals of the pure state may be rather different from the corresponding NOs, while the occupied KS orbitals of the ensemble solution can be considered as plausible approximations to the corresponding NOs.

Key words: Density functional theory – Fractional occupation numbers – Ensemble-v-representability – Electron correlation – Kohn-Sham

1 Introduction

For the formulation of the Hohenberg-Kohn density functional theory (DFT) the v-representability of densities $\rho(\mathbf{r})$ is an important issue: given a proper density $\rho(\mathbf{r})$ (nonnegative, integrating to N electrons), does there exist a local potential $v(\mathbf{r})$ such that $\rho(\mathbf{r})$ corresponds to: (a) a ground state Ψ of the system of interacting electrons moving in the potential $v(\mathbf{r})$ [Pure-state v-representable (PS-V); Ψ may be degenerate or not]; or (b) an ensemble of ground-state wave functions in case of degenerate ground state [ensemble v-representable (E-V)]? The ensemble representable density is obtained from a density matrix $\hat{M}$ representing a mixture (ensemble) of ground states,

$$\begin{aligned} &\hat{M} = \sum d_i |\Psi_{0,i}\rangle\langle\Psi_{0,i}| \quad d_i = d_i^* \geq 0, \sum d_i = 1 \\ &\rho(\mathbf{r}) = tr\hat{M}\hat{\rho}(\mathbf{r}) = \sum d_i \rho_{0,i}(\mathbf{r}) \end{aligned} \tag{1}$$

The diagonal density obtained is a convex combination of pure ground-state densities. Such densities have acquired special significance in the theory since it was proven by Levy [1] and by Lieb [2] that a convex combination of ground-state densities is not, in general, pure-state v-representable. The implication is that there is a large set of perfectly "normal looking" densities that do not correspond to a ground-state wave function.

We will not primarily consider the question of representability of densities by ground states of interacting electron systems, but we will be concerned with a similar problem that arises in the context of the Kohn-Sham (KS) approach of DFT. Kohn and Sham introduce a system of noninteracting electrons moving in a local potential $v_s(\mathbf{r})$, which has the same density $\rho(\mathbf{r})$ as the interacting system considered [3]. We will restrict ourselves here to the case that the interacting system has a nondegenerate ground state. Then the basic Ansatz of the KS theory is the v_s-representability of the given $\rho(\mathbf{r})$ [1,2,4-6] which can be formulated as follows. Suppose that, for an interacting system, $\rho(\mathbf{r})$ follows from a nondegenerate ground-state wave function Ψ, then there exists a local potential $v_s(\mathbf{r})$ such that $\rho(\mathbf{r})$ also corre-

---

Correspondence to: E.J. Baerends

sponds to a one-determinantal ground-state wave function Ψ_s of the noninteracting system. The original KS assumption may be extended, in view of the importance of ensemble-representable densities, to include the case that $v_s(\mathbf{r})$ leads to a degenerate ground state of the noninteracting system and $\rho(\mathbf{r})$ corresponds to an ensemble of ground states $\Psi_{s,i}$. We denote these possibilities as pure-state v_s-representable (PS-V_s) or ensemble v_s-representable (E-V_s).

One may wonder if a density corresponding to a nondegenerate ground state of the interacting system will ever fail to be PS-V_s representable. It is the purpose of this paper to identify such a situation and to emphasize that the E-V_s representability is not just an issue of mathematics but that there is a clear connection with the physical phenomenon of electron correlation in the interacting system. The v_s representability is studied for the case of the nondegenerate lowest singlet states of the molecules CH_2 ($^1A_1$ in C_{2v}) and C_2($^1\sum_g^+$ in $D_{\infty h}$), the KS solution being obtained from the corresponding correlated CI wave function Ψ with the iterative procedure of [7] (cf. also [8]). The molecules CH_2 and C_2 have been chosen as examples of systems with an essentially multideterminantal character of Ψ. For the lowest singlet state of CH_2 the importance of inclusion in the CI expansion of a configuration with $\sigma \rightarrow \pi$ excitation has been established, and for many years calculations of this molecule served as a benchmark for ab initio quantum chemistry [9–11]. For the ground state of C_2 the CI expansion exhibits for the equilibrium bond distance R_e(C–C) an exceptionally large contribution of excited configurations, compared to other dimers A_2 of the second period [12]. The main configuration D_0 changes with changing distance R(C–C): for shorter R(C–C) (including R_e) D_0 represents a valence state with two $2p$-based π-bonds and no σ-bond [configuration $\ldots(2\sigma_g)^2(2\sigma_u)^2(3\sigma_g)^0(1\pi_u)^4$], while for longer R(C–C) D_0 represents a valence state with one π- and one $2p$-based σ-bond [configuration $(2\sigma_g)^2(2\sigma_u)^2(3\sigma_g)^2(1\pi_u)^2$]. Another important configuration, in particular at short distances, is $(2\sigma_g)^2(2\sigma_u)^0(3\sigma_g)^2\,(1\pi_u)^4$, with nominally two π bonds and two σ bonds, one $2s$-based $[(2\sigma_g)^2]$ and one $2p$-based $[(3\sigma_g)^2]$.

There is an obvious connection between the use of fractional occupation numbers (FON) in (approximate) DFT calculations and the question of E-V_s representability of a ground state density, since a density matrix representing an ensemble of pure determinantal states leads to a diagonal density that can be written as a sum of orbital densities with fractional occupations [13]. Even prior to the discussions of v- and v_s-representability, and the recognition of the importance of ensemble-representable densities, FONs had been introduced by Slater et al. [14] in the Xα method in an intuitive manner. Dunlap [15, 16] observed in Xα calculations on C_2 and Si_2 that an improved $^1\sum_g^+$ (not the ground state in Xα) potential energy curve was obtained when allowing for fractional occupation of the π_u and $3\sigma_g$ orbitals. Recently Wang and Schwarz [17] have given a clear exposition of the FON method, demonstrating for a series of modern functionals that it gives an improved potential energy curve for the transition from the ring (D_{3h}) to open (C_{2v}) structure of O_3, and similarly for SO_2. These authors have implemented an automatic FON optimization, requiring practically no additional computational effort, into the Amsterdam density function (ADF) code [18, 19]. The FONs receive a rigorous basis if densities have to be represented by an ensemble of pure (determinantal) states of the noninteracting KS system. In this paper we wish to explore and establish the connection between (strong) electron correlation and the need for E-V_s representation of the density. For this purpose we avoid the use of approximate functionals, but generate the noninteracting density and KS potential directly from an accurate correlated ab initio electron density of the interacting system.

This paper is organized as follows. In Sect. 2 alternative types of KS solutions are considered and the situation with respect to the proof of the representability of ρ for both interacting and noninteracting systems is discussed. In Sect. 3 the effect of electron correlation on the type of the KS solution and on the form of the KS orbitals is discussed. For an interacting system ρ is conventionally represented with the natural orbitals (NOs)$\{\chi_i\}$ [20], while for a noninteracting system it is represented with the KS orbitals $\{\psi_i\}$. Naturally, the question (related to the representability problem) arises: what are the differences and similarities between the NOs and KS orbitals? The example of the (stretched) H_2 molecule is considered, for which the NOs and KS orbitals are shown to be distinctly different from each other.

An iterative procedure for the construction of the KS orbitals and the KS potential from the ab initio CI wave function is outlined. The situation is considered where this procedure combined with the constraint of integer occupations of the KS orbitals, leads to a solution with unoccupied orbitals having an energy lower than that of the highest occupied orbitals ("holes" below the Fermi level). A procedure denoted "evaporation of the hole below the Fermi level" [15, 17] is employed in this case, which leads to an ensemble KS solution with accidental degeneracy of the corresponding Slater determinants.

In Sect. 4 the KS solution obtained from the CI wave function of the CH_2 molecule is presented. In this case the electron correlation is just not strong enough to enforce E-V_s representability. The dependence of the KS solution on the size of the basis set and the CI expansion used is investigated. As will be shown in Sect. 4, in borderline cases where there possibly is a weak contribution of a second KS determinant, the type of solution may actually depend on the basis set size. In CH_2 it becomes possible only with a fairly large basis to obtain a one-determinantal (pure-state) KS solution, while for a smaller basis the ensemble solution is obtained.

In Sect. 5 the dependence of the type of KS solution on the bond distance R(C–C) is studied for the C_2 molecule. In this case the possibility will be demonstrated, for a density corresponding to a nondegenerate pure state of the interacting system, of an essentially accurate ensemble KS solution with accidental (not symmetry dictated) degeneracy of the KS determinants. Upon increase of

R(C–C), the KS solution evolves from a pure state to an ensemble with a "weak" involvement of an accidentally degenerate state, to an ensemble with a "strong" degeneracy, and then to another ensemble with a "weak" accidental degeneracy. For both CH_2 and C_2 one- and two-dimensional plots of the constructed exchange-correlation KS potentials v_{xc} are presented and interpreted. As will be shown, for the pure-state KS solution the orbitals may be rather different from the corresponding NOs, while the ensemble solution brings KS orbitals closer to NOs. In Sect. 6 the conclusions are drawn.

2 Representation of the electron density in the KS theory

As was noted in Sect. 1 the representability of the exact density ρ of an interacting electron system by means of a noninteracting system is the central problem of the KS theory. Let ρ be the density of a nondegenerate ground state (pure state, PS) Ψ of the interacting closed-shell system with the external potential v_{ext} so that it is denoted as a PS-V representable density. Then the following three possibilities may, in principle, occur for its representability by a noninteracting system.

2.1 PS-V_{s} representability

ρ is also the density of a pure ground state Ψ_{s} of a noninteracting KS system with the potential v_{s}, so that it is denoted as PS-V_{s} representable. The KS potential v_{s} is related to the external potential v_{ext} as follows:

$$v_{\mathrm{s}}(\mathbf{r}) = v_{\mathrm{ext}}(\mathbf{r}) + v_{\mathrm{H}}(\mathbf{r}) + v_{\mathrm{xc}}(\mathbf{r}), \tag{2}$$

where v_{H} is the Hartree potential of the electrostatic electron repulsion and v_{xc} is the exchange-correlation potential. In general, Ψ_{s} is a few-determinantal function built from K degenerate one-determinantal ground states $D_{\mathrm{s}j}$

$$\Psi_{\mathrm{s}} = \sum_{j=1}^{K} c_j D_{\mathrm{s}j} \tag{3}$$

$$\hat{H}_{\mathrm{s}} D_{\mathrm{s}j} = E_{\mathrm{s}} D_{\mathrm{s}j} \tag{4}$$

$$\hat{H}_{\mathrm{s}} = \sum_{i=1}^{N} \left\{ -\frac{1}{2}\nabla_i^2 + v_{\mathrm{s}}(\mathbf{r}_i) \right\} \tag{5}$$

In the simplest case the KS ground state is nondegenerate and Ψ_{s} is just one determinant D_{s}. In this case ρ is expressed in the form

$$\rho(\mathbf{r}) = 2\sum_{i=1}^{N/2} |\psi_i(\mathbf{r})|^2 \tag{6}$$

with integer occupations $(n_i = 2)$ of the $N/2$ lowest-energy KS orbitals. Consider the case that the interacting Ψ for a closed-shell N electron system is nondegenerate and its configuration interaction (CI) expansion is dominated by a single Slater determinant D_0 built from $N/2$ lowest Hartree-Fock (HF) orbitals φ_i. Then, Ψ_{s} is also expected to be a determinant D_{s} built from $N/2$ lowest KS orbitals ϕ_i. This is confirmed by the essentially accurate KS solutions obtained from ab initio CI wave functions for closed shell atoms [21, 22] and molecules [23–27] with a nondegenerate ground state. The KS solutions Ψ_{s} obtained until now are all single determinants D_{s}. Below in the text we shall consider only the one-determinantal pure KS state, so that (unless otherwise stated explicitly) the abbreviation PS for KS systems will refer to just these one-determinantal states.

2.2 E-V_{s} representability

ρ corresponds to an N-particle density matrix $\hat{M}_{\mathrm{s}}$ of the KS system, which represents an ensemble of L wave functions $\Psi_{\mathrm{s}j}$ of the type (3)

$$\hat{M}_{\mathrm{s}} = \sum_{j=1}^{L} d_j |\Psi_{\mathrm{s}j}\rangle\langle\Psi_{\mathrm{s}j}| \tag{7}$$

Obviously, the set of E-V_{s} representable ρ contains all the PS-V_{s} representable ρ, since (3) is a particular case of (7) with $L = 1$. In the simplest case Ψ_{s} in (7) are just ground-state determinants D_{sj} of (4)

$$\hat{M}_{\mathrm{s}} = \sum_{j=1}^{L} d_j |D_{\mathrm{s}j}\rangle\langle D_{\mathrm{s}j}| \tag{8}$$

Since the D_{sj} are degenerate ground-state wave functions of the noninteracting system, with a noninteracting energy E_{s} that is equal to the sum of the orbital energies, the D_{sj} have the same KS orbitals ψ_i with energies below the Fermi level, $\varepsilon_i < \varepsilon_{\mathrm{F}}$ and they differ from each other in, at least, one highest occupied orbital ψ_i^{HO}, the latter orbitals being all degenerate at $\varepsilon_{\mathrm{F}}, \varepsilon_i^{\mathrm{HO}} = \varepsilon_{\mathrm{F}}$. The orbitals ψ_i^{HO} may have a symmetry-related degeneracy, i.e., belong to different subspecies of a certain irreducible representation of the symmetry group of the molecule. They may also belong to different irreducible representations, i.e., exhibit accidental degeneracy. In that case two determinantal wave functions D_{sj} and D_{sk} may still be of the same symmetry (totally symmetric if for instance each of them has a highest occupied orbital that belongs to a one-dimensional irreducible representation and is doubly occupied). In that case two states of the noninteracting system of the same symmetry exhibit an accidental degeneracy, a situation that rarely (if ever) occurs in the interacting case. It may occur more readily in noninteracting systems since for two determinants that differ in two spin orbitals there will not be an interaction matrix element. Such accidental degeneracy allows for E-V_{s} representability of the ground-state density which we will see is intimately related to nondynamical correlation effects. Suppose the ground state Ψ of the interacting system is nondegenerate (pure state), but has an essentially multideterminantal character, i.e., apart from the main configuration D_0, some excited configurations D_i bring relatively large contributions to the CI expansion of Ψ

$$\Psi = c_0 D_0 + \sum_i c_i D_i + \ldots \tag{9}$$

The natural question arises, whether in this case just one KS determinant D_{s}, with the orbitals representing independent electrons moving in a local potential, is capable of representing the interacting ρ including the contributions from both the determinants D_0 as well as D_i, or whether the interacting ρ can only be represented by an ensemble of degenerate KS determinants $D_{\mathrm{s}i}$.

If ρ is represented by an ensemble for the noninteracting system (E-V$_{\mathrm{s}}$ representability), it is expressed via the KS orbitals as follows:

$$\rho(\mathbf{r}) = 2\sum_{i:\varepsilon_i<\varepsilon_{\mathrm{F}}} |\psi_i(\mathbf{r})|^2 + \sum_{i:\varepsilon_i=\varepsilon_{\mathrm{F}}} n_i^{\mathrm{HO}} |\psi_i^{\mathrm{HO}}(\mathbf{r})|^2 \tag{10}$$

with fractional occupations n_i^{HO} of the accidentally degenerate orbitals ψ_i^{HO}. The occupations n_i^{HO} are combinations of the weights d_j of $D_{\mathrm{s}j}$ in the ensemble (8) and the integer occupations $\theta_{ji}^{\mathrm{HO}}$ of the orbitals ψ_i^{HO} in the densities $\rho_{\mathrm{s}j}$, which correspond to $D_{\mathrm{s}j}$

$$\rho_{\mathrm{s}j}(\mathbf{r}) = \langle D_{\mathrm{s}j}|\hat{\rho}(\mathbf{r})|D_{\mathrm{s}j}\rangle = \sum_i \theta_{ji}|\psi_i(\mathbf{r})|^2$$

$$n_i^{\mathrm{HO}} = \sum_{j=1}^{L} d_j \theta_{ji}^{\mathrm{HO}} \tag{11}$$

The kinetic energy T_{s} in this case has the form

$$\begin{aligned} T_{\mathrm{s}} &= 2\sum_{i:\varepsilon_i<\varepsilon_{\mathrm{F}}} \int \psi_i^*(\mathbf{r})\left(-\frac{1}{2}\nabla^2\right)\psi_i(\mathbf{r})d\mathbf{r} \\ &\quad + \sum_{i:\varepsilon_i<\varepsilon_{\mathrm{F}}} n_i^{\mathrm{HO}} \int \psi_i^{\mathrm{HO}*}(\mathbf{r})\left(-\frac{1}{2}\nabla^2\right)\psi_i^{\mathrm{HO}}(\mathbf{r})d\mathbf{r} \end{aligned} \tag{12}$$

E^{KS}, which is the expectation value of the full interacting hamiltonian for the KS wave function (or ensemble) taken as a trial wave function for the interacting system, to be distinguished from the noninteracting energy $E_{\mathrm{s}} = \sum_i n_i \varepsilon_i$, has the form

$$E^{\mathrm{KS}} = T_{\mathrm{r}}\{\hat{M}_{\mathrm{s}}\hat{H}\} = \sum_{j=1}^{L} d_j \langle D_{\mathrm{s}j}|\hat{H}|D_{\mathrm{s}j}\rangle \tag{13}$$

so that the exchange energy E_{x} can be defined as a weighted sum of the exchange energies of the individual determinants

$$E_{\mathrm{x}} = \sum_{j=1}^{L} d_j E_{\mathrm{xc}}[\rho_j] \neq E_{\mathrm{xc}}\left[\sum d_j \rho_j\right] \tag{14}$$

where

$$E_{\mathrm{x}}[\rho_j] = -\frac{1}{4}\sum_i \sum_k \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\theta_{ji}\theta_{jk}\psi_i^*(\mathbf{r}_1)\psi_k(\mathbf{r}_1)\psi_k^*(\mathbf{r}_2)\psi_i(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{15}$$

Below we shall denote as E-V$_{\mathrm{s}}$ representable the densities that cannot be represented with a single KS determinant, but that can be represented with an ensemble (8)

2.3 Break-down of the KS approach

ρ cannot be represented with either a pure ground state (including states that are a linear combination of ground-state determinants) or an ensemble, as sect. 2.1 and 2.2. In this case of non-v_{s} representability, it is still possible that ρ can be expressed in the form

$$\rho(\mathbf{r}) = \sum_i n_i|\psi_i(\mathbf{r})|^2 \tag{16}$$

with partial occupation $(n_j < 1)$ of some orbitals ψ_i with energies ε_i below the Fermi level (holes below the Fermi level) [28]. For a noninteracting system, the ground state must obey the Aufbau principle, i.e., the *ground* state of the KS system cannot contain holes, thus (16) corresponds to some excitation of the KS system. This case could manifest a breakdown of the KS Ansatz of noninteracting v_{s}-representability of all interacting ground-state densities.

Unfortunately, although several mathematical studies on the v- and v_{s}-representability problems have appeared [1, 2, 4–6], the v_{s}-representability has not been proven yet for an arbitrary v-representable density. The original [29] HK functional F_{HK} has been extended to the functional $F_{\mathrm{L}}[\rho]$ [2] which searches the infimum of $T_{\mathrm{r}}\{\hat{M}(\hat{T}+\hat{W})\}$ for all density matrices $\hat{M}$ which yield the density ρ

$$\begin{aligned} &F_{\mathrm{L}}[\rho] = \inf_{\hat{M}\to\rho} T_{\mathrm{r}}\{\hat{M}(\hat{T}+\hat{W})\} \\ &\hat{M} = \sum d_i |\Psi_i\rangle\langle\Psi_i| \quad d_i = d_i^* \geq 0, \sum d_i = 1 \end{aligned} \tag{17}$$

The advantage of $F_{\mathrm{L}}[\rho]$ is that it is a convex functional defined on the convex set S_N of all nonnegative densities that integrate to N (not just the set of ground-state densities A_N on which F_{HK} is only defined). $F_{\mathrm{L}}[\rho]$ is differentiable, with v as a tangent functional, at all PS-V and E-V densities (and nowhere else). If we turn to noninteracting systems, for which the two-body interaction is zero, $\hat{W} = 0$, the functional $F_{\mathrm{L}}[\rho]$ reduces to just the kinetic energy functional for noninteracting electrons, $T_{\mathrm{L}}[\rho]$. Again, the differentiability of $T_{\mathrm{L}}[\rho]$ will be assured for PS-V$_{\mathrm{s}}$ and E-V$_{\mathrm{s}}$ densities, with v_{s} as the tangent functional. However, the KS method would only be put on firm ground if one could also prove that every interacting density, i.e., every E-V density, or at least every PS-V density, would also belong to the set of noninteracting v_{s}-representable densities. Since this crucial step has not been taken, it is too optimistic to consider the validity of the KS scheme as being rigorously established [13].

Chen and Stott [30, 31] have considered the problem of v_{s} representability of ρ for atomic-like systems in a central-field external potential. Their analysis deals with the E-V$_{\mathrm{s}}$ representability of a given density by establishing the degeneracy of the highest occupied level and the weights of determinantal states in the ensemble (fractional occupations of the degenerate highest orbitals). They analyzed how the topology of the orbital energy surfaces then determines the type of ensemble that is obtained. They did not prove that every interacting density is v_{s}-representable, since they discussed so-called reasonable densities for which it was assumed a potential v_{s} could be found on a whole domain in (fractional) occupation number space. However, we do not know whether the molecular systems considered have a "reasonable" interacting ground-state density in this sense, so that the v_s-representability of their densities is not assured.

We have, however, not encountered a break-down of the (extended) KS approach.

3 The effect of electron correlation on the form of the KS orbitals and the construction of the KS solution

The effect of electron correlation on the KS orbitals and problems associated with the construction of the KS solution can be illustrated with the simple example of the stretched H_2 molecule considered in detail in Refs. [32, 33]. In this case, as for any other two-electron closed-shell system, the form of the single molecular orbital (MO) ψ_g of the pure-state KS solution is defined (up to a phase factor) straightforwardly from the density ρ as

$$\psi_g(\mathbf{r}) = \sqrt{\frac{1}{2}\rho(\mathbf{r})} \ . \tag{18}$$

From this, one can see a distinct difference between the KS expansion of the density in a set of orbitals and the conventional expansion in quantum chemistry of ρ in terms of the NOs. Indeed, in the KS case ρ is represented for this two-electron system with the single MO (18), which belongs to the totally symmetric irreducible representation (irrep) $\sum_g^+$ of the molecular symmetry point group $D_{\infty h}$. In contrast, the NO expansion for stretched H_2, with the elongated bond distance R(H–H), has two dominant contributions from the NO χ_{1g} of $\sum_g^+$ symmetry and the NO χ_{1u} of $\sum_u^+$ symmetry

$$\rho(\mathbf{r}) \approx n_{1g}^{\mathrm{NO}}|\chi_{1g}(\mathbf{r})|^2 + n_{1u}^{\mathrm{NO}}|\chi_{1u}(\mathbf{r})|^2 \tag{19}$$

The difference between the KS and NO expansions can be considered in quantitative way, if we expand the KS MOs $\{\psi_i\}$ and the NOs $\{\chi_i\}$ in terms of the HF MOs $\{\varphi_i\}$

$$\psi_i(\mathbf{r}) = \sum_j a_j^{i(\mathrm{KS})}\varphi_j(\mathbf{r}) \tag{20}$$

$$\chi_i(\mathbf{r}) = \sum_j a_j^{i(\mathrm{NO})}\varphi_j(\mathbf{r}) \tag{21}$$

and then form the KS and NO representations b_{ij}^{KS} and b_{ij}^{NO} of ρ in terms of products $\varphi_i^*\varphi_j$ of the HFMOs

$$\rho(\mathbf{r}) = \sum_i\sum_j b_{ij}^{\mathrm{KS}}\varphi_i^*(\mathbf{r})\varphi_j(\mathbf{r}) = \sum_i\sum_j b_{ij}^{\mathrm{NO}}\varphi_i^*(\mathbf{r})\varphi_j(\mathbf{r}) \tag{22}$$

where

$$b_{ij}^{\mathrm{KS}} = \sum_k n_k^{\mathrm{KS}} a_i^{k*(\mathrm{KS})} a_j^{k(\mathrm{KS})} \tag{23}$$

$$b_{ij}^{\mathrm{NO}} = \sum_k n_k^{\mathrm{NO}} a_i^{k*(\mathrm{NO})} a_j^{k(\mathrm{NO})} \tag{24}$$

The b_{ij}^{NO} are just the elements of the one-electron density matrix in the HF MO basis, but the b_{ij}^{KS} are not. Note that the NOs χ_{1g} and χ_{1u} for the stretched H_2 consist, mainly, of the corresponding HF MOs φ_{1g} and φ_{1u}, respectively, which in their turn are, essentially, the bonding and antibonding combinations of 1*s* atomic orbitals (AOs) of the H atoms. Thus, the population b_{1u1u}^{NO} of φ_{1u} in the NO expansion (24) is relatively large. For example, full CI calculation of H_2 at R(H-H) = 5 a.u. in a basis with five *s*- and two *p*-type contracted Gaussian functions and an additional *d*-type Gaussian [32,34,35] yields the values $n_{1u} = 0.736$, $a_{1u}^{1u(\mathrm{NO})} = 0.976$, and $b_{1u1u}^{\mathrm{NO}} = 0.704$. Contrary to this, the HF expansion (20) for the KS orbital ψ_g cannot contain the orbital φ_{1u} of a different symmetry, so that $a_{1u}^{1g(\mathrm{KS})} = 0$ and φ_{1u} has zero population in the KS expansion (23), $b_{1u1u}^{\mathrm{KS}} = 0$. This difference between the KS and NO coefficients b_{1u1u}^{KS} and b_{1u1u}^{NO} determined for the same ρ and the same HF MO product $|\varphi_{1u}|^2$ becomes possible due to the established redundance of the basis of orbital products $\{\varphi_i\varphi_j\}$ [36] in one-particle space.

Lacking the direct contribution of φ_{1u}, the HF MO expansion (20) of the KS orbital ψ_g simulates the contribution of φ_{1u} to the correlated density (19) by the inclusion of the higher orbitals φ_{ng} of $\sum_g^+$ symmetry which consist, mainly, of the *p*- and *d*-AOs of the H atoms. Taken with the sign opposite to that of φ_{1g}, these higher orbitals bring, by means of the product terms $b_{n1}^{\mathrm{KS}}\varphi_{ng}^*(\mathbf{r})\varphi_{1g}(\mathbf{r})$, the hybrid cross-terms of the type $-sp_z$ and $-sd_{z^2}$ (z is the molecular axis) into the density $|\psi_g(\mathbf{r})|^2 = \rho(\mathbf{r})$. The cross-terms reduce the density around the bond midpoint, reproducing effectively the contribution to ρ of the orbital φ_{1u}, which has a nodal plane passing through the bond midpoint. One may wonder if the necessity to build correlation induced features of the density into the occupied KS orbitals leads to special requirements for the primitive basis set. Such a situation has been identified in the case of the dissociating electron pair bond of H_2 in [33]. In that case the σ_g and σ_u NOs can be very well represented in a minimal basis by

$$\psi_{g,u}(\mathbf{r}) = \frac{1}{\sqrt{2 \pm 2S}}[a(\mathbf{r}) \pm b(\mathbf{r})] \tag{25}$$

where $a(\mathbf{r})$ and $b(\mathbf{r})$ are the 1*s* AOs of the H atoms A and B, and S is the overlap integral

$$S = \int a(\mathbf{r})b(\mathbf{r})d\mathbf{r} \tag{26}$$

However, such a minimal basis fails to reproduce the features of the true σ_g KS orbital (18) around the bond midpoint. There is, however, no indication that demands on basis sets in the KS case are comparable to those in MP2 or other correlated calculations. The GGA functionals $E_{xc}[\rho(\mathbf{r})]$ are in fact not sensitive to such small features in the density as we are dealing with at present, and basis sets that are adequate at the SCF level will usually suffice.

In this paper the KS potential v_s and KS orbitals $\{\psi_i\}$ have been obtained from the *ab initio* CI density ρ by iterative solution [7] of the KS equations, with the orbitals ψ_i being expanded in the same basis $\{\varphi_i\}$ of HF MOs as was used in the CI calculations. In this paper a variant of the iterative procedure has been employed, according to which the components $\langle \varphi_i | v_{\mathrm{ext}} | \varphi_j \rangle$ and $\langle \varphi_i | v_{\mathrm{H}}[\rho] | \varphi_j \rangle$ of the matrix elements $\langle \varphi_i | v_{\mathrm{s}} | \varphi_j \rangle$ are fixed and only the unknown exchange-correlation potential $v_{\mathrm{xc}}(\mathbf{r})$ is updated during iterations, starting from some initial guess $v^0_{\mathrm{xc}}(\mathbf{r})$. At the mth iteration, the KS equations are solved with the potential $v^m_{\mathrm{xc}}(\mathbf{r})$

$$v_{\mathrm{s}}(\mathbf{r}) = v_{\mathrm{ext}}(\mathbf{r}) + v_{\mathrm{H}}(\mathbf{r}) + v^m_{\mathrm{xc}}(\mathbf{r}) \tag{27}$$

$$v^m_{\mathrm{xc}}(\mathbf{r}) = f^m(\mathbf{r}) v^{m-1}_{\mathrm{xc}}(\mathbf{r}) \tag{28}$$

calculated from the $v^{m-1}_{\mathrm{xc}}(\mathbf{r})$ of the previous iteration with the correction factor f^m. The latter is obtained from the density ρ^{m-1} from the $(m-1)$th iteration and the target density ρ

$$f^m(\mathbf{r}) = \frac{\rho(\mathbf{r}) + a}{\rho^{m-1}(\mathbf{r}) + a} \tag{29}$$

The parameter a smooths out the effect of the remote exponential density tails on the procedure. Matrix elements $\langle \varphi_i | v^m_{\mathrm{xc}} | \varphi_j \rangle$ have been calculated using a numerical integration with grids according to [19]. The accuracy of the resultant KS solution is characterized by the value of the absolute integral error $\Delta\rho$ at the m-th iteration

$$\Delta\rho = \int |\rho^m(\mathbf{r}) - \rho(\mathbf{r})| d\mathbf{r} \tag{30}$$

As the first option, we always attempt to construct the pure state KS solution, so that a trial density ρ^m is formed according to (6) with the doubly occupied KS counterparts ψ_i of the $N/2$ lowest energy HF orbitals φ_i. When the pure state is attainable, this leads to the KS solution, which reproduces (up to a small error $\Delta\rho$) the target correlated density and complies with the "Aufbau principle", i.e. all the occupied orbitals have lower energies than the unoccupied ones. On the other hand, if the pure state is unattainable, the iterative procedure (27–29) converges to a "non-Aufbau" solution, for which $\Delta\rho$ is, usually, also small, but the lowest unoccupied MO (LUMO) has a lower energy than the highest occupied MO (HOMO), i.e., a hole below the Fermi level appears. If the LUMO has a different symmetry than the HOMO, this situation may arise during the iterations when a fixed number of orbitals is occupied in each irreducible representation. An attempt to obtain the KS solution with reversed occupation, with the hope to achieve Aufbau, fails, since with this occupation pattern the iterative procedure does not converge to a solution with a small $\Delta\rho$. In case one is performing a standard self-consistent KS calculation using approximate functionals, this failure to accurately describe the correlated density would of course not be noticed, but in this case reversal of the occupation pattern will typically again lead to a non-Aufbau situation, since filling the hole in the orbital below the Fermi level will raise the orbital energy of the former hole orbital, and lower the energy of the former "HOMO" that is now de-occupied, to the effect that the orbital energies cross and Aufbau is again violated (cf. Fig. 1 of [17]).

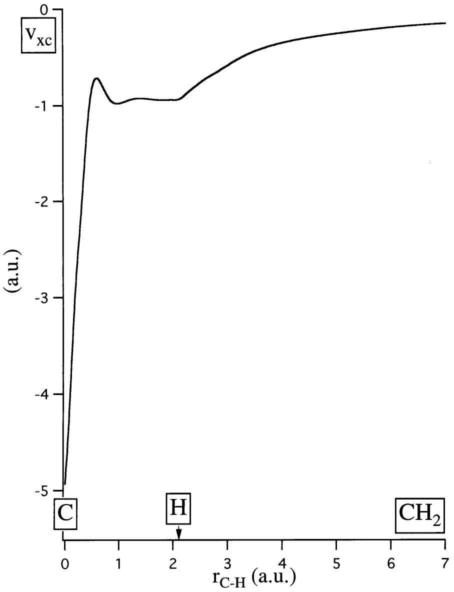


Fig. 1. The exchange-correlation potential v_{xc} along the C–H bond of CH_2. The C nucleus is in the origin and the H nucleus is at 2.099 bohr

In order to provide the KS solution for this case we have used a procedure of "evaporation of the hole below the Fermi level" as given in [37]. It consists in fractional occupation of the LUMO $\psi_{(N/2+1)}$, starting with $n_{(N/2+1)} = 0$ with simultaneous depopulation of the HOMO $\psi_{N/2}$ when forming a trial density ρ^m (we continue to number these orbitals as $N/2+1$ and $N/2$ respectively, although their order has reversed). This redistribution of the electrons increases the energy of the former and decreases the energy of the latter orbital. At certain fractional occupations $n_{(N/2+1)}$ and $n_{N/2}$ the orbital energies $\varepsilon_{N/2}$ and $\varepsilon_{(N/2+1)}$ are equal to each other and the target density ρ is reproduced with an ensemble KS solution. These energies will define the Fermi level energy ε_{F}, $\varepsilon_{(N/2+1)} = \varepsilon_{N/2} = \varepsilon_{\mathrm{F}}$, so that there will be no fractional occupations below the Fermi level. The corresponding density assumes the form (10) of the ensemble representable (E-V$_{\mathrm{s}}$)ρ. The procedure is analogous to the one applied in Xα [15] and generalized gradient approximation (GGA) [17] calculations, the difference being that our KS calculations are not approximate, but are aimed at reproducing the correlated density. Our results therefore will not be artifacts of the Xα or some local-density (LDA) or density-gradient approximation (DGA).

Fortunately, in our actual calculations we managed in all cases to represent ρ either with the pure-state form (6) or with the ensemble form (10). Thus, we did not encounter the breakdown of the KS theory (option 3 of the preceding section) though, in principle, one cannot exclude such a possibility.

4 Results for CH_2

The correlated density ρ of the CH_2 and C_2 molecules, the basic quantity for construction of the KS solution, has been obtained with configuration interaction (CI) calculations performed by means of the ATMOL package [38]. The lowest singlet state of CH_2 has been calculated at its equilibrium geometry [39], with the bond distance $R(\text{C–H}) = 2.099$ a.u. and angle $\angle(HCH) = 102.4°$. The KS potential and orbitals have been constructed in the same basis of HF MOs as was used in the CI calculations by means of a Gaussian orbital density functional code [32, 40] based on the ATMOL package.

The correlation-consistent polarized core-valence (cc-pCV) basis sets of Woon and Dunning [41] of contracted Gaussian functions have been used for the calculations. In order to describe properly various correlation effects, a large number of polarization functions of higher angular momenta of both core and valence size are included in these sets. In order to study the effect of the basis on the KS solution for CH_2, the calculations in a triple zeta basis augmented with additional polarization functions (aug-cc-pCVTZ) are compared with those performed in a quadruple zeta (cc-pCVQZ) basis. To study the effect of the size of the CI expansion, CI calculations with all single and double excitations from the single reference HF configuration (SRCI) are compared with multireference CI(MRCI) calculations. Using all possible excitations of the HF configuration in the internal space of six lowest HF MOs produced 90 reference configurations for the MRCI. The MRCI calculation performed in the largest cc-pCVQZ basis yields the total energy $E = -39.115$ Hartrees, which is only 0.002 Hartrees higher than that obtained in the benchmark calculation of CH_2 with the coupled cluster method with all single and double excitations (CCSD) by Grev and Schaefer [42].

The electron correlation produces a peculiar effect on the electron distribution of CH_2. The main HF configuration in the CI expansion for the singlet state is $1a_1^2 2a_1^2 1b_2^2 3a_1^2$, which corresponds to two σ-bonds C–H ($2a_1^2$ and $1b_2^2$) and the σ-electron lone pair ($3a_1^2$). This can be seen from Table 1, where the populations b_{ij}^{NO} Eq. (24) of the products $\varphi_i \varphi_j$ of HF MOs in the NO representation of ρ are presented. However, admixture of the excited HF configuration $1a_1^2 2a_1^2 1b_2^2 1b_1^2$ in the CI expansion provides a relatively small, but non-negligible population $b_{1b_1 1b_1}^{\text{NO}} = 0.067$ of the orbital $\varphi(1b_1)$, which is the π-type orbital oriented perpendicular to the molecular plane.

Construction of the KS solution in the largest cc-pCVQZ basis and with the density obtained from the largest (MR)CI produces a pure state with the configuration of the KS orbitals $1a_1^2$ (i.e., all the occupied KS orbitals are of σtype). The coefficients b_{ij}^{KS} of the corresponding KS representation of ρ in terms of the HF MO products $\varphi_i \varphi_j$ are also displayed in Table 1. In complete analogy with the stretched H_2 molecule discussed in the previous section, the HF MO $\varphi(1b_1)$ has zero population in the KS representation $b_{1b_1 1b_1}^{\text{KS}} = 0$, since its symmetry ($\pi$type) is different from the σtype symmetry of all occupied KS orbitals. The contribution of $|\varphi(1b_1)|^2$ to the target density ρ in the NO expansion is replaced in the KS representation by an admixture of higher orbitals φ_i of a_1 symmetry in the HF MO expansion of the occupied KS orbitals $\psi(2a_1)$ and $\psi(3a_1)$, which results in enhanced populations of the cross-products $\varphi(2a_1)\varphi(na_1)$ and $\varphi(3a_1)\varphi(na_1)$ $(n > 3)$. In Table 1 the populations of the cross-products $\varphi(2a_1)\varphi(4a_1)$ and $\varphi(3a_1)\varphi(4a_1)$ are compared for the NO and KS representation and, indeed, the populations for the latter representation are considerably larger than those for the former representation.

Table 1. Populations b_{ij}^{NO} and b_{ij}^{KS} of the products of the Hartree-Fock MOs φ_i and φ_j in the Kohn-Sham and natural orbital representations of the correlated density for CH_2

b_{ij}	$1a_1 1a_1$	$2a_1 2a_1$	$1b_2 1b_2$	$3a_1 3a_1$	$1b_1 1b_1$	$2a_1 4a_1$	$3a_1 4a_1$
NO	1.998	1.964	1.960	1.899	0.067	0.011	−0.020
KS	2.000	1.998	2.000	1.998	0.0	0.046	−0.063

Thus, the basis of HF MOs, in which all calculations have been performed, should include a sufficient number of orbitals of the relevant symmetry (a_1 orbitals in this case) in order to represent properly the correlation effect on the form of KS orbitals. The basis set quality may even affect the type (one-determinantal pure state or ensemble) of KS solution that is obtained. Table 2 compares the energies and occupations of the frontier KS orbitals $\psi(3a_1)$ and $\psi(1b_1)$ as well as the KS and HF energies obtained with the aug-cc-pCVTZ and cc-pCVQZ basis sets and with SRCI and MRCI expansions. One can see from Table 2 that the type of the KS solution depends essentially on the basis set size. In particular, at the TZ level the pure state "Aufbau" KS solution is obtained for a limited (SR) CI, with the fully occupied KS orbital $\psi(3a_1)$ having a lower energy than the empty orbital $\psi(1b_1)$. However, when more correlation is included via the MRCI, calculation at the triple zeta level fails to produce a PS solution. Instead, a "non-Aufbau" situation results, with the empty orbital $\psi(1b_1)$ having a lower energy than the fully occupied $\psi(3a_1)$ (a hole below the Fermi level). The procedure of "evaporation of a hole below the Fermi level" described in the previous section produces in this case an ensemble KS solution with a very small occupation (0.008) of the orbital $\psi(1b_1)$, which becomes degenerate with $\psi(3a_1)$. Only at the QZ level do calculations for both SRCI and MRCI yield the pure state "Aufbau" KS solution.

Note that the basis set effect on the KS energy characteristics is not very pronounced (Table 2). The KS kinetic energies T_s are somewhat increased when going from an aug-cc-pCVTZ a-pCVQZ basis, being consistently higher than their HF counterparts. However, the DFT exchange energies E_x and the KS expectation values E^{KS} calculated at various basis set and CI levels appear to be rather close to each other and to the corresponding exchange E_x^{HF} and total E^{HF} HF energies. This similarity for these specific energies (as opposed to other energy components) for KS and HF solutions has been observed before [25].

In Fig.1 the exchange-correlation potential v_{xc} of the pure state constructed for CH_2 with the most extensive cc-pCVQZ MRCI calculation is plotted along the C–H

Table 2. Kohn-Sham and Hartree-Fock energy characteristics (a.u.) and occupations of the frontier Kohn-Sham orbitals for CH_2. The Kohn-Sham orbitals reproduce target densities obtained from various types of CI calculations. SR: all singles and doubles from the HF reference configuration; MR: all singles and doubles from 90 reference configurations; TZ: aug-cc-pCVTZ basis [41]; QZ: cc-pCVQZ basis [41]

ρ from CI: KS solution:	SR-TZ PS	MR-TZ Hole	MR-TZ Ensemble	SR-QZ PS	MR-QZ PS
Kohn-Sham					
$\Delta\rho$	0.0061	0.0069	0.0062	0.0057	0.0057
$\varepsilon(3a_1)$	−0.391	−0.387	−0.390	−0.394	−0.389
$\varepsilon(1b_1)$	−0.363	−0.398	−0.390	−0.357	−0.384
$n(3a_1)$	2.0	2.0	1.992	2.0	2.0
$n(1b_1)$	0.0	0.0	0.008	0.0	0.0
T_s	38.910	38.900	38.900	38.940	38.932
E_x	−5.781	−5.776	−5.777	−5.786	−5.780
E^{KS}	−44.912	−44.909	−44.909	−44.913	−44.911
Hartree-Fock					
Basis:	TZ			QZ	
T^{HF}	38.846			38.860	
$E_x{}^{HF}$	−5.788			−5.790	
E^{HF}	−44.916			−44.918	

bond axis as a function of the distance z from the C nucleus. In v_{xc} there is a well around the C nucleus, with a depth of -5 a.u. that agrees well with that found in [22] for a free C atom. The well is terminated with the typical local maximum (intershell peak) between the core and valence regions of atom C. In the region of the C–H bond v_{xc} has a flat form and, remarkably, the location of the H nucleus at 2.1 a.u. can be recognized only by the change of the slope of v_{xc} at that point. This is a characteristic feature of a covalent bond X–H, which has been established previously also for the BH molecule [24]. For larger z the potential v_{xc} approaches smoothly its Coulombic asymptotics $v_{xc}(z) \rightarrow -1/z$.

Figure 2 is a contour plot of v_{xc} in the plane that is perpendicular to the molecular plane and crosses it along the bisector of the valence angle H–C–H, thus representing the region of the electron lone pair of CH_2. The C nucleus is at the position (0,0) and the positive part of the z axis is the bisector of the acute angle H–C–H. The distribution of v_{xc} appears to characterize the nonbonding (lone-pair) character of the HOMO $\psi(3a_1)$, being slightly more attractive in the direction away from the H atoms (negative z values). Along the z axis v_{xc} displays intershell peaks near $z = \pm 0.6$ a.u. and then local wells near $z = \pm 1.0$ a.u., which indicate the region of the lone pair of the C atom at negative z values. Note that, in the direction perpendicular to the molecular plane, v_{xc} lacks the intershell peak and that it increases monotonically with increasing distance from the C nucleus.

The results of this section show that even for CH_2, for which the wave function Ψ has a relatively weak multideterminantal character, great care should be exercised in order to obtain a KS solution of the proper form. Only at the QZ level does it become possible to construct the pure-state solution. As will be shown in the next section, in calculations at this level for C_2, for which the wave function Ψ has a stronger multideterminantal character, a great variety of KS solutions is produced for various bond distances R(C–C), ranging from a pure state through an ensemble with a weak accidental degeneracy to an ensemble with a strong accidental degeneracy.

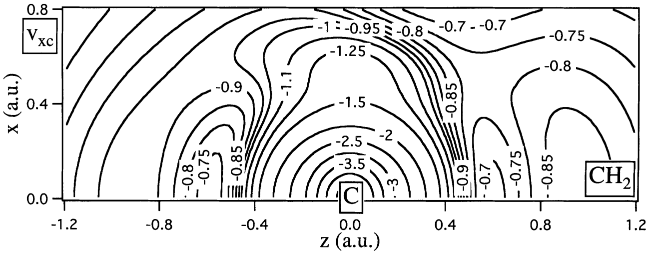


Fig. 2. Contour plot of the exchange-correlation potential v_{xc} of CH_2 in a plane perpendicular to the molecular plane. The C nucleus is at the position (0,0) and the positive part of the z axis is the bisector of the acute angle H-C-H; the electron lone pair of CH_2 is located around the negative part of z axis

5 Results for C_2

CI calculations for the singlet ground state of the C_2 molecule with subsequent construction of the KS solution have been performed at its equilibrium bond distance $R_e = 2.348$ a.u. [39], as well as at the shorter distance R(C–C) = 1.8 a.u. and at larger distances R(C–C) = 2.8, 3.0, 3.2 and 4.0 a.u. The large cc-pCVQZ basis set described in the previous section has been used for the C atoms and MRCI calculations have been performed to obtain the correlated ρ. The MRCI employed 27 reference configurations selected in the active space of the nine lowest HF MOs. All the configurations needed to provide the proper dissociation limit of C_2[34] are included in the reference space. At this level the MRCI calculation performed at R_e is able to recover 90% of the Coulomb correlation energy of C_2 estimated empirically from spectroscopic data [43].

C_2 is an exceptional molecule since at the equilibrium bond distance R_e(C–C) it can be approximately described as having two π bonds and no σ bond. Whole classes of chemical compounds are known (the saturated hydrocarbons, for example), which have only σ bonds and no π bonds, but the reverse situation is very rare, indeed. One can see this feature of C_2 from the data in Table 3 where the populations b_{ij}^{NO} (Eq. 24) of the products $\varphi_i\varphi_j$ of HF MOs in the NO representation of ρ are presented for the bond distances considered. The main HF configuration in the CI expansion at R_e(C–C) is $2\sigma_g^2 2\sigma_u^2 1\pi_u^4$ (we omit the core orbitals $1\sigma_g$ and $1\sigma_u$), corresponding to two π bonds ($1\pi_u^4$) and no σ bond (both the 2*s*–2*s* bonding and antibonding orbitals are occupied: $2\sigma_g^2 2\sigma_u^2$). The weight of this configuration is $c_0 = 0.857$, and the admixture of the excited configuration $2\sigma_g^2 1\pi_u^4 3\sigma_g^2$ (both a 2*s* and a 2*p* σ bond) with the coefficient $c_1 = -0.333$ produces net depopulation of the σ antibonding $2\sigma_u$ HF MO and population of the p_z–p_z σ bonding MO $3\sigma_g$ (Table 3). The configuration $2\sigma_g^2 2\sigma_u^2 1\pi_u^2 3\sigma_g^2$, corresponding to excitation of two π electrons to the 2*p*–2*p* σ bond $3\sigma_g$, does not contribute much at R_e.

At a shorter distance R(C–C) = 1.8 a.u., the relative destabilization of the $2\sigma_u$ MO leads to stronger involvement of $2\sigma_g^2 1\pi_u^4 3\sigma_g^2$, i.e., the $2\sigma_u$ population further decreases and that of $3\sigma_g$ increases, so that one can speak of partial formation of both 2*s* and 2*p* based σ bonds, while the two π bonds are fully retained. At these short distances one cannot unequivocally identify occupation of $3\sigma_g$ with formation of a 2*p*σ bond, since the $2p_z$ orbitals do not overlap particularly favorably due to the nodal plane, and in addition the $2p_z$ has to be made orthogonal to the opposite 2*s* (orthogonality of $3\sigma_g$ to $2\sigma_g$). Elongation of the C–C bond from its R_e value stabilizes the $2\sigma_u$ MO, which becomes completely occupied (note the equivalence of $2\sigma_g^2 2\sigma_u^2$ to the presence of two 2*s* lone pairs). It also leads to a relative weakening of the 2*p*-based π bonds compared to the 2*p*-based σ bond. As a result, the population of $3\sigma_g$ again increases, this time in expense of the π bonding MO $1\pi_u$. At R(C–C) = 2.8 a.u. $1\pi_u^4 \rightarrow 1\pi_u^2 3\sigma_g^2$ becomes the leading excitation and at R(C–C) = 3.0 a.u. the wave function Ψ has a strong multideterminantal character, with the determinant $\Psi_0 = |\sigma_g^2 2\sigma_u^2 1\pi_u^4|$ and the configuration state function (CSF) $\Psi_1 = |(2\sigma_g^2 2\sigma_u^2 1\pi_u^2 3\sigma_g^2)^1 \sum_g^+\rangle = \frac{1}{\sqrt{2}}(|2\sigma_g^2 2\sigma_u^2 1\pi_{ux}^2 3\sigma_g^2| + |2\sigma_g^2 2\sigma_u^2 1\pi_{uy}^2 3\sigma_g^2|)$ being mixed with the coefficients $c_0 = 0.638$ and $c_1 = -0.624$. This result in a large population of the HF MO $3\sigma_g$ (Table 3), which at these longer distances is clearly $2p_z$–$2p_z$ σ bonding.

At R(C–C) = 3.2 a.u. the CSF $|(2\sigma_g^2 2\sigma_u^2 1\pi_u^2 3\sigma_g^2)^1 \sum_g^+\rangle$ starts to dominate with the coefficient $c_1 = -0.854$, while $2\sigma_g^2 2\sigma \mathrm{u}^2 1\pi_u^4$, which dominates at shorter R(C–C), has a much smaller contribution $c_0 = 0.302$. At R(C–C) = 4.0 a.u. the contribution of $2\sigma_g^2 2\sigma_u^2 1\pi_u^4$ reduces further, $c_0 = 0.212$, and the population of the HF MO $1\pi_u$ reduces to $b^{NO}_{1\pi_{ux}1\pi_{ux}} + b^{NO}_{1\pi_{uy}1\pi_{uy}} = 1.738$, while that of $3\sigma_g$ increases to $b^{NO}_{3\sigma_g 3\sigma_g} = 1.782$. Note also an increase of population of the π-antibonding MO $1\pi_g$ with increasing bond distance. The observed trends are in qualitative agreement with those reported in [12].

The demonstrated essentially multideterminantal character of the wave function has a remarkable effect on the corresponding KS solution for the C_2 molecule. Table 4 presents the density error (30) of the KS solution, energies of the valence KS orbitals and changing total occupation $n(1\pi_u)$ of the KS orbitals $\psi(1\pi_{ux})$ and $\psi(1\pi_{uy})$ and occupation $n(3\sigma_g)$ of the orbital $\psi(3\sigma_g)$ (the KS orbitals $\psi(2\sigma_g)$ and $\psi(2\sigma_u)$ are always doubly occupied and $\psi(1\pi_g)$ is always empty). At the shortest distance R(C–C) = 1.8 a.u. a one-determinantal pure-state KS solution is obtained with the configuration $2\sigma_g^2 2\sigma_u^2 1\pi_u^4$ and with $\psi(2\sigma_u)$ as the HOMO at −0.493 a.u. just above the fully occupied $1\pi_u$ at −0.505 a.u. and just below the $3\sigma_g$ at −0.478 a.u.. However, at R_e and R(C–C) = 2.8,3.0 a.u. attempts to obtain pure-state so-

Table 3. Populations b_{ij}^{NO} of the products of the Hartree-Fock MOs φ_i and φ_j in the natural orbital representation of the correlated density for C_2. The b_{ij}^{NO} are the elements of the one-electron density matrix in the HF MO basis. For $1x_u$ $1\pi_u$ the sum of the *x* and *y* components is given, $b^{NO}_{1\pi_{ux}1\pi_{ux}} + b^{NO}_{1\pi_{uy}1\pi_{uy}}$

R(C–C), a.u.	1.8	2.348	2.8	3.0	3.2	4.0
			b_{ij}^{NO}			
$2\sigma_g 2\sigma_g$	1.973	1.966	1.962	1.961	1.958	1.952
$2\sigma_u 2\sigma_u$	1.415	1.669	1.819	1.896	1.928	1.934
$1\pi_u 1\pi_u$	3.896	3.804	3.502	2.824	2.062	1.738
$3\sigma_g 3\sigma_g$	0.560	0.316	0.371	0.962	1.722	1.782
$1\pi_g 1\pi_g$	0.016	0.060	0.112	0.113	0.097	0.209
$2\sigma_g 3\sigma_g$	−0.152	−0.201	−0.209	−0.130	0.00	−0.013

Table 4. Energy characteristics (a.u.), occupations of the frontier Kohn-Sham orbitals and density errors $\Delta\rho$ of the Kohn-Sham (pure state or ensemble) and Kohn-Sham-like (hole below Fermi level) solutions for C_2

R(C–C), a.u.	1.8	2.348	2.8	3.0	3.2	4.0
Pure state (at 1.8 a.u.) and ensemble (other R) Kohn-Sham solutions						
ε($2\sigma_g$)	−0.950	−0.895	−0.831	−0.806	−0.792	−0.710
ε($2\sigma_u$)	−0.493	−0.538	−0.564	−0.573	−0.580	−0.617
ε($1\pi_u$)	−0.505	−0.456	−0.431	−0.420	−0.414	−0.381
ε($3\sigma_g$)	−0.478	−0.456	−0.431	−0.420	−0.414	−0.381
ε($1\pi_g$)	−0.085	−0.187	−0.253	−0.270	−0.284	−0.323
$n(1\pi_u)$	4.0	3.912	3.652	3.019	2.201	2.050
$n(3\sigma_g)$	0.0	0.088	0.348	0.981	1.799	1.950
$\Delta\rho$	0.012	0.004	0.008	0.008	0.007	0.015
Kohn-Sham solutions with a hole below the Fermi level						
ε($2\sigma_g$)	–	−0.894	−0.843	−0.864	−0.788	−0.708
ε($2\sigma_u$)	–	−0.551	−0.604	−0.695	−0.574	−0.616
ε($1\pi_u$)	–	−0.450	−0.418	−0.399	−0.429 (h)	−0.385 (h)
ε($3\sigma_g$)	–	−0.476 (h)	−0.504 (h)	−0.631 (h)	−0.384	−0.372
ε($1\pi_g$)	–	−0.185	−0.247	−0.267	−0.296	−0.326
$n(1\pi_u)$	–	4.0	4.0	4.0	2.0	2.0
$n(3\sigma_g)$	–	0.0	0.0	0.0	2.0	2.0
$\Delta\rho$	–	0.008	0.017	0.058	0.020	0.018

lution with the same configuration lead to a single-determinant solution with the LUMO $\psi(3\sigma_g)$ having lower energy (Table 4) than the HOMO $\psi(1\pi_u)$ (a hole below the Fermi level). The $2\sigma_u$ drops significantly and will be fully occupied without doubt. Similarly, an attempt to obtain the KS solution at $R(\text{C–C}) = 3.2$ and 4.0 a.u. with the occupation pattern $2\sigma_g^2 2\sigma_u^2 1\pi_u^2 3\sigma_g^2$ also produces a hole below the Fermi level. This time the hole is present in the partially occupied orbitals $\psi(1\pi_{ux})$ and $\psi(1\pi_{uy})$, which have lower energy than the highest doubly occupied $\psi(3\sigma_g)$.

The procedure of "evaporation of a hole below the Fermi level" applied at $R > 1.8$ a.u. produces a KS solution with accidental degeneracy of the orbital $\psi(3\sigma_g)$ with $\psi(1\pi_{ux})$ and $\psi(1\pi_{uy})$, so that all three orbitals have the same energy (Table 4). This solution corresponds to a three-determinantal ensemble with the density matrix $\hat{M}_s$

$$\begin{aligned} \hat{M}_s &= d_0|D_{s0}\rangle\langle D_{s0}| + d_1\left(\tfrac{1}{2}|D_{s1}\rangle\langle D_{s1}| + \tfrac{1}{2}|D_{s2}\rangle\langle D_{s2}|\right) \\ d_0 &= \frac{n(1\pi_u) - n(3\sigma_g)}{4}, \quad d_1 = \frac{n(3\sigma_g)}{2} \end{aligned} \tag{31}$$

where D_{s0} is the KS determinant $|2\sigma_g^2 2\sigma_u^2 1\pi_u^4|$, D_{s1} is the determinant $|2\sigma_g^2 2\sigma_u^2 1\pi_{ux}^2 3\sigma_g^2|$ and D_{s2} is the determinant $|2\sigma_g^2 2\sigma_u^2 1\pi_{uy}^2 3\sigma_g|$. The density matrix $\hat{M}_s$ provides the E-V_s representability of the correlated ρ with fractional occupations $n(1\pi_u)$ and $n(3\sigma_g)$ of the orbitals $\psi(1\pi_u)$ and $\psi(3\sigma_g)$, respectively.

The contributions of the configurations D_{s0} and D_{s1}, D_{s2} change dramatically with bond distance following the trend for the CI expansion of Ψ discussed above. At the equilibrium distance R_e a KS ensemble is obtained with a "weak" accidental degeneracy in the sense that a low occupation $n(3\sigma_g) = 0.088$ of $\Psi(3\sigma_g)$ is enough to make this orbital degenerate with $\Psi(1\pi_u)$, as expected from the small difference between the orbital energies in the single-determinantal solution ($\epsilon_{3\sigma_g} = -0.476$, $\epsilon_{1\pi_u} = -0.450$). As a result, the KS solution in this case is close to the pure-state D_{s0} with low weights $n(3\sigma_g)/4 = 0.022$ of the determinants D_{s1} and D_{s2} in the density matrix (32).

At a longer distance $R(\text{C–C}) = 2.8$ a.u. a "medium" ensemble is obtained, while at $R(\text{C–C}) = 3.0$ a.u. a "strong" ensemble is formed with $n(3\sigma_g) = 0.981$ and the weights $n(3\sigma_g)/4 = 0.245$ and $\{n(1\pi_u) - n(3\sigma_g)\}/4 = 0.510$ of the determinants D_{s1}, D_{s2} and D_{s0}, respectively. The formation of this KS ensemble with a strong accidental degeneracy reflects the abovementioned strong multideterminantal character of the function Ψ of the interacting system at this distance. At $R(\text{C–C}) = 3.2$ a.u. the mixture of the determinants in the ensemble is reduced, since the weight of D_{s0} reduces to $\{n(1\pi_u) - n(3\sigma_g)\}/4 = 0.100$, while those of D_{s1} and D_{s2} increase to $n(3\sigma_g)/4 = 0.450$, and the ensemble (31) becomes rather close to a (by symmetry equal) mixture of the determinants D_{s1} and D_{s2}. At $R(\text{C–C}) = 4.0$ a.u. the accidental degeneracy of the ensemble can be again called "weak", the weight $\{n(1\pi_u) - n(3\sigma_g)\}/4 = 0.025$ of the configuration D_{s0} in (32) becomes really small.

From Table 4 one can compare the characteristics of the ensemble KS solution with those of the KS-like solution with a hole below the Fermi level. In all cases the density error $\Delta\rho$ of the ensemble solution is smaller than that for the KS-like solution, though the corresponding difference is not large for weak and medium ensembles. However, the one-determinantal solution clearly fails to reproduce the ρ corresponding to the function Ψ with a strong multideterminantal character at $R(\text{C–C}) = 3.0$ a.u., the corresponding $\Delta\rho$ amounts to 0.058 elections more than 7 times as large as $\Delta\rho$ of the ensemble solution. In this case the hole in the $3\sigma_g$ is more than 0.2 a.u. below the Fermi level. The inferior quality of the KS-hole solution follows also from a comparison of the kinetic energies T_s, which will be made below.

At R_e and $R(\text{C–C}) = 2.8$ and 3.0 a.u. the occupation of the $3\sigma_g$ is increased upon forming the ensemble solution and the energies of all σ orbitals for the ensemble are shifted upwards compared to those of the solution with a hole (with the only exception of the orbital $2\sigma_g$ at R_e), while the energies of all π-orbitals are shifted downwards. At $R(\text{C–C}) = 3.2$ and 4.0 a.u. the opposite trend takes place, in agreement with the fact that at these distances the occupation of $1\pi_u$ is increased when going to the ensemble solution. These trends are also reflected in the form of the exchange-correlation potentials v_{xc} for both types of solution, which will be discussed later in this section.

Table 5 displays the populations b_{ij}^{KS} (Eq. 23) of the products $\varphi_i\varphi_j$ of HF MOs in the KS representation of ρ, which can be compared with the populations b_{ij}^{NO} in the NO representation from Table 3. The differences illustrate the well-known fact that the KS solution does not provide the full one-electron density matrix but just the diagonal density. At $R(\text{C–C}) = 1.8$ a.u. the comparison reveals a distinct difference between the KS orbitals of the pure determinantal state and the NOs, analogous to that established for the stretched H_2 molecule in Sect. 3 and the CH_2 molecule in the previous section. The population $b_{3\sigma_g 3\sigma_g}^{KS} = 0.016$ of the σ bonding HF orbital $\varphi(3\sigma_g)$ in the KS representation is very small, while $\varphi(3\sigma_g)$ is relatively highly populated in the NO representation, $b_{3\sigma_g 3\sigma_g}^{NO} = 0.560$. On the other hand, the population $b_{2\sigma_u 2\sigma_u}^{KS} = 1.994$ of the σ antibonding HF orbital $\varphi(2\sigma_u)$ is higher than its NO counterpart $b_{2\sigma_u 2\sigma_u}^{NO} = 1.415$. The NO coefficients reflect an appreciable population of the HF MO $\varphi(3\sigma_g)$ and depopulation of $\varphi(2\sigma_u)$, in agreement with considerable admixture of the excited configuration $2\sigma_u^2 \to 3\sigma_g^2$ in the CI. The NOs $\chi(3\sigma_g)$ and $\chi(2\sigma_u)$ are very similar to the corresponding HF MOs and their occupation numbers therefore are close to the diagonal elements $b_{3\sigma_g 3\sigma_g}^{NO}$ and $b_{2\sigma_u 2\sigma_u}^{NO}$. The KS coefficients reflect the fact that the KS pure-state determinant with configuration $2\sigma_g^2 2\sigma_u^2 1\pi_u^4$ lacks the $3\sigma_g$ orbital. The effect on the diagonal density of admixture of $2\sigma_u^2 \to 3\sigma_g^2$ in the CI wavefunction has to be built in by changing the orbitals (particularly $2\sigma_g$) away from the HF MO character, which can only be effected by mixing with unoccupied HF MOs of the same symmetry. The KS orbital $\psi(2\sigma_g)$, while still being predominantly the HF $\varphi(2\sigma_g)$, has some contribution, with minus sign, of the HF MO $\varphi(3\sigma_g)$. In analogy with the H_2 and CH_2 cases, the effect of excitations such as $2\sigma_u^2 \to 3\sigma_g^2$ and $1\pi_u^4 \to 1\pi_u^2 3\sigma_g^2$ is reproduced in the KS representation by enhanced populations of the cross-products $\varphi_i\varphi_j$ of

Table 5. Populations $b^{\mathrm{KS}}_{\mathrm{ij}}$ of the products of the Hartree-Fock MOs φ_{i} and φ_{j} in the Kohn-Sham representations of the correlated density for C_2. The $1\pi_u\,1\pi_u$ contributions are summed over the *x* and *y* components

R(C–C), a.u.	1.8	2.348	2.8	3.0	3.2	4.0
Pure state (at 1.8 a.u.) and ensemble (other *R*) Kohn-Sham solutions						
				$b_{ij}^{\mathrm{KS-ensemble}}$		
$2\sigma_g2\sigma_g$	1.979	1.987	1.989	1.992	1.999	1.999
$2\sigma_u2\sigma_u$	1.994	1.996	1.998	1.998	1.996	1.994
$1\pi_u1\pi_u$	3.999	3.910	3.648	3.014	2.192	2.034
$3\sigma_g3\sigma_g$	0.016	0.098	0.353	0.976	1.799	1.948
$1\pi_g1\pi_g$	0.00	0.00	0.00	0.00	0.00	0.00
$2\sigma_g3\sigma_g$	–0.359	–0.300	–0.253	–0.169	0.00	0.00
Kohn-Sham solutions with a hole below the Fermi level						
				$b_{ij}^{\mathrm{KS-hole}}$		
$2\sigma_g2\sigma_g$	–	1.985	1.982	1.955	1.998	1.998
$2\sigma_u2\sigma_u$	–	1.993	1.981	1.904	1.987	1.992
$1\pi_u1\pi_u$	–	3.996	3.990	3.972	1.992	1.984
$3\sigma_g3\sigma_g$	–	0.012	0.014	0.026	1.999	1.998
$1\pi_g1\pi_g$	–	0.00	0.00	0.00	0.00	0.00
$2\sigma_g3\sigma_g$	–	–0.315	–0.338	–0.448	0.00	0.00

the HF MOs. Indeed, the population of the product $\varphi(2\sigma_g)\varphi(3\sigma_g)$ in the KS representation $b^{\mathrm{KS}}_{2\sigma_g3\sigma_g} = -0.359$ is much larger than $b^{\mathrm{NO}}_{2\sigma_g3\sigma_g} = 0.152$ in the NO representation. The diagonal b^{KS}_{ij} in Table 5 shows that the coefficients of the $\varphi(2\sigma_g)$ and $\varphi(3\sigma_g)$ in the KS $\psi(2\sigma_g)$ are 0.995 and −0.089 respectively, leading to the significant off-diagonal $b^{\mathrm{KS}}_{2\sigma_g3\sigma_g} = -0.359$.

Similar observations can be made with respect to the comparison of the NO representation at R_e and $R(C–C) = 2.8$ and 3.0 a.u. with that for the KS-like single determinant solution with a hole below the Fermi level (Table 5). It is interesting to note that in going from 3.0 to 3.2 bohr the hole-solution changes configuration from a hole in $3\sigma_g$ to a hole in $1\pi_u$ (cf. populations in Table 4), which is reflected in an abrupt change in the diagonal $b^{\mathrm{KS-hole}}_{3\sigma_g3\sigma_g}$ from 0.026 to 1.999, and in $b^{\mathrm{KS-hole}}_{1\pi_u1\pi_u}$ from 3.972 to 1.992. According to Table 4, the error $\Delta\rho$, although larger than for the ensemble solutions, is not excessive. Apparently the orbitals in these hole determinants are able to reorganize significantly so as to continue to describe the target CI density reasonably well, almost irrespective of the configuration (for instance, either the KS $3\sigma_g$ fully occupied or empty). This must have important consequences for shape of the orbitals and of the KS potentials as well. The sudden change $b^{\mathrm{KS-hole}}_{2\sigma_g3\sigma_g}$ from −0.448 at $R = 3.0$ a.u. to 0.00 at $R = 3.2$ a.u. is testimony to this change (of the $2\sigma_g$ orbital in this case).

It is interesting to note that at these larger distances, judging from the coefficients b^{KS}_{ij} and b^{NO}_{ij}, the KS orbitals of the ensemble solution are closer to the NOs than the orbitals of the one-determinantal hole solution, which is especially so for the "medium" and "strong" ensembles at R(C–C) = 2.8 and 3.0 a.u. For the ensemble at 3.0 bohr the population of the HF MO $\varphi(3\sigma_g)$, $b^{\mathrm{KS-ensemble}}_{3\sigma_g3\sigma_g} = 0.976$, is much larger than that for the KS-hole solution, $b^{\mathrm{KS-hole}}_{3\sigma_g3\sigma_g} = 0.026$. The latter has to be built up fully by mixing of the HF $\varphi(3\sigma_g)$ into the KS $2\sigma_g$ whereas in the ensemble the KS $3\sigma_g$ consisting predominantly of $\varphi(3\sigma_g)$, already has an occupation number $n(3\sigma_g) = 0.981$. The population of the HF $\varphi(3\sigma_g)$ in the ensemble solution is in fact rather close to that for the NO representation ($b^{\mathrm{NO}}_{3\sigma_g3\sigma_g} = 0.962$, Table 3). Similarly, the populations $b^{\mathrm{KS-ensemble}}_{1\pi_u1\pi_u} = 3.014$ and $b^{\mathrm{KS-ensemble}}_{2\sigma_g2\sigma_g} = -0.169$ for the ensemble are much closer to the $b^{\mathrm{NO}}_{1\pi_u1\pi_{ui}} = 2.834$ and $b^{\mathrm{NO}}_{2\sigma_g3\sigma_g} = -0.130$ for the NO representation as compared to $b^{\mathrm{KS-hole}}_{1\pi_u1\pi_u} = 3.972$ and $b^{\mathrm{KS-hole}}_{2\sigma_g3\sigma_g} = -0.448$ for the KS-hole solution. The ensemble configuration of the KS orbitals $2\sigma_g^2 2\sigma_u^2 1\pi_u^{3.019} 3\sigma_g^{0.981}$ (occupation numbers from Table 4) contains the fractionally occupied orbital $\psi(3\sigma_g)$, which has a dominant contribution of the corresponding HF MO $\varphi(3\sigma_g)$, and the depopulated orbital $\psi(1\pi_u)$. This configuration is close to the configuration of the NOs $2\sigma_g^{1.966} 2\sigma_u^{1.898} 1\pi_u^{2.826} 3\sigma_g^{0.970}$ (NO occupation numbers used), which represents the correlated ρ. Thus, one can consider the occupied KS orbitals of the ensemble as plausible approximations for the corresponding NOs. They will of course not be exactly equal.

In Table 6 the energy characteristics obtained for the KS pure state or ensemble and for the KS-like solutions with a hole below the Fermi level at various distances R(C–C) are compared with the corresponding HF characteristics. Considering first the ensemble solutions we note that the kinetic energies T_s of the KS ensemble solution are consistently higher than their HF counterparts T^{HF}, with the corresponding difference being increased with increasing R(C–C). As was established previously for other molecules [24, 25], this is due to the contraction of the correlated ρ around the nuclei as compared with the HF density ρ^{HF}, and the increasing non-dynamical correlation at larger bond distances, which is neglected in the HF approximation. On the other hand, the ensemble exchange energies E_x are close to the HF ones E_x^{HF} for the ensembles with a weak accidental degeneracy, and E_x are somewhat larger (in absolute magnitude) than E_x^{HF} for medium and strong ensembles. When summing up, differences between individual KS and HF energy components tend to compensate each other, so that for the distances R(C–C) up to 3.0 a.u. the KS expectation values E^{KS} are rather close to the HF energies E^{HF} (Table 6). For R(C–C) = 3.2 and 4.0 a.u. the ensemble E^{KS} becomes definitely higher than E^{HF} of the configuration $2\sigma_g^2 2\sigma_u^2 1\pi_u^2 3\sigma_g^2$.

Table 6. Kohn-Sham and Hartree-Fock energy characteristics (a.u.) for C_2

R(C–C), a.u.	1.8	2.348	2.8	3.0	3.2	4.0
Pure state (at 1.8 a.u.) and ensemble (other R) Kohn-Sham solutions						
T_s	77.287	75.517	75.068	75.088	75.101	75.029
E_x	−10.381	−10.047	−9.913	−9.922	−9.946	−9.809
E^{KS}	−95.148	−90.720	−88.206	−87.347	−86.611	−84.233
KS solutions with a hole below the Fermi level						
T_s	–	75.518	75.083	75.210	75.105	75.030
E_x	–	−10.032	−9.861	−9.788	−9.945	−9.808
E^{KS}	–	−90.713	−88.167	−87.153	−86.614	−84.233
Hartree-Fock						
T^{HF}	77.191	75.392	74.854	74.756	74.885	74.688
E_x^{HF}	−10.389	−10.041	−9.876	−9.827	−9.760	−9.810
E^{HF}	−95.179	−90.738	−88.210	−87.316	−86.686	−84.300

Comparing next to the KS solutions with a hole below the Fermi level, we note that the T_s values for the ensemble KS solution are consistently lower. This difference is marginal for weak ensembles at R_e and $R(\mathrm{C{-}C}) = 4.0$ a.u. as well as at $R(\mathrm{C{-}C}) = 3.2$ a.u., however, it becomes appreciable for medium and strong ensembles at $R(\mathrm{C{-}C}) = 2.8$ and 3.0 a.u. (Table 6). Since a true KS solution must have the minimal (for a certain ρ) kinetic energy, one can conclude that the ensemble represents such a true KS solution. The one-determinantal solution with a hole is of inferior quality, which is especially clear for the case when the wave function Ψ of the interacting system has a strong multideterminantal character. For the hole solutions the E^{KS} values are clearly inferior at distances up to 3.0 bohr.

Figure 3 presents the KS exchange-correlation potential v_{xc} constructed at $R(\mathrm{C{-}C}) = R_e$ 2.8, 3.0 and 4.0 a.u. The potentials are plotted along the C–C bond axis as functions of the distance z from the bond midpoint, so that half of symmetrical picture for the C_2 dimer is presented. The ensemble v_{xc} displays all the typical features of the exchange-correlation potentials for a dimer molecule A_2, which have been established and interpreted previously for the dimers Li_2, N_2, F_2 [44]. These are a deep well around the nucleus, atomic intershell peaks, Coulombic asymptotics at larger z and a plateau in the bonding region (small z). In particular, a plateau reflects the form in the σ-bonding region of the exchange (Fermi) hole, since the potential of this hole brings a dominant contribution to v_{xc}. In this region the Fermi hole is delocalized symmetrically over both C atoms and is essentially static, i.e., it does not change shape when the reference position is changed around the bond midpoint [44, 45]. Increase of R(C–C) produces a change of the form of v_{xc} in the σ-bonding region around the bond midpoint from a plateau to a more Coulombic-like shape.

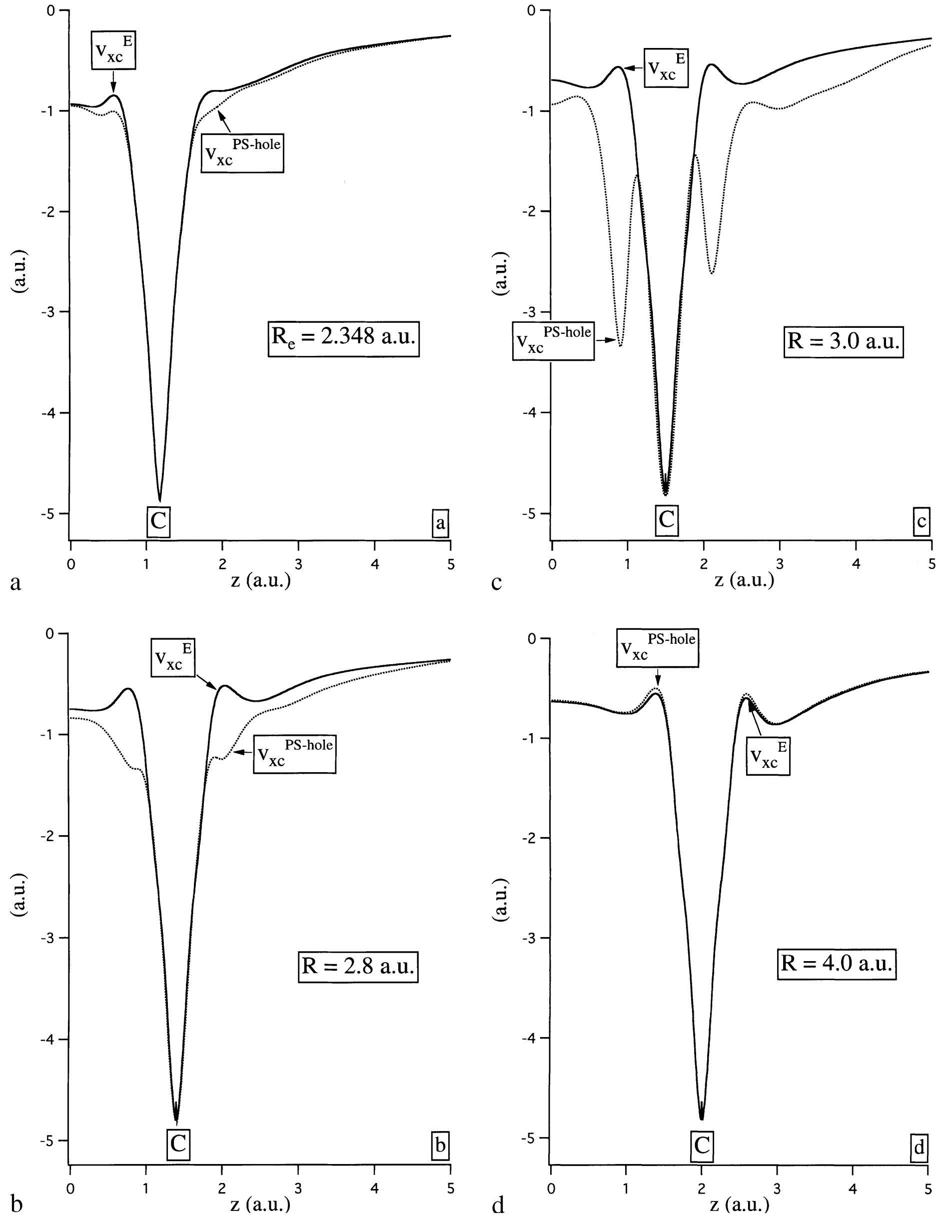
While v_{xc} at all R(C–C) considered looks"normal" for the ensemble, the potential for the KS-like solution with a hole is heavily distorted at $R(\mathrm{C{-}C}) = 2.8$ and 3.0 a.u. (Fig. 3*b, c*). This appears to be the price for an attempt to reproduce with just one determinant the density from a CI wave function Ψ that has a strong multideterminantal character. We already noted that the occupied KS orbitals in this case have to change shape considerably in order to incorporate the character of HF MOs that are occupied in the CI and in the ensemble solution. This of course requires substantial changes in the potential, which have to be effected through v_{xc}. If Ψ consists of a relatively weak mixture of determinants, at R_e and $R(\mathrm{C{-}C}) = 4.0$ a.u., the potentials for both types of solution are rather close to each other, though at R_e the potential for the KS-hole solution does not possess the characteristic atomic intershell peaks. For R_e and $R(\mathrm{C{-}C}) = 2.8$ and 3.0 a.u. the potential for the KS-hole solution is consistently more attractive than that for the ensemble solution. This correlates with the downward shift of energies of the σ-orbitals of the former solution as compared with those of the latter solutions. Taken together with the abovementioned relatively large density error $\Delta\rho$ and high kinetic energy T_s these results confirm the inferior quality of a one-determinantal KS-like solution with a hole and confirm that the ensemble is the proper KS solution in this case.

The behavior of v_{xc} with changing R(C–C) in the π-bonding region is shown in Fig. 4. It represents contour plots of the ensemble v_{xc} at R_e and $R(\mathrm{C{-}C}) = 2.8$ and 4.0 a.u. in the plane containing the bond axis (axis z of the plot), with the bond midpoint being at the position (0,0). The variation of the shape of v_{xc} correlates with the gradual decrease of the π-bonding with increasing R(C–C). Indeed, at R_e and away from the z axis (at $x = 0.5-0.8$ a.u.) the potential has lower values (contours of −0.9 and −0.85 Hartrees) closer to the bond midplane, thus favoring an accumulation of the electron density in the region of "banana" bonds (Fig. 4a). At larger distances the contour of −0.85 Hartrees reduces to a small closed curve and moves in the nonbonding region on the outer side of the C nucleus.

6 Summary and conclusions

In this paper the type of Kohn-Sham solution and the form of the KS orbitals has been studied for cases (the CH_2 and C_2 molecules) where the nondegenerate pure-state wave function of the corresponding interacting system has essentially multideterminantal character. We find the CH_2 ground state at equilibrium geometry to be PS-V_s representable, but C_2 at R_e and longer bond

Fig. 3a–d The exchange-correlation potential v_{xc} along the bond axis of C_2. v_{xc}^{E} (solid lines) is the potential corresponding to the ensemble solution; $v_{xc}^{PS-hole}$ (*dotted lines*) corresponds to the single-determinant solutions with a hole below the Fermi level. **a** R_e, **b** 2.8 a.u., **c** 3.0 a.u., **d** 4.0 a.u.



distances to be E-V_s representable, (or class *b*, as described in Sect. 1).

Applying the results of Levy [1] and Lieb [2] to noninteracting systems, we know that there are many densities that are not PS-V_s representable but that are E-V_s representable. It is thus certainly possible that a given interacting ground state density does not belong to the set of PS noninteracting ground state densities, but does belong to the set of ensemble-representable noninteracting densities. Our results prove this not to be only an academic possibility, but E-V_s representability is in fact called for to handle cases with strong electron correlation, i.e., essentially multideterminantal character of the interacting ground-state wavefunction.

The KS solution for the lowest singlet states of CH_2 and C_2 has been constructed from the *ab* initio CI ρ. To obtain the ensemble solution in the cases when an attempt to construct the PS leads to a non-Aufbau solution, the procedure of "evaporation of the hole below the Fermi level" has been employed [15, 17, 37]. Already for CH_2, for which the wave function has a relatively weak multideterminantal character, calculation in the TZ basis yields the ensemble solution and only with the extended QZ basis the pure-state one-determinantal KS solution has been obtained.

With the example of C_2, the possibility has been demonstrated of an essentially accurate ensemble KS solution with accidental degeneracy (E-V_s representability) for a ρ that is PS-V representable for the interacting system. A variety of KS solutions has been obtained depending on the *R*(C–C) distance, ranging from a single-determinantal pure-state at *R*(C–C) = 1.8 a.u. through an ensemble with a weak accidental degeneracy at R_e to an ensemble with a strong degeneracy

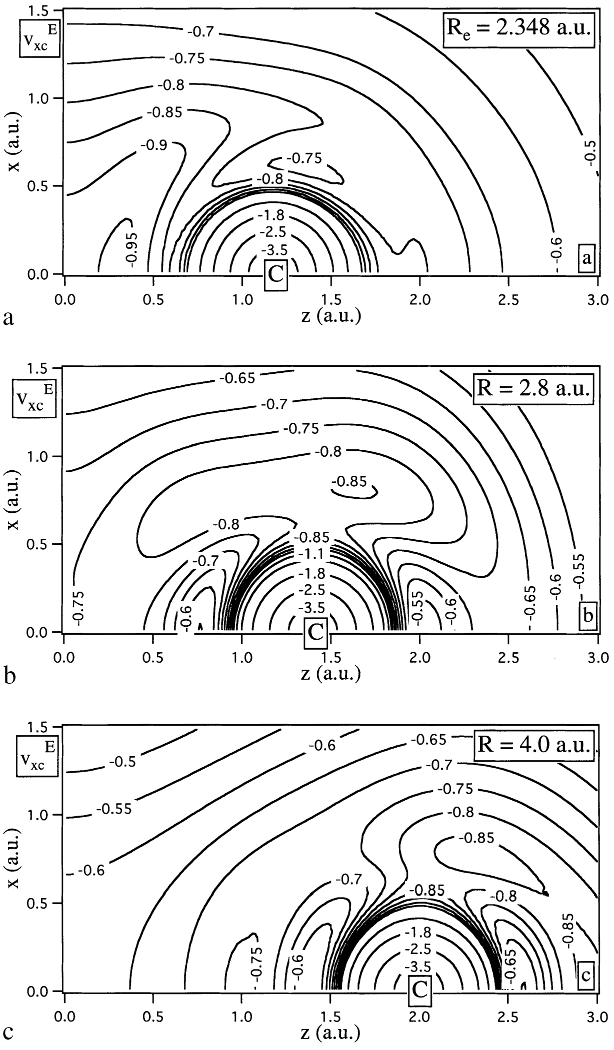


Fig. 4a-c Contour plots of the exchange-correlation potential v_{xc} in the plane containing the bond axis of C_2. The bond midpoint is at (0,0). **a** R_e, **b** 2.8a.u. **c** 4.0 a.u.

at $R(C–C) = 3.0$ a.u. and then to another ensemble with a weak accidental degeneracy at $R(C–C) = 4.0$ a.u. KS-like solutions with a hole below the Fermi level have also been constructed as alternatives to the ensemble solution. When the multideterminantal character of the wave function Ψ of the interacting system is weak, the density errors $\Delta\rho$ and energies such as T_s and E^{KS} of both solutions are in fact rather close to each other. However, when Ψ has a strong multideterminantal character, the quality of the solution with a hole is inferior to that of the ensemble since $\Delta\rho$ is considerably higher for the former solution. This is probably related to the unusual form that v_{xc} has to take in this case in order to "distort" the fully occupied KS orbitals sufficiently so that they can build the density. The potential obtained for the ensemble solution looks "normal", i.e., its form with characteristic atomic intershell peaks is reminiscent of that obtained previously for monohydrides XH (X = Li, B, F) [24, 46] and for dimers X_2 (X = Li, N, F) [44]. The relatively large $\Delta\rho$ and the strange shape of v_{xc} are, of course, not proof that the hole solutions are not correct, since we cannot exclude the possibility that the numerical procedure of constructing v_{xc} may be refined so as to achieve a better $\Delta\rho$ by an even more elaborate v_{xc}. The solutions with a hole below the Fermi level are, however, not acceptable since they are not true ground states of the noninteracting system with the obtained potential v_s. Taken together, these results confirm that the ensemble is the proper KS solution in this case.

The KS orbitals can be compared to the NOs. When the density is still PS-V_s representable, but configuration mixing is already becoming significant in the CI wave function, the occupied KS orbitals may start to differ from the NOs, since they have to incorporate the effect of the configuration mixing on the density. As an example we have discussed the stretched H_2 molecule. The expansion of the KS σ_g orbital in terms of the canonical Hartree-Fock orbitals has a specific form, since unoccupied HF orbitals of σ_g symmetry have to be included in order to represent the correlation effect on the density of electronic excitation to the σ_u orbital (of a different symmetry) in the CI expansion. Because of this specific expansion, the construction of accurate KS orbitals in a finite basis requires a sufficiently large basis set. The expansion indicated here causes the KS orbitals for a pure state to be distinctly different from the natural orbitals representing the same ρ. When being expanded in terms of the HF orbitals φ_i, the KS orbitals and NOs provide different representations of the same ρ in terms of the products $\varphi_i\varphi_j$. This has been observed in C_2 at 1.8 bohr, where the density can be represented by a noninteracting pure state, and in particular for the KS-like solutions that consists of a pure state (single determinant) with a hole below the Fermi level. The KS orbitals of the ensemble solutions are closer to the NOs.

The ensemble solutions obtained for C_2 provide a representation of ρ with fractional occupations of KS orbitals. Thus, our results support the use of fractional occupation of KS orbitals within the computational DFT schemes based on the LDA and GGA in the cases where the standard procedure creates holes below the Fermi level in the KS spectrum [15,17]. Still, care should be exercised in this case, since the form of the KS solution might be sensitive to the size of the basis and to specific DFT approximation used.

The elongation of the C–C bond can be considered as a simple example of a dissociation reaction. The complex electron rearrangement with a σ-bond forming and a π-bond breaking is described in the KS theory by the appearance and transformation of the ensemble solution with an accidental degeneracy presented above. One can expect that such ensembles play an important role in the KS description of chemical reactions and, especially, of their transition state [17]. This will be substantiated in our further work.

References

1. Levy M (1982) Phys Rev A 26: 1200
2. Lieb EH (1983) Int J Quantum Chem 24: 243

3. Kohn W, Sham LJ (1965) Phys Rev 140A: 1133
4. Lieb EH (1981) Phys Rev Lett 46: 457
5. Englisch H, Englisch R (1983) Phys Stat Sol 123: 711
6. Englisch H, Englisch R (1983) Phys Stat Sol 124: 373
7. van Leeuwen R, Baerends EJ (1994) Phys Rev 49: 2421
8. Schipper PRT, Gritsenko OV, Baerends EJ (1997) Theor Chem Acc 98: 16
9. Shavitt I (1985) Tetrahedron 41: 1531
10. Schaefer HF (1986) Science 231: 1100
11. Yamaguchi Y, Sherril CD, Schaefer HF (1996) J Mol Spectrosc 179: 263
12. Kraemer WP, Roos BO (1987) Chem Phys 118: 345
13. Dreizler RM, Gross EKU (1990) Density functional theory: an approach to the many-body problem; Springer, Berlin, Heidelberg, New York
14. Slater JC, Mann JB, Wilson TM, Wood JH (1969) Phys Rev 184: 672
15. Dunlap BI, Mei WN (1983) J. Chem Phys 78: 4997
16. Dunlap BI (1984) J Chem Phys 29: 2902
17. Wang SG, Schwarz WHE (1996) J Chem Phys 105: 4641
18. Baerends EJ, Ellis DE, Ros P (1973) Chem Phys 2: 41
19. te Velde G, Baerends EJ (1992) J Comput Phys 99: 84
20. McWeeny R (1989) Methods of molecular quantum mechanics Academic,: New York
21. Zhao Q, Morrison RC, Parr RG (1994) Phys Rev A 50: 2138
22. Morrison RC, Zhao Q (1995) Phys Rev A 51: 1980
23. Gritsenko OV, van Leeuwen R, Baerends EJ (1995) Phys Rev A 52: 1870
24. Gritsenko OV, van Leeuwen R, Baerends EJ (1996) J Chem Phys 104: 8535
25. Gritsenko OV, Schipper PRT, Baerends EJ (1997) J Chem Phys 107: 5007
26. Ingamells VE, Handy NC (1996) Chem Phys Lett 248: 373
27. Tozer DJ, Ingamells VE, Handy NC (1996) J Chem Phys 105: 9200
28. Levy M, Perdew JP (1985) In: Dreizler R M, da Providencia J (eds) Density functional methods in physics vol 123. Plenum, New York p 11
29. Hohenberg P, Kohn W (1964) Phys Rev 136B: 864
30. Chen J, Stott MJ (1991) Phys Rev A 44: 2809
31. Chen J, Stott MJ (1993) Phys Rev A 47: 153
32. Buijse MA, Baerends EJ, Snijders JG (1989) Phys Rev A 40: 4190
33. Gritsenko OV, Baerends EJ (1997) Theor Chem Acc 96: 44
34. Lie GC, Clementi E (1974) J Chem Phys 60: 1275
35. Poirier K, Kari P, Csizmadia IG (1985) Handbook of Gaussian basis sets. Elsevier, Amsterdam
36. Hoch DE, Harriman JE (1995) J Chem Phys 102: 9590
37. Averill FW, Painter GS (1992) Phys Rev B 46: 2498
38. Saunders VR, van Lenthe JH (1983) Mol Phys 48: 923
39. Dickson RM, Becke AD (1993) J Chem Phys 99: 3898
40. Buijse MA (1991) Electron correlation. Fermi and Coulomb holes, dynamical and nondynamical correlation. Thesis, Vrije Universiteit, Amsterdam
41. Woon DE, Dunning TH (1995) J Chem Phys 103: 4572
42. Grev RS, Schaefer HF (1992) J Chem Phys 97: 8389
43. Savin A, Stoll H, Preuss H (1986) Theo Chim Acta 70: 407
44. Schipper PRT, Gritsenko OV, Baerends EJ (1998) Phys Rev A 57: 1729
45. Buijse MA, Baerends EJ (1995) In: Ellis DE (ed) Electronic density functional theory of molecules, clusters and solids, Kluwer, Dordrecht, p 1
46. Gritsenko OV, Baerends EJ (1996) Phys Rev A 54: 1957