

# New insights from domain-averaged Fermi holes and bond order analysis into the bonding conundrum in C<sub>2</sub>

David L. Cooper,<sup>1,\*</sup> Robert Ponec,<sup>2</sup> and Miroslav Kohout<sup>3</sup>

<sup>1</sup> *Department of Chemistry, University of Liverpool, Liverpool L69 7ZD, United Kingdom*

<sup>2</sup> *Institute of Chemical Processes, The Czech Academy of Sciences, Rozvojová 135, 165 02 Prague 6, Czech Republic*

<sup>3</sup> *Max Planck Institute for Chemical Physics of Solids, Nöthnitzer Strasse 40, 01187 Dresden, Germany*

\* **Corresponding Authors:** E-mail: dlc@liv.ac.uk; ponec@icpf.cas.cz

**Abstract.** The bonding in the ground state of C<sub>2</sub> is examined using a combined approach based on the analysis of domain-averaged Fermi holes and of the contributions to covalent bond orders that can be associated with individual localised natural orbitals. The  $\sigma$  system in this molecule turns out to be particularly sensitive, evolving from a description that includes a fairly traditional shared electron pair  $\sigma$  bond, for a range of intermediate nuclear separations, to a somewhat different situation near equilibrium geometry, where non-classical repulsive interactions are particularly important. The various results provide further support for the view that the electronic structure of this molecule sufficiently exceeds the scope of traditional bonding paradigms that attempts to classify the bonding in terms of a classical bond multiplicity are highly questionable.

**Keywords:** Peculiarity of C<sub>2</sub> bonding; Domain-averaged Fermi holes (DAFH); Localized natural orbitals; Cioslowski covalent bond order (CBO); Wiberg-Mayer index; Quantum theory of atoms in molecules (QTAIM).

## 1. Introduction

The nature of the bonding in the ground state of  $C_2$  near its equilibrium geometry  $R_e$  remains controversial, with multiple claims and counterclaims [1-17]. Recent interest in this molecule has been driven by certain sets of valence bond (VB) calculations that could be interpreted in terms of three two-centre two-electron bonds ( $1 \times \sigma$  and  $2 \times \pi$ ) and a somewhat weaker ‘inverted’  $\sigma$  bond, based on the interaction of outwards pointing hybrids [6,7,9,14]. This VB description essentially identifies the system as possessing four bonds. Of course, there is at first sight nothing particularly remarkable about the value of  $R_e$  or the magnitude of the dissociation energy for the  $X^1\Sigma_g^+$  ground state of  $C_2$ , which dissociates smoothly to two ground state  $C(^3P)$  atoms. The VB study did of course include a calculation of an accurate estimate of the dissociation energy to ground state atoms but, in order to substantiate a high bond multiplicity, its authors have, amongst other things, invoked the significant  $2 \times C(^5S)$  character of  $C_2$  near  $R_e$  to argue that one should also look at the somewhat larger energy difference relative to two  $C(^5S)$  atoms. The various claims of enhanced bonding in the ground state of  $C_2$  have been challenged by Frenking and Hermann who, amongst other concerns, maintained that a more direct *in situ* measure of the strength of the bonding near  $R_e$  is provided by the value of the vibrational force constant [8]. Based on that criterion, no particularly strong bonding appears to be present in this system. The model favoured by Hermann and Frenking [17] is based instead on a  $\pi$  double bond augmented with weakly bonding donor-acceptor  $\sigma$  interactions. There are similarities here to the ‘ $\nu$  bonding’ model of Weinhold and Landis [18]. Xu and Dunning [13] have also argued against a strong quadruple bond in this system, highlighting the importance of an antiferromagnetic component to the bonding.

An important issue that has not been settled in previous studies is whether computed values of the total bond order for the ground state of  $C_2$  near  $R_e$  might arise from factors other than bonding. The main inspiration for the present work is that we were able to show previously [19] that a combination of domain-averaged Fermi hole analysis and an examination of orbital-resolved bond orders could clearly distinguish between  $N_2$ , for which a total  $\sigma$ -only bond order of *ca.* 1.0 really does correspond to a fairly ordinary two-centre two-electron  $\sigma$  bond, and  $Be_2$  which features a bond order of *ca.* 0.9 for a comparable nuclear separation, even though it is demonstrably unbound at this geometry. The obvious question that arises is whether the analogous  $\sigma$ -only bond order in  $C_2$  is primarily due to bonding, as in  $N_2$ , or whether the  $\sigma$  system is in fact more similar to that in  $Be_2$ . It is of course important to

bear in mind, when considering sufficiently short nuclear separations, that increases in various measures of bond order as two atoms are brought still closer together [20] need not correspond to increased bonding: not all electron sharing between atomic domains leads to stabilization. This is of course especially true of Pauli repulsion. The relatively high  $\sigma$  bond order in  $\text{Be}_2$ , which is not so much smaller than the corresponding  $\sigma$ -only values for  $\text{N}_2$  for these geometries, does in fact correspond to net destabilizing interactions. This observation appeared to be very clearly highlighted by the analysis that we performed [19] and so it seems timely to apply the same sort of methodology to the controversial case of  $\text{C}_2$ .

A key aspect that we aim to address in the present work is the extent to which the electronic structure of the  $X^1\Sigma_g^+$  ground state of  $\text{C}_2$  near  $R_e$  resembles one of two simple models. The first of these descriptions, which is essentially the one promoted by Shaik and coworkers [6,7,9,14], corresponds to ethyne,  $\text{HC}\equiv\text{CH}$ , from which both hydrogen atoms have notionally been removed but retaining the triple bond ( $1\times\sigma$  and  $2\times\pi$ ) plus two outwards pointing hybrids (which might or might not interact so as to generate a fourth bond). The second description of the  $\sigma$  system is in essence much the same as that in  $\text{Be}_2$  (for this geometry) but it is augmented with two  $\pi$  bonds. Given that the analysis we intend to deploy could clearly distinguish the situation in  $\text{Be}_2$  from that in  $\text{N}_2$  [19] (and, presumably, in  $\text{HC}\equiv\text{CH}$ ) there is every reason to suppose that it is very appropriate for the task at hand. After identifying the nature of the  $\sigma$  system near equilibrium geometry, it should then prove very interesting to use the same techniques to monitor the changes to the electronic structure that accompany the smooth dissociation of the  $X^1\Sigma_g^+$  ground state of  $\text{C}_2$  into two ground state atoms. Of special interest is the information that these modes of analysis can provide about the changes to the number and to the nature of the bonding electron pairs as the nuclear separation is increased. In particular, we can expect to observe individual bonds being formed/broken at different CC distances.

## 2. Theoretical and computational details

Domain-averaged Fermi hole (DAFH) analysis [21-31] has been shown to provide useful insights into the electronic structure of a wide range of molecules and solids. This includes systems with non-trivial bonding patterns [23,24,32-38], such as multicentre bonding, metal-metal bonding, hypervalency, and so on. In order to investigate the nature of the bonding in the  $X^1\Sigma_g^+$  ground state of  $\text{C}_2$ , such DAFH analysis has been augmented in the

present work with an examination of localised natural orbitals (LNOs) and of the relative importance of different contributions to certain measures of bond order. This is essentially the same combination of techniques as was successfully used previously to identify features that clearly distinguish the somewhat different  $\sigma$  systems in the ground states of  $\text{Be}_2$  and  $\text{N}_2$  [19].

In order to carry out the DAFH analysis, as well as to examine LNOs and bond orders, we do of course require appropriate wavefunctions for all of the systems in which we are interested. First of all, single-configuration restricted Hartree-Fock (RHF) descriptions were computed at  $R=1.2425 \text{ \AA}$  for the lowest  $^1\Sigma_g^+$  state of  $\text{C}_2$  ( $1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 2\sigma_u^2 1\pi_u^4$ ) and for the ‘excited’ configurations ( $^3\Sigma_u^+$ ,  $^1\Sigma_u^+$  and  $^1\Sigma_g^+$ ) that correspond to excitations from  $2\sigma_u$  to  $3\sigma_g$ . In practice, such single-configuration RHF descriptions were generated using appropriate complete active space self-consistent field (CASSCF) constructions with very limited active spaces. All of these RHF and CASSCF [39,40] calculations were carried out in  $D_{2h}$  symmetry using the general-purpose quantum chemistry program package MOLPRO [41,42] and made use of the standard cc-pVQZ basis set in spherical harmonic form. An additional RHF calculation at  $R=1.2425 \text{ \AA}$  was carried out for the lowest  $^1\Sigma_g^+$  state of  $\text{C}_2$  using the smaller 3-21G basis.

Various CASSCF calculations were then carried out for  $\text{C}_2$ ,  $\text{Be}_2$ ,  $\text{N}_2$  and  $\text{HC}\equiv\text{CH}$  using standard cc-pVQZ basis sets, adopting a fixed bond length of  $1.2425 \text{ \AA}$ . In the case of  $\text{HC}\equiv\text{CH}$ , we used  $R_{\text{CH}}=1.06 \text{ \AA}$ . Results from our various types of analysis were subsequently compared for different choices of CASSCF active space including: (a) the full-valence space and (b) only the  $\sigma$  part of the full-valence space. For the lowest  $^1\Sigma_g^+$ ,  $^3\Sigma_u^+$  and  $^1\Sigma_u^+$  states of  $\text{C}_2$  we also employed a larger CASSCF(8,16) construction in which, relative to full-valence CASSCF, the number of active orbitals of each symmetry (in  $D_{2h}$ ) was doubled. This somewhat more flexible description was subsequently used to monitor the evolution of the bonding in the  $X^1\Sigma_g^+$  ground state of  $\text{C}_2$  as the nuclear separation was decreased from  $3.25 \text{ \AA}$  towards  $R_e$ . So as to be certain of obtaining  $^1\Sigma_g^+$  solutions, instead of switching to the  $\text{B}^1\Delta_g$  state (which corresponds to the same irreducible representation in  $D_{2h}$ ), we turned on a useful feature in MOLPRO [41,42] that selects states with a specific value of  $\Lambda$ .

We have chosen in the present work to use the quantum theory of atoms in molecules (QTAIM) [43] to define atomic domains,  $\Omega$ , as non-overlapping spatial regions of each molecular system. With the total electron densities expanded in terms of (real) orthonormal natural orbitals  $\phi_\mu$  with occupation numbers  $\omega_\mu$ , we also require for our subsequent analysis

the numerical values of the so-called domain-condensed overlap integrals,  $\langle \phi_\mu | \phi_\nu \rangle_\Omega$ , which take the following form:

$$\langle \phi_\mu | \phi_\nu \rangle_\Omega = \int_\Omega \phi_\mu(\mathbf{r}) \phi_\nu(\mathbf{r}) d\mathbf{r} \quad (1)$$

A particularly useful measure of the bond order between atoms  $A$  and  $B$  in a correlated singlet system is provided by the Wiberg-Mayer index,  $w$ , which may be defined in the following manner [44]:

$$w = \sum_{i \in A} \sum_{j \in B} [(\mathbf{DS})_{ij}(\mathbf{DS})_{ji} + (\mathbf{RS})_{ij}(\mathbf{RS})_{ji}] \quad (2)$$

in which  $\mathbf{D}$  and  $\mathbf{S}$  are the total one-electron density matrix and the overlap matrix, respectively. The matrix  $\mathbf{R}$ , which arises because of the difference between  $2(\mathbf{DS})$  and  $(\mathbf{DS})^2$ , is defined in Ref. 44 for any correlated system, not just singlet states. The notation  $i \in A$  signifies here that the particular summation in Equation (2) is restricted to atom-centred basis functions, labelled by  $i$ , that are associated with atom  $A$ . (We refer to this type of restriction on a summation as a Mulliken-like scheme [45,46].)

Straightforward manipulations of Equation (2), similar to those used by Ángyán *et al.* [47], lead to the following expression for the QTAIM-generalised Wiberg-Mayer index  $W$  between QTAIM domains  $A$  and  $B$  in a correlated singlet system:

$$W = \sum_\mu \sum_\nu \left[ \left( \omega_\mu \omega_\nu + \{ \omega_\mu (2 - \omega_\mu) \omega_\nu (2 - \omega_\nu) \}^{1/2} \right) \langle \phi_\mu | \phi_\nu \rangle_A \langle \phi_\mu | \phi_\nu \rangle_B \right] \quad (3)$$

We refer here to the resulting numerical values of  $W$  as the total W-M index. Given that there are no cross terms in Equation (3) between natural orbitals of  $\sigma$  and  $\pi$  symmetry, we may decompose this total W-M index into its separate W-M( $\sigma$ ) and W-M( $\pi$ ) components.

We also make substantial use here of the definition of the Cioslowski covalent bond order (CBO) [48]. First of all, application to the canonical natural orbitals  $\phi_\mu$  and their occupation numbers  $\omega_\mu$  of an implementation of Cioslowski's isopycnic transformation

scheme [49] leads to a set of LNOs  $\psi_p$  with occupation numbers  $\lambda_p$ . The Cioslowski covalent bond order  $C$  between QTAIM domains  $A$  and  $B$  can then be defined according to [48]:

$$C = \sum_p (\lambda_p)^2 \langle \psi_p | \psi_p \rangle_A \langle \psi_p | \psi_p \rangle_B \quad (4)$$

We refer here to the resulting numerical values of  $C$  as the total CBO value. Given that Equation (4) involves a single summation over the LNOs, not only can we decompose the total CBO value into its separate  $CBO(\sigma)$  and  $CBO(\pi)$  components, but we can also quantify the relative importance of the terms involving particular LNOs. As was shown in our previous work [19], and will be shown again here, the magnitudes of the relative contributions from different LNOs to the overall  $CBO(\sigma)$  and  $CBO(\pi)$  values turn out to be especially useful diagnostics of the molecular electronic structure. In addition to monitoring changes to visual depictions of the LNOs  $\psi_p$  and their occupations  $\lambda_p$ , we can also examine the localization number, which measures the effective number of domains that each of them spans [49]. This quantity, which we denote  $\tilde{n}_p$ , is defined according to [49]:

$$\frac{1}{\tilde{n}_p} = \sum_A \langle \psi_p | \psi_p \rangle_A \langle \psi_p | \psi_p \rangle_A \quad (5)$$

Alongside the inspection of LNOs and of relative contributions to bond orders, we have also performed DAFH analysis. Given that detailed accounts of the theoretical background and computational methodology have been presented before, we restrict ourselves here to some key features. The so-called ‘hole’  $g_\Omega(\mathbf{r}_1)$  for a domain  $\Omega$  may be defined as follows:

$$\begin{aligned} g_\Omega(\mathbf{r}_1) &= \rho^{(1)}(\mathbf{r}_1) \int_\Omega \rho^{(1)}(\mathbf{r}_2) d\mathbf{r}_2 - 2 \int_\Omega \rho^{(2)}(\mathbf{r}_1; \mathbf{r}_2) d\mathbf{r}_2 \\ &\equiv \sum_p n_{p\Omega} |\varphi_{p\Omega}(\mathbf{r}_1)|^2 \end{aligned} \quad (6)$$

in which  $\rho^{(1)}$  and  $\rho^{(2)}$  are the (spinless) one-electron and two-electron densities. For each domain in turn, we calculated a matrix representation of  $g_\Omega(\mathbf{r}_1)$  in the (orthonormal) natural orbital basis using the one- and two-electron density matrices, expressed in the same basis,

and the various domain-condensed overlap integrals. In each case, the eigenvectors and eigenvalues of this matrix representation of the ‘hole’ were then localised by means of an isopycnic transformation, thereby converting them into a set of (nonorthogonal) DAFH functions  $\varphi_{p\Omega}(\mathbf{r}_1)$  with occupations  $n_{p\Omega}$ .

It has been found that the most useful information tends to emerge from DAFH analysis when the domains are chosen according to some sort of physically-sound partitioning of the total electron density, such as QTAIM (as was used here), self-consistent Hirshfeld approaches [28] or even a simple Mulliken-like scheme. Pictorial representations of the resulting DAFH functions then often allow a straightforward association with such familiar concepts as chemical bonds, lone pairs, and so on. They also provide information about how, and to what extent, the electrons in a given domain are involved in interactions with the other domains in the molecule. Fortunately it turns out that the forms of such DAFH functions tend to be rather insensitive to the particular choice of partitioning scheme for the total electron density [28]. The same is true for the sums of complementary occupation numbers. On the other hand, individual values of  $n_{p\Omega}$  can be much more sensitive [28] and so it proves more appropriate to examine the coarse features of the occupation numbers than to focus on fine details.

It is our experience, for sensible choices of active space, that the results of the DAFH analysis for the active electrons of a given system are little changed whether or not the inactive space is also included in Equation (6). Accordingly, we have chosen throughout the present work to restrict our attention to the valence electrons of the various RHF wavefunctions and to the active spaces of the various post-RHF CASSCF constructions. (The QTAIM analysis, and thus the definition of the atomic domains, was of course always based on the total electron density.)

DAFH analysis can also be carried out for merged domains that are formed by combining multiple atomic domains. As well as identifying the electron pairs (chemical bonds, lone pairs, and so on) that are retained within a particular merged domain, such analysis can provide insights into the interactions between the various (merged) domains in the molecule. In order to enable direct comparison between  $\text{HC}\equiv\text{CH}$  and the various diatomic systems, it proved useful to analyze this polyatomic molecule in terms of two CH fragments, with each CH domain formed as the union of the corresponding C and H QTAIM domains.

In the particular case that a merged domain is taken to be an entire molecule, the resulting DAFH functions  $\varphi_{p\Omega}$  (with occupation numbers  $n_{p\Omega}$ ) must coincide with the LNOs

$\psi_p$  (with occupation numbers  $\lambda_p$ ). Note that all of the isopycnic transformations (for the DAFH analysis and for the generation of LNOs) were carried out for the various diatomic molecules using the domain-condensed overlaps for individual atomic QTAIM domains, whereas for HC $\equiv$ CH we used instead the analogous domain-condensed overlaps for each of the CH domains. The QTAIM analysis, including the computation of all of the domain-condensed overlaps, was carried out using the AIMall program [50]. Our own codes were used for all of the DAFH, LNO, W-M and CBO analysis for the various valence or active spaces. Pictorial representations of DAFH functions and of LNOs were produced from Virtual Reality Markup Language (VRML) files that were generated with MOLDEN [51].

### 3. Results and discussion

#### 3.1. Single-configuration descriptions of $C_2$ at $R_e$

Even though it is well known that the restricted Hartree-Fock (RHF) description of the  $X^1\Sigma_g^+$  ground state of  $C_2$  is rather inadequate, it is useful to consider briefly certain single-configuration descriptions of  $C_2$  before moving on to more realistic multiconfiguration wavefunctions. In the RHF configuration,  $1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 2\sigma_u^2 1\pi_u^4$ , the  $2\sigma_g$  and  $2\sigma_u$  bonding and antibonding orbitals, respectively, are based on different hybrids so that they do not completely cancel one another (*cf.* the  $X^1\Sigma_g^+$  ground state of  $Be_2$  [52]). Instead of just a  $\pi$  double bond, as was supposed by Mulliken [1], we can expect there to be a small residual bonding component of  $\sigma$  symmetry. The electronic structure of this system can thus be likened (adapting a description in Ref. 7) to a ‘sandwich with a meagre filling’, *i.e.* the electronic structure is dominated by the  $\pi$  double bond (the bread) but with a small  $\sigma$  bonding component (the filling). The corresponding RHF configuration for  $Be_2$  is of course the same, except for the omission of the filled  $1\pi_u$  molecular orbital.

Our intended analysis of the RHF wavefunction for  $C_2$  ( $X^1\Sigma_g^+$ ) at  $R_e$  was unfortunately made more complicated by the presence of a QTAIM non-nuclear maximum (NNM) at the midpoint. Such NNMs are not present for larger  $R$  and do not appear, for any of the geometries we considered, in the subsequent calculations that took account of electron correlation. With this in mind, we have, on this occasion only, used instead the smaller 3-21G basis, for which there is no NNM. We checked that the results of the RHF-based DAFH analysis for the 3-21G and cc-pVQZ basis sets were rather similar when using instead a Mulliken-like approximation to the domain-condensed overlap integrals. Additionally, we



found that DAFH analysis of the RHF/3-21G wavefunction produced similar results when using the Mulliken-like and QTAIM-based approaches.

The resulting DAFH functions for one of the carbon atom domains are shown in the left-hand column of Figure 1, together with the corresponding occupation numbers. The corresponding functions for the other carbon atom domain (not shown) are of course the mirror images of these. We observe that the singly-occupied functions of  $\pi$  symmetry are consistent with standard notions of a  $\pi$  double bond but that the situation is somewhat unclear in the case of the  $\sigma$  bonding system. As well as a  $\sigma$  function that is mostly localised on the specific carbon atom domain (occupation *ca.* 1.6) we also observe its mirror image (occupation *ca.* 0.4) for the other domain. Neither of these functions looks well adapted to classical shared electron pair  $\sigma$  bonding. Looking instead at the LNOs (right-hand column of Figure 1), the interpretation is more straightforward. The two doubly-occupied  $\sigma$  LNOs are identical to the corresponding DAFH functions; they do of course each account for 50% of the  $\text{CBO}(\sigma)$  value of *ca.* 1.2. Such a situation, in which a significant  $\text{CBO}(\sigma)$  value arises from LNOs that do not take the anticipated forms for bonding, arose in our previous work for  $\text{Be}_2$  [19] (see also later) for which the RHF description is of course the same as for  $\text{C}_2$ , except for the omission of the filled  $\pi$  orbitals. On the other hand, the situation was entirely different for  $\text{N}_2$  and, as will be shown later, the analysis of  $\text{HC}\equiv\text{CH}$  fairly closely resembles that of  $\text{N}_2$ . It thus appears that the RHF description of the  $\sigma$  system in  $\text{C}_2$  is much more alike to descriptions of  $\text{Be}_2$  (for the same geometry) than it is to those of  $\text{HC}\equiv\text{CH}$ .

«Figure 1 near here»

According to a simple analytic model (see Section S1 in the Supplemental data), the DAFH functions and the LNOs of  $\sigma$  symmetry correspond to normalised in-phase and out-of-phase combinations of the  $2\sigma_g$  and  $2\sigma_u$  RHF orbitals. The various occupation numbers and, indeed, the values of  $\text{W-M}(\sigma)$  and  $\text{CBO}(\sigma)$  are then given by simple expressions that involve only the value of a domain condensed overlap integral. Such a domain condensed overlap between  $2\sigma_g$  and  $2\sigma_u$  does of course represent a direct measure of how similar/different are the hybrids from which these two molecular orbitals are constructed. Additional numerical values for this  $\sigma$  system, as well as the various analytic expressions, are reported in Section S-A.2 in the Supplemental data. (The values of  $\text{W-M}(\pi)$  and  $\text{CBO}(\pi)$  for the RHF description of  $\text{C}_2$  are exactly 2.)

The key  $\sigma$  excitation in the subsequent correlated descriptions of the  $X^1\Sigma_g^+$  ground state of  $\text{C}_2$  is of course from  $2\sigma_u$  to  $3\sigma_g$ . Singly occupying each of these two orbitals corresponds to

an RHF description of the lowest  $^3\Sigma_u^+$  state (as well as a much higher lying  $^1\Sigma_u^+$  configuration). We return later to correlated descriptions of this triplet state ( $c^3\Sigma_u^+$ ) which, unlike the corresponding descriptions of  $X^1\Sigma_g^+$ , turn out to be reminiscent of  $\text{HC}\equiv\text{CH}$ . In order to generate instead an excited RHF configuration of  $^1\Sigma_g^+$  symmetry we must of course doubly excite from  $2\sigma_u$  to  $3\sigma_g$ . Unsurprisingly, the corresponding energy is higher than that of  $^1\Sigma_u^+$  – the RHF energy of this doubly-excited  $^1\Sigma_g^+$  configuration is nearly 244 millihartree above that of the lowest closed-shell RHF solution (cc-pVQZ basis set). Given that there is no longer any occupancy of the  $2\sigma_u$  antibonding orbital to counter the doubly-occupied  $2\sigma_g$  bonding orbital, this  $1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 3\sigma_g^2 1\pi_u^4$  excited configuration should in principle correspond to greater  $\sigma$  bonding. There is clearly a sense in which the supposed enhanced multiplicity of the  $\sigma$  bonding in the ground state of  $\text{C}_2$  is directly linked to the relative importance of this excited configuration in the final wavefunction. (All of the RHF energies are available in Table S1 in the Supplemental data.)

### 3.2. CASSCF descriptions at $R=1.2425\text{\AA}$

Various studies have employed a full-valence CASSCF description ('8 electrons in 8 orbitals') in order to investigate the nature of the bonding in  $\text{C}_2$  near  $R_e$ . We initially chose to do the same but it is important to bear in mind there must be significant changes to the character of the wavefunction in the region of the avoided crossing between the two  $^1\Sigma_g^+$  states ( $X^1\Sigma_g^+$ ,  $B^1\Sigma_g^+$ ) which dissociate to ground state atoms [53]. When increasing the nuclear separation, so as to monitor the changes that accompany dissociation of the ground state, we found that the changes in the natural orbital occupation numbers  $\omega_\mu$  were unrealistically sharp, over a relatively short range of  $R$ . In order to ameliorate this situation, we chose to use instead a larger CASSCF(8,16) construction, doubling the number of active orbitals of each symmetry (in  $D_{2h}$ ), and then found that the corresponding ground state  $\omega_\mu$  values changed somewhat more smoothly with increasing  $R$ . An alternative strategy could have been to perform appropriate state-averaged CASSCF calculations for the  $X^1\Sigma_g^+$  and  $B^1\Sigma_g^+$  states.

Given that the key excitation is in the  $\sigma$  space, we constructed also a somewhat more compact ' $\sigma$ -only' valence CASSCF(4,4) wavefunction at  $R_e$ . The occupation numbers of the  $2\sigma_g$ ,  $2\sigma_u$ ,  $3\sigma_g$  and  $3\sigma_u$  natural orbitals are 1.9938, 1.5773, 0.4232 and 0.0056, respectively. (We note that the occupation number of  $2\sigma_g$  is close to 2 and that of  $3\sigma_u$  is close to zero so that an even smaller CASSCF(2,2) wavefunction should give comparable results.) The dominant components of the CASSCF(4,4) wavefunction are of course the original

(...)2 $\sigma_g^2$ 2 $\sigma_u^2$ 1 $\pi_u^4$  configuration (78.6%) and the anticipated doubly-excited (...)2 $\sigma_g^2$ 3 $\sigma_g^2$ 1 $\pi_u^4$  configuration (21.0%). The corresponding 2 $\sigma_u$  and 3 $\sigma_g$  occupation numbers in the much more flexible CASSCF(8,16) construction are 1.6083 and 0.3876, respectively. The weight of the (...)2 $\sigma_g^2$ 3 $\sigma_g^2$ 1 $\pi_u^4$  configuration is now only 12.8% but the accumulated weight of all configuration in which there are single or double excitations from 2 $\sigma_u$  to 3 $\sigma_g$  (perhaps alongside other excitations) is *ca.* 20.4%.

We compare in Figure 2a the resulting DAFH functions  $\phi_{p\Omega}$  for one of the carbon atom domains when using the CASSCF(8,16), full-valence CASSCF and  $\sigma$ -only CASSCF wavefunctions. Also reported in Figure 2a are the occupation numbers  $n_{p\Omega}$ . The corresponding LNOs  $\psi_p$  and their occupation numbers  $\lambda_p$  are shown in Figure 2b, with the quantities in brackets being the relative contributions to the overall CBO( $\sigma$ ) and CBO( $\pi$ ) values. Probably the most immediate observation is that the various pictorial results, and indeed also the numerical ones, are relatively insensitive to the flexibility of the CASSCF construction that we used. (All of the total energies are listed in Table S2 in the Supplemental data.) The previously reported results for DAFH analysis of a spin-coupled (or full generalised VB) wavefunction for C<sub>2</sub> [16] also closely resemble those shown in Figure 2a.

«Figure 2 near here»

It proves convenient to start with an examination of the DAFH functions and LNOs from the  $\sigma$ -only CASSCF calculation, returning later to the description of the  $\pi$  bonding. In order to put these various  $\sigma$  functions in a proper context, we have reproduced them in Figure 3 alongside the corresponding results based on  $\sigma$ -only CASSCF calculations for Be<sub>2</sub> and HC≡CH, for the same central bond length. In the case of HC≡CH, with  $R_{CH}=1.06\text{Å}$ , the analysis has been done in terms of two CH domains, each formed as the union of two atomic QTAIM domains. It is clear from Figure 3 that it is the various pictorial representations for Be<sub>2</sub> that more closely resemble those for HC≡CH, with the results for C<sub>2</sub> being more different from those for HC≡CH. Similarly, we observe that the numerical results for C<sub>2</sub> are somewhat more similar to those for Be<sub>2</sub> than they are to those for HC≡CH, which, in turn, resemble those presented previously for N<sub>2</sub> [19] (see also Figures S1 and S2 in the Supplemental data.)

«Figure 3 near here»

The first DAFH function for HC≡CH, with occupation close to 2, takes the form expected for a CH  $\sigma$  bond (see Figure 3a). The second one, with occupation close to 1, can be

associated with the contribution of the dangling valence from this domain to a two-centre two-electron CC  $\sigma$  bond. Looking instead at the LNOs (see Figure 3b), we observe a pair of essentially doubly-occupied functions for the CH bonds and then a further practically doubly-occupied orbital for the CC  $\sigma$  bond. In contrast to the three  $\sigma$  electrons per domain in HC $\equiv$ CH there are of course only two for Be<sub>2</sub>. The occupation of the first DAFH function (see Figure 3a) is reduced by *ca.*  $\frac{3}{4}$  with the remaining  $\frac{1}{4}$  being taken from the occupation of the second DAFH function, which also becomes less symmetric. There are corresponding reductions in the occupation numbers of the LNOs (see Figure 3b), with the value for the ‘bond-like’ one becoming 0.76. It is clear from the pictorial representation of this function that it is somewhat less localised in the central region of this molecule – there are now clearly significant ‘outwards’ contributions.

Turning now to C<sub>2</sub>, the first DAFH function (see Figure 3a) appears to be somewhat less focussed in the ‘outwards’ region than was the case for Be<sub>2</sub> and HC $\equiv$ CH. The same observation applies to the first pair of LNOs (see Figure 3b). Comparing C<sub>2</sub> with Be<sub>2</sub>, the second DAFH function is somewhat less symmetric and the ‘bond-like’ LNO exhibits still larger ‘outwards’ contributions, with reduced localization in the central region. As such, these functions appear to be somewhat less well suited to describing shared electron pair bonding than was the case for Be<sub>2</sub> and (especially) HC $\equiv$ CH. Alongside these various changes to the shapes, there are also some small shifts in the occupation numbers from Be<sub>2</sub> to C<sub>2</sub>, again at the expense of the functions that appear to be most suited to describing bonding.

Although it is clear from Figure 3 that we do observe a  $\sigma$  DAFH function for C<sub>2</sub> that takes the form expected for the dangling valence of a traditional two-centre  $\sigma$  bond, its occupation number is just 0.65 at this geometry. There is of course an analogous function for the other C domain but, instead of any highly questionable attempts to associate these  $2 \times 0.65 = 1.3$  electrons with net  $\sigma$  bonding, it is important to notice that the corresponding value for Be<sub>2</sub> is higher,  $2 \times 0.75 = 1.5$ . We do of course know that Be<sub>2</sub> is unbound at the geometry we have considered [52]. As such, we must accept that the various pictorial depictions of DAFH functions (and of LNOs) seem to correspond to hypothetical bonding descriptions that do not take full account of repulsive interactions such as Pauli repulsion. Only when those repulsive interactions are less important will the results closely resemble our usual expectations for shared electron pair bonding. To a large extent, the latter appears to be true for N<sub>2</sub> and for HC $\equiv$ CH, but not for Be<sub>2</sub> or C<sub>2</sub>.

Fortunately, a rather clear cut distinction between Be<sub>2</sub> (dissociative) and N<sub>2</sub> (two-centre two-electron  $\sigma$  bond) emerged in the previous study in which we used much the same methodology as we have employed here [19]. In the case of a  $\sigma$ -only CASSCF description of Be<sub>2</sub>, we found that the significant CBO( $\sigma$ ) value arose predominantly from LNOs that do not take the anticipated forms for describing shared electron pair bonding. For N<sub>2</sub>, on the other hand, more than 95% of the CBO( $\sigma$ ) value was associated with a ‘bond-like’  $\sigma$  LNO. For the various CASSCF constructions considered in the present work, the ‘bond-like’  $\sigma$  LNOs contribute *ca.* 96% and 99%, respectively, to the CBO( $\sigma$ ) values of 1.0 for N<sub>2</sub> and HC $\equiv$ CH (see Figures S1b and S2b in the Supplemental data and Figure 3b). The corresponding contribution to the CBO( $\sigma$ ) value of 0.9 for Be<sub>2</sub> is *ca.* 16% (see Figure 3b) whereas it is even lower for the X<sup>1</sup> $\Sigma_g^+$  ground state of C<sub>2</sub> near its equilibrium geometry: depending on the choice of CASSCF construction, the contribution from the ‘bond-like’  $\sigma$  LNO to the CBO( $\sigma$ ) value of *ca.* 1.0 ranges from just 5.5% to 7% (see Figure 2b). (The numerical values of W-M( $\sigma$ ), W-M( $\pi$ ), CBO( $\sigma$ ) and CBO( $\pi$ ) for Be<sub>2</sub>, C<sub>2</sub>, N<sub>2</sub> and HC $\equiv$ CH are listed in Table S3 in the Supplemental data.)

All in all, the description of the  $\sigma$  system of the ground state of C<sub>2</sub> that emerges from our analysis is rather different from that for HC $\equiv$ CH. Indeed, there are stronger similarities for C<sub>2</sub> to the case of Be<sub>2</sub>, albeit that the latter does not also feature  $\pi$  electrons. Returning to the CASSCF(8,16) and full valence CASSCF results for C<sub>2</sub> that are depicted in Figure 2, we observe approximately singly-occupied  $\pi$  DAFH functions that take the anticipated forms for the dangling valences of conventional  $\pi$  bonds. The corresponding functions for N<sub>2</sub> and HC $\equiv$ CH (see Figures S1 and S2 in the Supplemental data) are rather similar to these. We observe for all three molecules that it is the ‘bond-like’  $\pi$  LNOs which account for practically all of the CBO( $\pi$ ) value, but that excitations from  $\pi_u$  (bonding) into  $\pi_g$  (antibonding) do lead to reductions in the occupation numbers and  $\pi$ -only bond orders from the RHF values of 2. We note that the partial population of antibonding natural orbitals that accompanies the inclusion of electron correlation leads to the same general trend in occupation numbers and bond orders for other molecules.

There is clearly no evidence in any of our analysis so far for an ethyne-like description of C<sub>2</sub> at its equilibrium geometry. On the other hand, as was anticipated earlier, such a picture does turn out to be fairly appropriate for the lowest <sup>3</sup> $\Sigma_u^+$  excited state of C<sub>2</sub> ( $c^3\Sigma_u^+$ ), for which we have carried out a full-valence CASSCF calculation at the equilibrium geometry of the

ground state. The various DAFH functions (for the domain of one C atom) are shown in Figure 4a and the corresponding LNOs are depicted in Figure 4b. We observe ‘outwards’  $\sigma$  hybrids with occupation numbers a little above 1 for DAFH functions and LNOs. Alongside these, there is a ‘bond-like’  $\sigma$  LNO (as well as a corresponding DAFH function that looks like the dangling valence of a two-centre  $\sigma$  bond). Given our earlier warnings that such pictures need not take proper account of repulsive interactions it is important to check the contribution from the ‘bond-like’  $\sigma$  LNO to the  $\text{CBO}(\sigma)$  value: it exceeds 98%. The DAFH functions and LNOs of  $\pi$  symmetry present no surprises. All in all, the description that emerges here for the  $c^3\Sigma_u^+$  excited state of  $\text{C}_2$  does indeed correspond fairly closely to the description of the triple bond in  $\text{HC}\equiv\text{CH}$ , augmented with two outwards pointing hybrids. (Also shown in Figure 4 are the corresponding RHF results for  $\text{C}_2$  ( $^3\Sigma_u^+$ ) which, in general terms, are rather similar to those based on the full-valence CASSCF wavefunction. Values of  $\text{CBO}(\sigma)$  and  $\text{CBO}(\pi)$  of 0.85 and 2 for the RHF description become 0.82 and 1.78, respectively, at the full-valence CASSCF level.)

«Figure 4 near here»

According to the multireference average quadratic coupled cluster (MR-AQCC) calculations of Müller *at al.* [54] the equilibrium bond length in the  $c^3\Sigma_u^+$  state of  $\text{C}_2$  (which also dissociates to ground state atoms) is shorter than that in the  $X^1\Sigma_g^+$  ground state by more than 0.03 Å, but this excited state lies more than 1 eV higher. Because our various modes of analysis are primarily designed to reveal the nature of bonding interactions, they are inherently unable to address directly questions concerning the relative energies of different electronic states. The resolution of the apparent paradox that a state with an ethyne-like triple bond lies so much higher in energy than the ground state would thus require the availability of additional information taken from explicit quantum chemical calculations and/or from auxiliary qualitative models.

### 3.3. Geometry dependence of the bonding in the $X^1\Sigma_g^+$ ground state of $\text{C}_2$

We turn now to the evolution of the bonding in the  $X^1\Sigma_g^+$  ground state of  $\text{C}_2$  as the nuclear separation is decreased towards  $R_e$ , starting from ground state  $\text{C}(^3\text{P})$  atoms. As indicated earlier, we employed for this purpose a CASSCF(8,16) construction in which, relative to full-valence CASSCF, the number of active orbitals of each symmetry (in  $D_{2h}$ ) was doubled. The resulting potential energy curve for geometries from 1 Å to 3.25 Å is displayed in Figure 5a. (All of the numerical values depicted in Figure 5 are listed in Tables S4, S5 and S6 in the

Supplemental data.) It seems likely that the very slight change of curvature that is visible in the potential curve near 1.55 Å, which is in the vicinity of the crossing with the  $B^1\Delta_g$  state and of an avoided crossing with the  $B'^1\Sigma_g^+$  state [53], would be ‘smoothed out’ by further extensions to the flexibility of the wavefunction and/or by state-averaging.

«Figure 5 near here»

For very large nuclear separation  $R$ , the relative orientation of the 2p orbitals on the two  $C(^3P)$  atoms is of course entirely arbitrary from the point of view of the total energy. However, following the  $X^1\Sigma_g^+$  potential to shorter  $R$  using the CASSCF(8,16) construction we observe that the system adopts by 3 Å a specific relative orientation such that the total number of  $\pi$  electrons  $n(\pi)$  (*i.e.* the sum of the  $b_{3u}$ ,  $b_{2u}$ ,  $b_{2g}$  and  $b_{3g}$  natural orbital occupation numbers in  $D_{2h}$  symmetry) is a little above 2. Specifically,  $n(\pi)=2.12$  for  $R=3\text{Å}$ . We speculate that this is the preferred relative orientation of the two moieties for putative  $\sigma$  bonding interactions to be formed. The W-M index is already 0.19 for this geometry and the value of CBO is 0.10. Both of these values are dominated by the contributions from the  $\sigma$  electrons.

The geometry dependence of  $n(\pi)$  for the CASSCF(8,16) description of the  $X^1\Sigma_g^+$  ground state of  $C_2$  is shown in Figure 5b for nuclear separations from  $R_e$  to 3.25 Å. We observe that  $n(\pi)$  initially increases relatively slowly as  $R$  decreases but that there is then a dramatic switch to  $n(\pi)\sim 4$ , with the most rapid changes occurring for nuclear separations from 1.5 Å to 1.6 Å. This rapid change is of course associated with the avoided crossing between the  $X^1\Sigma_g^+$  and  $B'^1\Sigma_g^+$  states. We show in Figure 5c the geometry dependence of the total W-M index: it increases fairly steadily as  $R$  is reduced, with the fastest changes occurring in the same region as the most rapid changes to  $n(\pi)$ . Also shown in Figure 5c are the separate  $\sigma$  and  $\pi$  contributions to the total W-M index. To a first approximation,  $W-M(\pi)$  varies in a similar fashion to the total index, but the behaviour of  $W-M(\sigma)$  is somewhat different. As  $R$  is reduced,  $W-M(\sigma)$  initially increases towards a maximum value of *ca.* 0.9 before it decreases towards a local minimum in the region of the rapid change of  $n(\pi)$ . At still shorter  $R$ , approaching  $R_e$ ,  $W-M(\sigma)$  increases again to a value of *ca.* 1.

The geometry dependences of  $CBO(\sigma)$ ,  $CBO(\pi)$  and the total CBO value, shown in Figure 5d, are analogous to those depicted in Figure 5c except that some of the changes are clearly somewhat more abrupt. The clear observation of two distinct ranges of  $R$  in which  $CBO(\pi)$  changes rather rapidly appears to suggest that the breaking of the  $\pi$  bonds proceeds in two steps, which would be rather unusual. Nothing like this has been observed in previous work for the splitting of ordinary triple bonds, such as in  $N_2$ , for which the dissociation of the

two  $\pi$  components proceeds in a single one-step process (albeit at a different geometry than for the breaking of the  $\sigma$  bond) [55,56].

Taken together, the geometry dependencies shown in Figures 5b-5d already hint at an interesting evolution of the bonding in the  $X^1\Sigma_g^+$  ground state of  $C_2$ . In order to elucidate the nature of the various changes over this range of  $R$ , we turn now to an examination of the DAFH functions and of the LNOs at four representative geometries. Pictorial results of our DAFH and LNO analysis, together with relevant numerical values, are collected in Figures 6a and 6b, respectively, with successive rows corresponding to nuclear separations of 2.1 Å, 1.75 Å, 1.5 Å and 1.2425 Å. We also show in Figure 5e the geometry dependence of the occupation numbers for the ‘bond-like’  $\sigma$  LNO and for the first LNO of  $\pi$  symmetry. (The corresponding changes to the occupancies of the degenerate pair of  $\sigma$  LNOs are very much smaller.)

Looking first at the results for a nuclear separation of 2.1 Å (top row of Figures 6a and 6b), we observe that the forms and occupation numbers of the first DAFH function and of the corresponding degenerate pair of LNOs are indicative of almost doubly-occupied lone pairs on the two centres. These LNOs contribute relatively little (2.2% each) to the value of  $CBO(\sigma)$ . Instead, it is a ‘bond-like’ LNO of  $\sigma$  symmetry (occupation 1.81) which accounts for almost 95% of the  $CBO(\sigma)$  value of 0.86. The form of this LNO suggests a ‘stretched’  $\sigma$  bond and the second DAFH function (occupation 0.97) looks like the dangling valence of this long bond. (A further  $\sigma$  LNO, not shown in Figure 6b, has an occupation of 0.15 but it contributes less than 0.7% to the  $CBO(\sigma)$  value.) As for the  $\pi$  system, it is helpful to imagine that there are singly-occupied unperturbed  $2p_x$  orbitals on each carbon centre but then to recall, because of the requirement for cylindrical symmetry, that the wavefunction must correspond to the average of this hypothetical arrangement and the one in which there are instead singly-occupied unperturbed  $2p_y$  orbitals on each carbon centre. As such, the net populations of each of the various DAFH functions and LNOs of  $\pi$  symmetry should be  $\frac{1}{2}$ . In practice, for the CASSCF(8,16) description of the  $X^1\Sigma_g^+$  ground state of  $C_2$  for this value of  $R$ , those populations are 0.54 and there are already some interactions between the  $2p_\pi$  functions on the two centres, such that  $CBO(\pi)=0.12$ . The values of  $W-M(\sigma)$  and  $W-M(\pi)$  at this geometry are 0.78 and 0.36, respectively.

«Figure 6 near here»

As the carbon atoms are brought closer together, to  $R=1.75$  Å, the changes to the shapes of the  $\sigma$  DAFH functions and LNOs of  $\sigma$  symmetry (see second row of Figures 6a and 6b)



simply reflect a shortening of the ‘stretched’  $\sigma$  bond and there is just a small increase in  $\text{CBO}(\sigma)$  to 0.88. We observe that very much larger changes occur in the  $\pi$  system, corresponding to the ongoing formation of effectively a net single  $\pi$  bond, with  $\text{CBO}(\pi)$  increasing to 0.51. In order to interpret the occupations of the  $\pi$  LNOs at this geometry, we can imagine a doubly-occupied  $\pi_x$  orbital but then, in order to respect the cylindrical symmetry, consider instead the average of this hypothetical arrangement with the one in which it is instead a  $\pi_y$  orbital that is doubly occupied. In this way, we would expect a degenerate pair of  $\pi_x/\pi_y$  LNOs, each with net occupancy 1. This is in effect what is observed, with each of these orbitals contributing 49.3% of the  $\text{CBO}(\pi)$  value. The values of  $\text{W-M}(\sigma)$  and  $\text{W-M}(\pi)$  at this geometry are 0.88 and 0.60, respectively.

Continuing to shorter  $R$ , we start to observe the formation of a fully populated  $\pi$  double bond, which is reflected in an increase in the occupation of each of the  $\pi$  LNOs from 1.0 at 1.75 Å to 1.7 at 1.5 Å (see third row of Figures 6a and 6b). There is similarly a large increase in the value of  $\text{CBO}(\pi)$ , which reaches 1.52 for  $R=1.5$  Å. Of course, such increases in the total occupancy of the natural orbitals and LNOs of  $\pi$  symmetry (see also Figures 5b and 5e) must be at the expense of the  $\sigma$  system. Accordingly, it is clear from Figure 6 that there are quite dramatic changes between 1.75 Å (second row of Figures 6a and 6b) and 1.5 Å (third row of Figures 6a and 6b) in the occupancies of the various functions of  $\sigma$  symmetry. In particular, the occupancy of the ‘bond-like’ LNO of  $\sigma$  symmetry falls from 1.84 to 0.51. The value of  $\text{CBO}(\sigma)$  is still fairly high (0.84) but relatively little of it (7.6%) is still associated with the ‘bond-like’  $\sigma$  LNO. Instead, the value of  $\text{CBO}(\sigma)$  is dominated by the first degenerate pair of  $\sigma$  LNOs. Such a situation is of course reminiscent of what we observed for  $\text{Be}_2$ , albeit that the shapes of the formally ‘nonbonding’  $\sigma$  DAFH functions and LNOs in  $\text{Be}_2$  seemed more closely to resemble quasi lone pairs. The change to the shape of the first  $\sigma$  LNO of  $\text{C}_2$  between 1.75 Å and 1.5 Å corresponds to an increase from 1.01 to 1.31 in the localization number ( $\tilde{n}_p$ ), which measures the effective number of domains that it spans (see Equation (5)). Continuing to  $R=1.2425$  Å (fourth row of Figure 6b) we observe a further increase for  $\text{C}_2$  in the occupation of the  $\pi$  LNOs and a further reduction in the contribution made by the ‘bond-like’ LNO of  $\sigma$  symmetry to the value of  $\text{CBO}(\sigma)$ . The value of  $\tilde{n}_p$  for the first  $\sigma$  LNO increases to 1.46. (The corresponding values from  $\sigma$ -only CASSCF calculations at the same geometry for  $\text{Be}_2$ ,  $\text{C}_2$  and  $\text{N}_2$  are 1.40, 1.43 and 1.01, respectively.) The geometry dependence for the CASSCF(8,16) description of  $\text{C}_2$  of the values of  $\tilde{n}_p$  for the key LNOs is shown in Figure 5f. The full reasons for the increase in the localization index at short  $R$  for the degenerate pair of  $\sigma$  LNOs currently remain unclear and deserve further scrutiny. This is

an issue to which we intend to return in future work, comparing the geometry dependence of such localization indices for a wide range of molecules.

The basic picture that emerges from our analysis of the evolution of the pattern of bonding in the  $X^1\Sigma_g^+$  ground state of  $C_2$ , using a CASSCF(8,16) construction, is thus as follows. Bringing the atoms together, starting at 3.25 Å, the  $\sigma$  system involves *ca.* six electrons, of which four are accommodated in quasi lone pairs and the other two start to form a ‘stretched’  $\sigma$  bond. As  $R$  is reduced, this  $\sigma$  bonding is accompanied by the formation of effectively a net single  $\pi$  bond. It is only when the atoms are brought still closer together that the fully populated  $\pi$  double bond forms, ‘stealing’ the required electrons primarily from the  $\sigma$  bond rather than from the nonbonding  $\sigma$  orbitals. As a consequence, the  $\sigma$  system of  $C_2$  near  $R_e$  becomes somewhat more akin to that in  $Be_2$  than it is to that in  $HC\equiv CH$ .

Even though they feature two more  $\sigma$  valence electrons than do  $Be_2$  and  $C_2$ , the  $\sigma$  systems in the ground states of  $N_2$  and  $HC\equiv CH$  near their equilibrium geometries are somewhat simpler. In essence, this is because it is favourable to ‘allocate’ four of the six  $\sigma$  electrons to N lone pairs or CH bonds, leaving just two of them to be distributed in the central region. This leads of course to fairly ordinary two-centre two-electron  $\sigma$  bonding in both cases. The situation in  $Be_2$  and  $C_2$  (for short nuclear separations) is more complicated, with four  $\sigma$  electrons competing for the same region of space, where they are subject to strong mutual repulsion. Obviously, they need to be distributed in a fashion that reduces that repulsion. We note that  $C_2$ , but not  $Be_2$ , could relieve some of this Pauli repulsion by adopting the antiferromagnetic coupling scheme that was highlighted by Xu and Dunning [13].

Perhaps the simplest distribution for the four  $\sigma$  valence electrons of  $Be_2$  and  $C_2$  would be to allocate them to doubly-occupied quasi lone pairs. Not only might such an arrangement seem to suggest no  $\sigma$  bonding at all, but the Pauli repulsion between the quasi lone pairs could signify a net repulsive interaction. On the other hand, it is certainly the case for both of these molecules that the  $2\sigma_g$  (bonding) and  $2\sigma_u$  (antibonding) orbitals are based on different hybrids, so that they do not completely cancel one another. As such, we can expect there to be a small residual  $\sigma$  bonding component. Furthermore, the incompleteness of this cancellation could be enhanced in post-RHF treatments by the depopulation of  $2\sigma_u$  that arises primarily from excitations into the  $3\sigma_g$  orbital. On the other hand, these residual  $\sigma$  bonding components in both molecules will still be countered to varying degrees by Pauli repulsion.

There is of course a region of sufficiently short nuclear separations for which increases in various measures of bond order as two atoms are brought still closer together corresponds

predominantly to repulsive interactions [20]. Our analysis of  $\sigma$ -only CASSCF wavefunctions for the ground states of  $\text{Be}_2$  and  $\text{C}_2$  suggests that both of these systems could already be in this region at  $R=1.2425 \text{ \AA}$ . Whereas it is true that the values of  $W\text{-}M(\sigma)$  and  $\text{CBO}(\sigma)$  remain relatively high for the ground state of  $\text{C}_2$  near  $R_e$ , it now appears that these values arise to a large extent for much the same reasons as for  $\text{Be}_2$  (which is unbound at the same geometry) and so we cannot interpret them as signifying an ‘effective’ single  $\sigma$  bond.

This impossibility thus clearly suggests the limitations of the intuitive interpretation of nominal values of bond orders in terms of classical bond multiplicities. One of the reasons is that these concepts were originally designed for the description of bonding situations that correspond reasonably well to classical Lewis models of shared electron pair bonds and their usefulness for a given molecule thus depends to a significant degree on the extent to which the actual bonding situation conforms to such classical bonding paradigms. This problem turns out to be particularly important for the  $X^1\Sigma_g^+$  ground state of  $\text{C}_2$  near its equilibrium geometry: the various contributions within the  $\sigma$  system could be dominated by the non-classical repulsive interactions, making it inappropriate to interpret nominal values of bond orders as a direct indication of the bond multiplicity.

The description of the  $X^1\Sigma_g^+$  ground state of  $\text{C}_2$  near  $R_e$  that emerges from our analysis clearly does not correspond to the supposed model of an ethyne-like triple bond augmented with two outwards pointing hybrids (which may or may not interact so as to constitute a further bonding interaction). Whereas this sort of picture is indeed what we found when using comparable analysis for the  $c^3\Sigma_u^+$  excited state of  $\text{C}_2$  which also dissociates to ground state atoms, it turns out that the  $\sigma$  system of the  $X^1\Sigma_g^+$  ground state looks much more like the one we observed for  $\text{Be}_2$ . A fundamental difference between the ground states of  $\text{Be}_2$  and  $\text{C}_2$  for  $R=1.2425 \text{ \AA}$  is, of course, the presence also of the  $\pi$  double bond for  $\text{C}_2$ . Our analysis suggests that this  $\pi$  double bond is fairly similar (near  $R_e$ ) to those in  $\text{N}_2$  and  $\text{HC}\equiv\text{CH}$ .

## 4. Conclusions

The various modes of analysis deployed here provide interesting new insights into the unusual manner in which the bonding in the  $X^1\Sigma_g^+$  ground state of  $\text{C}_2$  changes with nuclear separation  $R$ . The  $\sigma$  system of this molecule turns out to be particularly sensitive, evolving from a description that includes a fairly traditional shared electron pair  $\sigma$  bond, for a range of intermediate  $R$  values, to a situation in which repulsive interactions could be dominant, near

$R_e$ . These dramatic changes are closely tied to the  $R$ -dependent redistribution of electrons between the  $\sigma$  and  $\pi$  systems. While at short  $R$ , close to  $R_e$ , the electronic structure of the  $X^1\Sigma_g^+$  ground state of  $C_2$  reflects the existence of four  $\sigma$  and four  $\pi$  valence electrons, an increase in the nuclear separation through the region of the avoided crossing with the  $B^1\Sigma_g^+$  state [53] results in the transfer of two of the  $\pi$  electrons into the  $\sigma$  framework, thereby increasing the number of  $\sigma$  valence electrons to six and reducing the number of  $\pi$  electrons to two. A direct consequence of this redistribution of valence electrons is the observed formation, at intermediate  $R$ , of a shared electron pair  $\sigma$  bond, but at the expense of a reduction in the bonding provided by the  $\pi$  system.

Our results, for any of the geometries that we have considered, do not even support notions of a triple bond ( $1\times\sigma$  and  $2\times\pi$ ), whether or not there is also a weaker ‘inverted’  $\sigma$  bond. According to our analysis, the  $\sigma$  system of  $C_2$  ( $X^1\Sigma_g^+$ ) near  $R_e$  has more in common with that of  $Be_2$  than it does with that of ethyne,  $HC\equiv CH$ , whereas the  $\pi$  system of  $C_2$  at such geometries turns out to be fairly similar to that in  $HC\equiv CH$ . As a consequence, the electronic structure of the  $X^1\Sigma_g^+$  ground state of  $C_2$  is dominated near  $R_e$  by the  $\pi$  double bond and the non-classical  $\sigma$  components. Similarly, because the formation of the two-centre two-electron  $\sigma$  bond at intermediate  $R$  requires the partial depopulation of the  $\pi$  system, the shared electron pair  $\sigma$  bond and a fully populated  $\pi$  double bond never coexist. Note that we do observe an ethyne-like triple bond in the case of the  $c^3\Sigma_u^+$  excited state of  $C_2$ .

The dramatic  $R$ -dependent coupling between the  $\sigma$  and  $\pi$  components, as well as the accompanying deep change in the character of the  $\sigma$  system, evidently transcend the scope of traditional bonding paradigms. As such, the present study provides further support for the view [13] that the classification of the bonding in this molecule in terms of classical concepts of bond order or bond multiplicity is highly questionable.

### **Disclosure statement**

No potential conflict of interest was reported by the author(s).

### **Supplemental data**

Supplemental data for this article can be accessed <http://dx.doi.org/10.1080/...>

**Acknowledgement**

We thank Gernot Frenking (Philipps-Universität Marburg) for helpful comments and for sending us a preprint of Ref. 17.

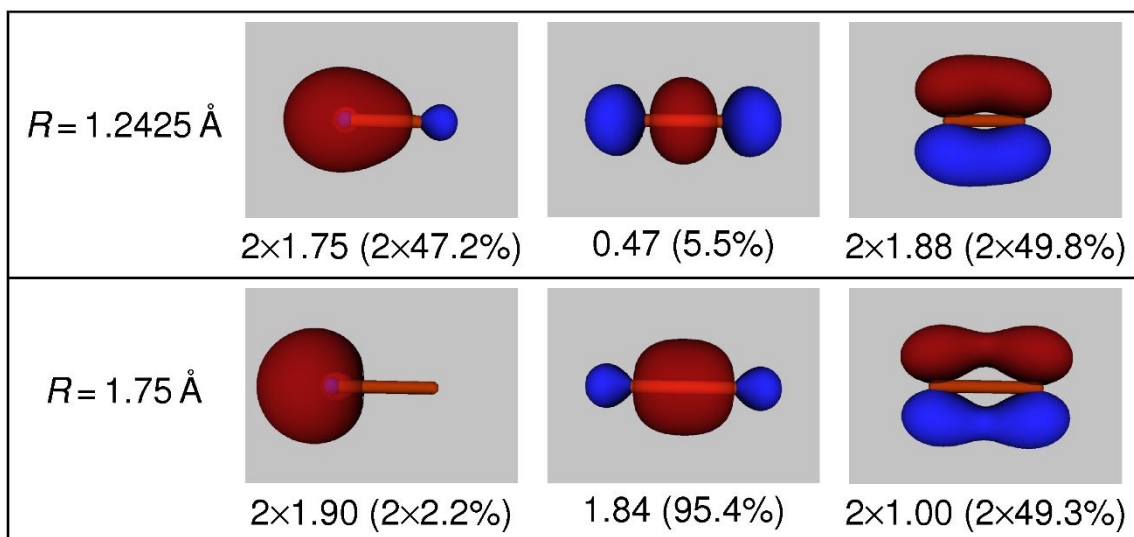
## References

- [1] R.S. Mulliken, Phys. Rev. **56** (8), 778-781 (1939).
- [2] P.v.R. Schleyer, P. Maslak J. Chandrasekhar, and R.S. Grev, Tetrahedron Lett. **34** (40), 6387-6390 (1993).
- [3] A. Krapp, F.M. Bickelhaupt, and G. Frenking, Chem. Eur. J. **12** (36), 9196-9216 (2006).
- [4] V. Bezugly, P. Wielgus, M. Kohout, and F.R. Wagner, J. Comp. Chem. **31** (7), 1504-1519 (2010).
- [5] P. Su, J. Wu, J. Gu, W. Wu, S. Shaik, and P.C. Hiberty, J. Chem. Theory Comp. **7** (1), 121-130 (2011).
- [6] S. Shaik, D. Danovich, W. Wu, P. Su, H.S. Rzepa, and P.C. Hiberty, Nature Chemistry **4** (3), 195-200 (2012).
- [7] S. Shaik, H.S. Rzepa, and R. Hoffmann, Angew. Chem. Int. Ed. **52** (10), 3020-3033 (2013).
- [8] G. Frenking and M. Hermann, Angew. Chem. Int. Ed. **52** (23), 5922-5925 (2013).
- [9] D. Danovich, S. Shaik, H.S. Rzepa, and R. Hoffmann, Angew. Chem. Int. Ed. **52** (23), 5926-5928 (2013).
- [10] E. Ramos-Cordoba, P. Salvador, and M. Reiher, Chem. Eur. J. **19** (45), 15267-15275 (2013).
- [11] J.M. Matxain, F. Ruipérez, I. Infante, X. Lopez, J.M. Ugalde, G. Merino, and M. Piris, J. Chem. Phys. **138** (15), 151102 (2013)
- [12] J.M.L. Martin, Mol. Phys. **112** (5-6), 785-793 (2014).
- [13] L.T. Xu and T.H. Dunning Jr., J. Chem. Theor. Comp. **10** (1), 195-201 (2014).
- [14] D. Danovich, P.C. Hiberty, W. Wu, H.S. Rzepa, and S. Shaik, Chem. Eur. J. **20** (21), 6220-6232 (2014).
- [15] K. Hendrickx, B. Braïda, P. Bultinck, and P.C. Hiberty, Comput. Theoret. Chem. **1053**, 180-188 (2015).
- [16] D.L. Cooper, F.E. Penotti, and R. Ponec, Comput. Theoret. Chem. **1053**, 189-194 (2015).
- [17] M. Hermann and G. Frenking, 'The chemical bond in C<sub>2</sub>' Chem. Eur. J. (submitted for publication).
- [18] F. Weinhold and C. Landis, *Valency and bonding: A Natural Bond Orbital donor-acceptor perspective* (Cambridge University Press, Cambridge, UK, 2005).

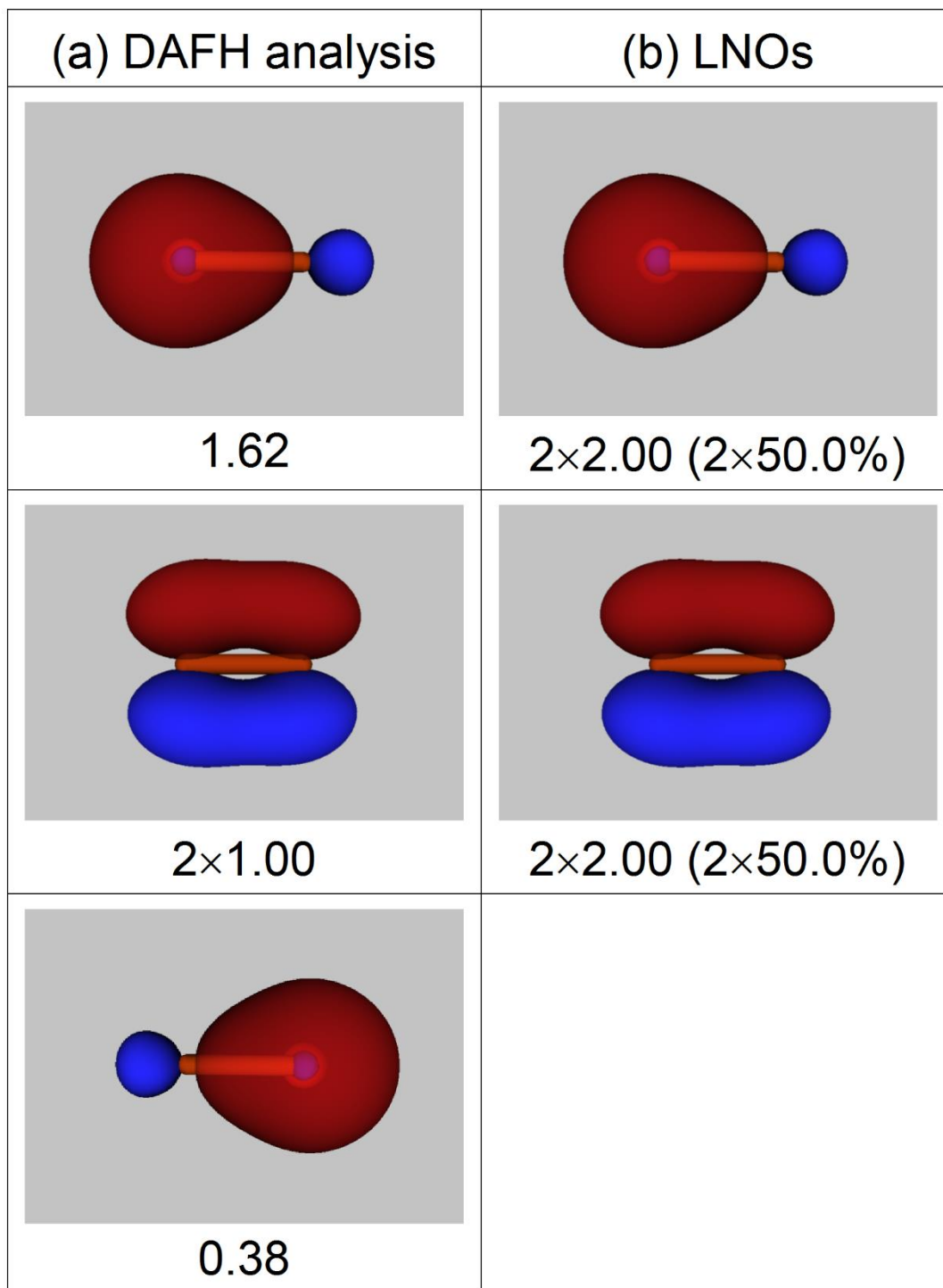
- [19] D.L. Cooper, R. Ponec, and M. Kohout, *Mol. Phys.* **113** (13-14), 1682-1689 (2015).
- [20] M. García-Revilla, P.L.A. Popelier, E. Francisco, and Á.M. Pendás, *J. Chem. Theory Comput.* **7** (6), 1704-1711 (2011).
- [21] R. Ponec, *J. Math. Chem.* **21** (3), 323-333 (1997).
- [22] R. Ponec, *J. Math. Chem.* **23** (1-2), 85-103 (1998).
- [23] R. Ponec and A.J. Duben, *J. Comp. Chem.* **20** (8), 760-771 (1999).
- [24] R. Ponec and J. Roithová, *Theor. Chem. Acc* **105** (4-5), 383-392 (2001).
- [25] R. Ponec and D.L. Cooper, *Faraday Disc.* **135**, 31-42 (2007).
- [26] R. Ponec and D.L. Cooper, *J. Phys. Chem. A* **111** (44), 11294-11301 (2007).
- [27] R. Ponec, D.L. Cooper, and A. Savin, *Chem.-Eur. J.* **14** (11), 3338-3345 (2008).
- [28] P. Bultinck, D.L. Cooper, and R. Ponec, *J. Phys. Chem. A* **114** (33), 8754-8763 (2010).
- [29] D. Tiana, E. Francisco, M.A. Blanco, P. Macchi, A. Sironi, and Á.M. Pendás, *Phys. Chem. Chem. Phys.* **13** (11), 5068-5077 (2011).
- [30] A.I. Baranov, R. Ponec, and M. Kohout, *J. Chem. Phys.* **137** (21), 214109 (2012).
- [31] E. Francisco, Á.M. Pendás, and A. Costales, *Phys. Chem. Chem. Phys.* **16** (10), 4586-4597 (2014).
- [32] R. Ponec and X. Gironés, *J. Phys. Chem. A* **106** (41), 9505-9511 (2002).
- [33] R. Ponec, G. Yuzhakov, X. Gironés, and G. Frenking, *Organometallics* **23** (8), 1790-1796 (2004).
- [34] R. Ponec, G. Yuzhakov, and D.L. Cooper, *Theor. Chem. Acc.* **112** (5-6), 419-430 (2004).
- [35] R. Ponec and G. Yuzhakov, *Theor. Chem. Acc.* **118** (4), 791-797 (2007).
- [36] R. Ponec, G. Lendvay, and J. Chaves, *J. Comp. Chem.* **29** (9), 1387-1398 (2008).
- [37] R. Ponec, G. Lendvay, and M.R. Sundberg, *J. Phys. Chem. A* **112** (40), 9936-9945 (2008).
- [38] R. Ponec and F. Feixas, *J. Phys. Chem. A* **113** (29), 8394-8400 (2009).
- [39] H.-J. Werner and P.J. Knowles, *J. Chem. Phys.* **82** (11), 5053-5063 (1985).
- [40] P.J. Knowles and H.-J. Werner, *Chem. Phys. Lett.*, **115** (3), 259-267 (1985).
- [41] H.-J. Werner, P.J. Knowles, G. Knizia, F.R. Manby, and M. Schütz, *WIREs Comput. Mol. Sci.* **2** (2), 242-253 (2012).
- [42] H.-J. Werner, P.J. Knowles, G. Knizia, F.R. Manby, M. Schütz, P. Celani, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K.R. Shamasundar, T.B. Adler, R.D. Amos, A. Bernhardsson, A. Berning, D.L. Cooper, M.J.O. Deegan, A.J. Dobbyn, F. Eckert, E.

- Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A.W. Lloyd, R.A. Mata, A.J. May, S.J. McNicholas, W. Meyer, M.E. Mura, A. Nicklass, D.P. O'Neill, P. Palmieri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A.J. Stone, R. Tarroni, T. Thorsteinsson, and M. Wang, *MOLPRO, Version 2012.1*. (Cardiff, U. K.). <www.molpro.net>
- [43] R.F.W. Bader, *Atoms in molecules: A quantum theory*. (Oxford University Press, Oxford, UK, 1990).
- [44] I. Mayer, Chem. Phys. Lett. **544**, 83-86 (2012).
- [45] R.S. Mulliken, J. Chem. Phys. **23**, 1833-1840 (1955).
- [46] R.S. Mulliken, J. Chem. Phys. **23**, 1841-1846 (1955).
- [47] J.G. Ángyán, M. Loos, and I. Mayer, J. Phys. Chem. **98** (20), 5244-5248 (1994).
- [48] J. Cioslowski and S.T. Mixon, J. Am. Chem. Soc. **113** (11), 4142-4145 (1991).
- [49] J. Cioslowski, Int. J. Quantum Chem. **S24**, 15-28 (1990).
- [50] T.A. Keith, *AIMAll Version 13.11.04* (TK Gristmill Software, Overland Park KS, USA, 2012). <aim.tkgristmill.com>
- [51] G. Schaftenaar and J.H. Noordik, J. Comput.-Aided Mol. Des. **14** (2), 123-134 (2000). <www.cmbi.ru.nl/molden>
- [52] M. El Khatib, G.L. Bendazzoli, S. Evangelisti, W. Helal, T. Leininger, L. Tenti, and C. Angeli, J. Phys. Chem. A **118** (33), 6664-6673 (2014).
- [53] J.S. Boschen, D. Theis, K. Ruedenberg, and T.L. Windus, Theor. Chem. Acc. **133** (2), 1425 (2014).
- [54] T. Müller, M. Dallos, H. Lischka, Z. Dubrovay, and P.G. Szalay, Theor. Chem. Acc. **105** (3), 227-243 (2001).
- [55] R. Ponec and D. L. Cooper, J. Mol. Struct. THEOCHEM **727** (1-3), 133-138 (2005).
- [56] D.L. Cooper and R. Ponec, Int. J. Quantum Chem. **109** (11), 2383-2392 (2009).

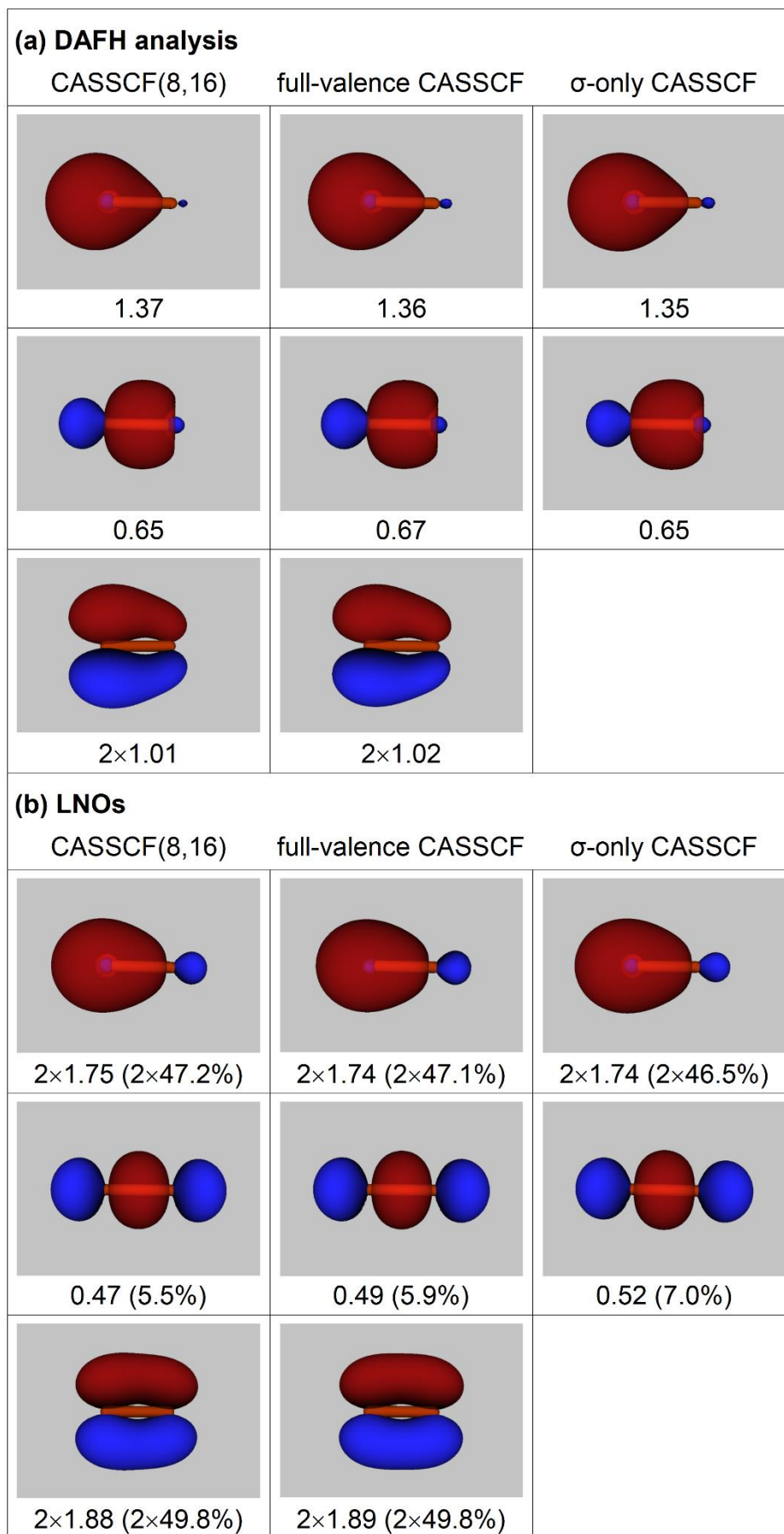




**Graphical Abstract**



**Figure 1.** DAFH functions and LNOs for the  $X^1\Sigma_g^+$  state of  $C_2$  (RHF/3-21G,  $R=1.2425 \text{ \AA}$ ), together with occupation numbers. The quantities in brackets signify the relative contributions to the overall  $CBO(\sigma)$  and  $CBO(\pi)$  values.



**Figure 2.** DAFH functions and LNOs for  $C_2 (X^1\Sigma_g^+)$ , together with occupation numbers. The quantities in brackets signify the relative contributions to the overall  $CBO(\sigma)$  and  $CBO(\pi)$  values. ( $R=1.2425 \text{ \AA}$ , cc-pVQZ basis.)

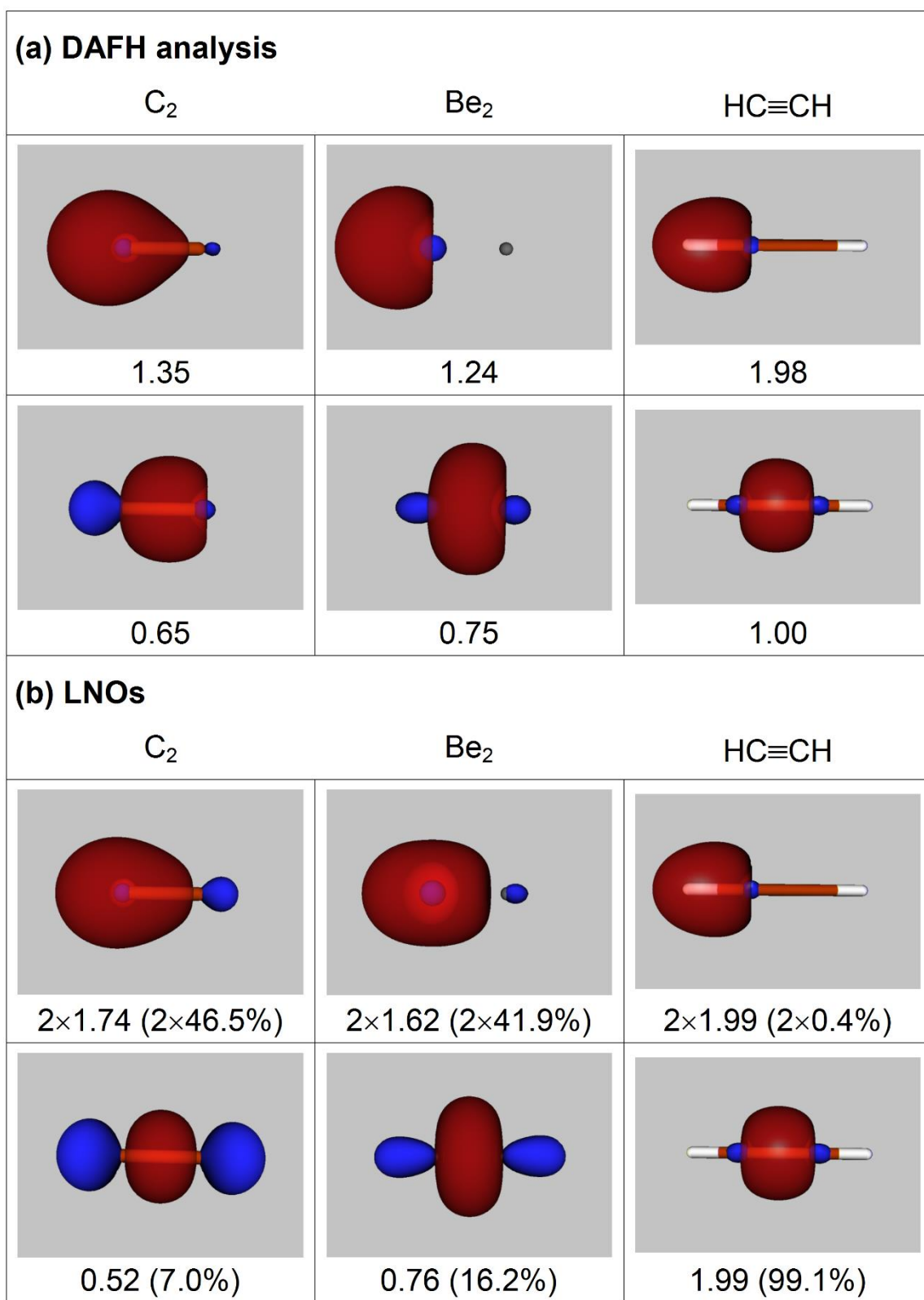


Figure 3. DAFH functions and LNOs for  $\sigma$ -only CASSCF descriptions of the ground states of  $C_2$ ,  $Be_2$  and  $HC\equiv CH$ , together with occupation numbers. Quantities in brackets signify the relative contributions to the overall  $CBO(\sigma)$  values. ( $R=1.2425 \text{ \AA}$ ,  $R_{CH}=1.06 \text{ \AA}$ , cc-pVQZ basis.)  $HC\equiv CH$  was analyzed as two CH domains.

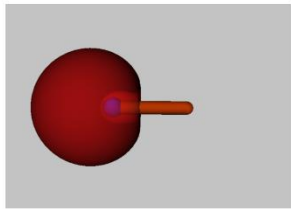
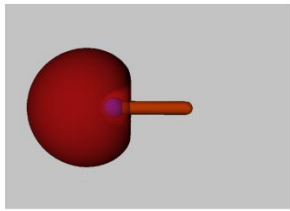
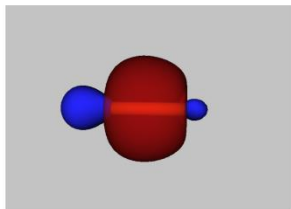
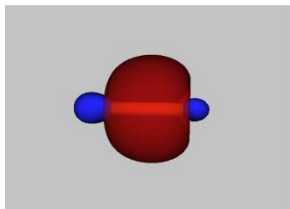
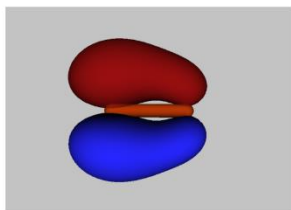
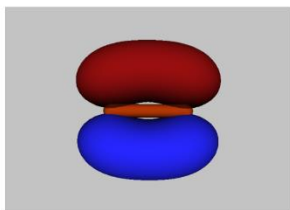
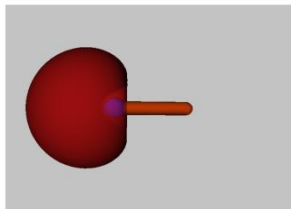
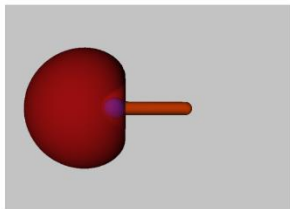
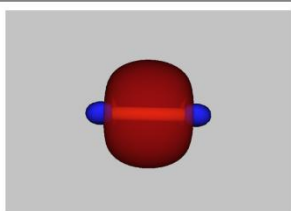
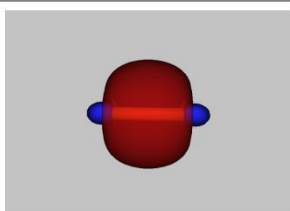
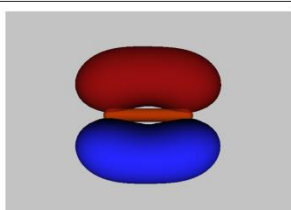
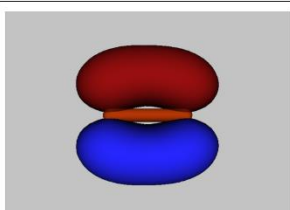
<b>(a) DAFH analysis</b>	
full-valence CASSCF	RHF
 1.14	 1.09
 0.85	 0.88
 2×1.01	 2×1.00
<b>(b) LNOs</b>	
full-valence CASSCF	RHF
 2×1.09 (2×0.8%)	 2×1.09 (2×0.8%)
 1.79 (98.3%)	 1.83 (98.4%)
 2×1.88 (2×49.8%)	 2×2.00 (2×50.0%)

Figure 4. DAFH functions and LNOs for the lowest  ${}^3\Sigma_u^+$  excited state of  $\text{C}_2$ , together with occupation numbers. The quantities in brackets signify the relative contributions to the overall  $\text{CBO}(\sigma)$  and  $\text{CBO}(\pi)$  values. ( $R=1.2425 \text{ \AA}$ , cc-pVQZ basis.)

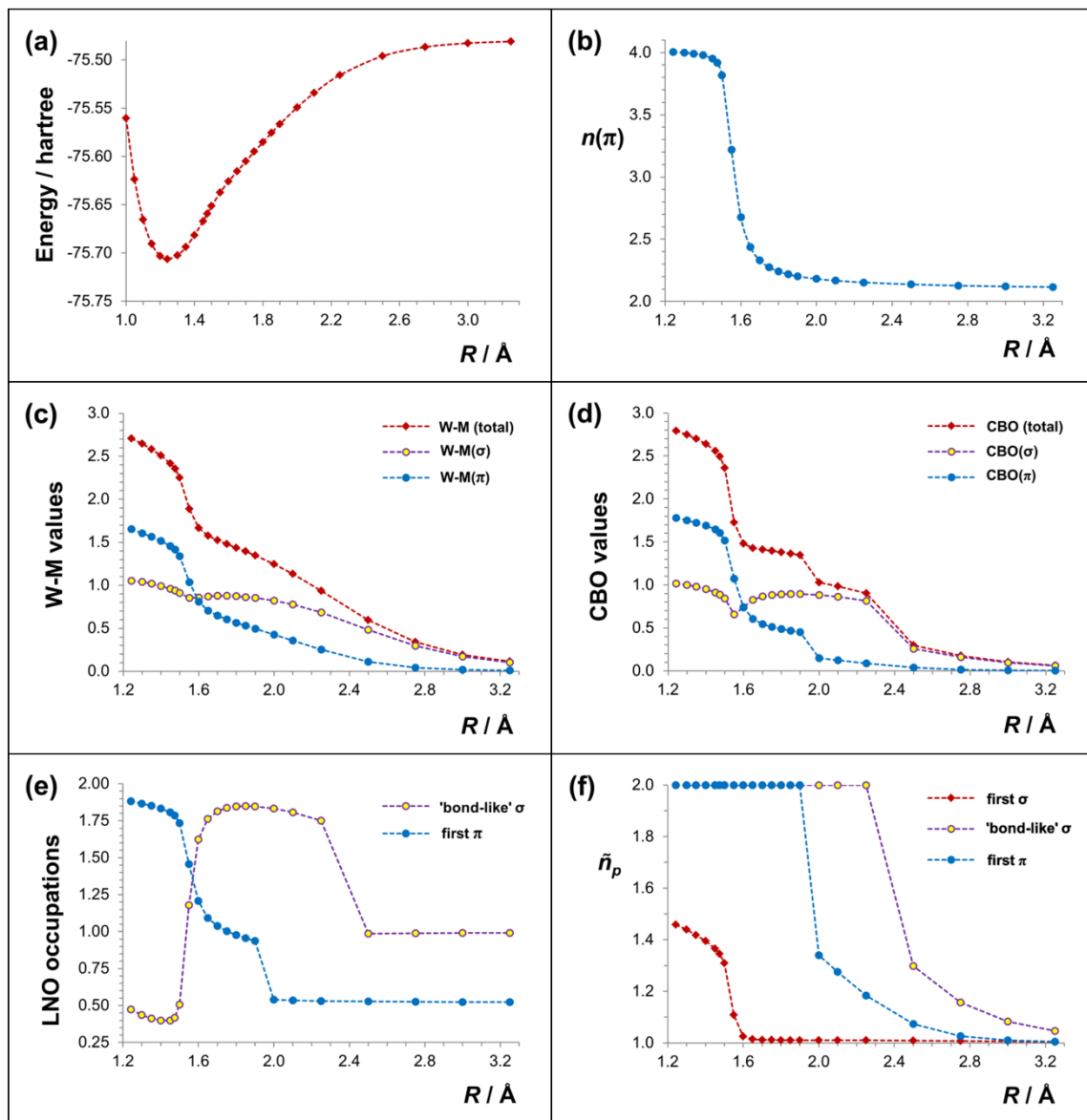


Figure 5. Geometry dependence of various quantities for the  $X^1\Sigma_g^+$  ground state of  $C_2$ : (a) Total energies; (b) Total number of  $\pi$  electrons; (c) W-M values; (d) CBO values; (e) LNO occupation numbers; (f) Localization numbers,  $\tilde{n}_p$ , for key LNOs. (CASSCF(8,16), cc-pVQZ basis.)

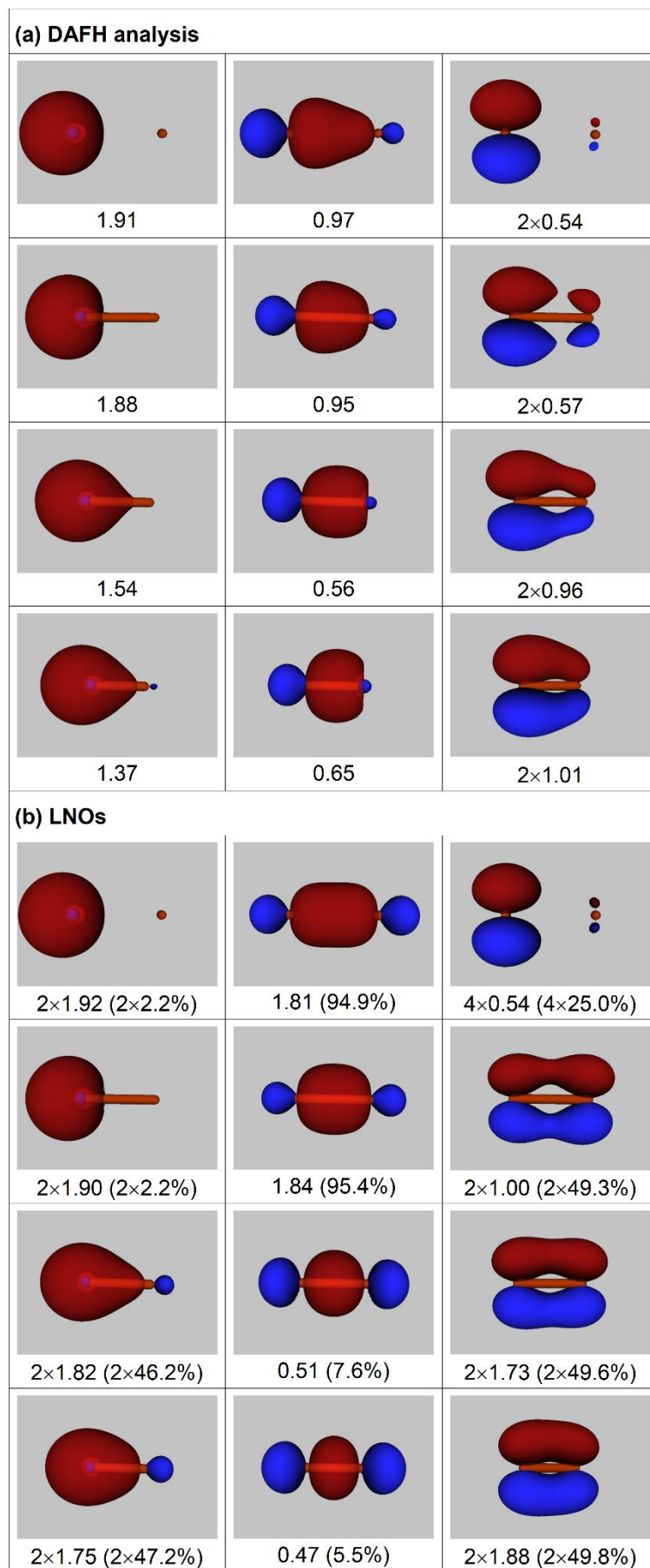


Figure 6. DAFH functions and LNOs for the  $X^1\Sigma_g^+$  ground state of  $C_2$ , together with occupation numbers. Successive rows correspond to nuclear separations of 2.1 Å, 1.75 Å, 1.5 Å and 1.2425 Å. Quantities in brackets signify relative contributions to the overall CBO( $\sigma$ ) and CBO( $\pi$ ) values. (CASSCF(8,16), cc-pVQZ basis.)

# Supplemental data

New insights from domain-averaged Fermi holes and bond order analysis into the bonding conundrum in  $C_2$

David L. Cooper,<sup>1,\*</sup> Robert Ponec,<sup>2</sup> and Miroslav Kohout<sup>3</sup>

<sup>1</sup> *Department of Chemistry, University of Liverpool, Liverpool L69 7ZD, United Kingdom*

<sup>2</sup> *Institute of Chemical Processes, The Czech Academy of Sciences, Rozvojová 135, 165 02 Prague 6, Czech Republic*

<sup>3</sup> *Max Planck Institute for Chemical Physics of Solids, Nöthnitzer Strasse 40, 01187 Dresden, Germany*

\* **Corresponding Authors:** E-mail: dlc@liv.ac.uk; ponec@icpf.cas.cz

## **Contents**

S-A: Single-configuration of $C_2$ descriptions at $R_e$	S2
S-A.1: Energies and energy differences	S2
S-A.2: Simple model	S2
S-B: CASSCF descriptions at $R=1.2425 \text{ \AA}$	S4
S-C: Geometry dependence of the bonding in the $X^1\Sigma_g^+$ ground state of $C_2$	S8



## ***S-A: Single-configuration descriptions of $C_2$ at $R_e$***

### ***S-A.1. Energies and energy differences***

Various single-configuration descriptions were generated by means of appropriate CASSCF calculations with very limited active spaces. The resulting RHF energies and energy differences ( $R=1.2425 \text{ \AA}$ , cc-pVQZ basis) are reported in Table S1. We notice that the single excitation from  $(\dots)2\sigma_g^2 2\sigma_u^2 1\pi_u^4 \ ^1\Sigma_g^+$  to  $(\dots)2\sigma_g^2 2\sigma_u^1 3\sigma_g^1 1\pi_u^4 \ ^3\Sigma_u^+$  actually corresponds to a significant energy lowering at this low level of theory.

**Table S1.** RHF energies and energy differences for  $C_2$  ( $R=1.2425 \text{ \AA}$ , cc-pVQZ basis).

Configuration	State	$E/\text{hartree}$	$\Delta E/\text{millihartree}$
$1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 2\sigma_u^2 1\pi_u^4$	$^1\Sigma_g^+$	-75.405765	0.0
$1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 2\sigma_u^1 3\sigma_g^1 1\pi_u^4$	$^3\Sigma_u^+$	-75.464875	-59.1
$1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 2\sigma_u^1 3\sigma_g^1 1\pi_u^4$	$^1\Sigma_u^+$	-75.193718	212.0
$1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 3\sigma_g^2 1\pi_u^4$	$^1\Sigma_g^+$	-75.162190	243.6

### ***S-A.2. Simple model***

Suppose that we have the valence configuration  $\sigma_g^2 \sigma_u^2$  and that there are two QTAIM domains, each containing one of the nuclei of a homonuclear diatomic system. The matrix of domain-condensed overlap integrals for the first domain will take the form

$$\begin{pmatrix} 1/2 & |x| \\ |x| & 1/2 \end{pmatrix}$$

so that the corresponding matrix for the second domain must be

$$\begin{pmatrix} 1/2 & -|x| \\ -|x| & 1/2 \end{pmatrix}$$

These two matrices must of course add to a unit matrix. The  $\mathbf{G}$  matrix for the first atomic domain is then

$$\begin{pmatrix} 1 & 2|x| \\ 2|x| & 1 \end{pmatrix}$$

which has eigenvalues  $1+2|x|$  and  $1-2|x|$ . The corresponding DAFH functions are  $(\sigma_g + \sigma_u)/\sqrt{2}$  and  $(\sigma_g - \sigma_u)/\sqrt{2}$ . (These cannot be made any more localised in the isopycnic transformation.)

The eigenvalues are of course the same for the second domain, but the order of the DAFH

functions is reversed. The LNOs, which are doubly occupied, necessarily take the same forms as the DAFH functions.

The values of  $W-M(\sigma)$  and  $CBO(\sigma)$  are the same for this model system, being given by the expression  $2-8x^2$ . (This expression coincides with twice the product of the two DAFH eigenvalues.) Clearly all of the numerical quantities depend only on the magnitude of the domain-condensed overlap integral,  $|x|$ , which can be interpreted as quantifying how similar/different are the atom-centred hybrids from which  $\sigma_g$  and  $\sigma_u$  are constructed.

In the limit that  $|x|=1/2$ , corresponding to  $\sigma_g$  and  $\sigma_u$  orbitals that are built from identical hybrids, only one of the eigenvalues of  $\mathbf{G}$  for a given domain is non-zero, so that there is double occupancy of either  $(\sigma_g + \sigma_u)/\sqrt{2}$  or  $(\sigma_g - \sigma_u)/\sqrt{2}$ . The value of  $W-M(\sigma)$  or  $CBO(\sigma)$  is of course then zero. Looking instead at the hypothetical limit  $|x|=0$ , there would be single occupancy of the functions  $(\sigma_g + \sigma_u)/\sqrt{2}$  and  $(\sigma_g - \sigma_u)/\sqrt{2}$ , and the value of  $W-M(\sigma)$  or  $CBO(\sigma)$  would be 2.

The actual value of the domain-condensed overlap integral between  $2\sigma_g$  and  $2\sigma_u$  is  $|x|=0.3095$  for the closed-shell RHF description of the  $X^1\Sigma_g^+$  state of  $C_2$  (3-21G basis). From the simple model, we would thus predict DAFH eigenvalues of 1.619 and 0.381, with a value for  $W-M(\sigma)$  or  $CBO(\sigma)$  of 1.234. This is precisely what we observe from the actual calculations.

### ***S-B: CASSCF descriptions at $R=1.2425\text{Å}$***

Total energies for the lowest  $^1\Sigma_g^+$ ,  $^3\Sigma_u^+$  and  $^1\Sigma_u^+$  states of  $C_2$  are reported in Table S2 for different CASSCF constructions ( $R=1.2425\text{Å}$ , cc-pVQZ basis). Various numerical results for the ground states of  $Be_2$ ,  $C_2$ ,  $N_2$  and  $HC\equiv CH$  are reported in Table S3 for different CASSCF wavefunctions ( $R=1.2425\text{Å}$ , cc-pVQZ basis).  $HC\equiv CH$  (with  $R_{CH}=1.06\text{Å}$ ) was analyzed as two CH domains.

**Table S2.** Energies and energy differences for the lowest  $^1\Sigma_g^+$ ,  $^3\Sigma_u^+$  and  $^1\Sigma_u^+$  states of  $C_2$  for different CASSCF constructions ( $R=1.2425\text{Å}$ , cc-pVQZ basis.)

	$X^1\Sigma_g^+$	$\underline{c}^3\Sigma_u^+$	$\underline{D}^1\Sigma_u^+$
RHF			
$E/\text{hartree}$	-75.405765	-75.464875	-75.193718
$\Delta E/\text{millihartree}$	0.0	-59.1	212.0
$\sigma$ -only CASSCF			
$E/\text{hartree}$	-75.503108	-75.473560	-75.204349
$\Delta E/\text{millihartree}$	0.0	29.5	298.8
full-valence CASSCF			
$E/\text{hartree}$	-75.643169	-75.600167	-75.408705
$\Delta E/\text{millihartree}$	0.0	43.0	234.5
CASSCF(8,16)			
$E/\text{hartree}$	-75.706245	-75.662188	-75.490547
$\Delta E/\text{millihartree}$	0.0	44.1	215.7

**Table S3.** Various numerical results for ground states of Be<sub>2</sub>, C<sub>2</sub>, N<sub>2</sub> and HC≡CH ( $R=1.2425 \text{ \AA}$ , cc-pVQZ basis.) HC≡CH (with  $R_{\text{CH}}=1.06 \text{ \AA}$ ) was analyzed as two CH domains.

(a)  $\sigma$ -only CASSCF

Molecule	W-M( $\sigma$ )	CBO( $\sigma$ )
Be <sub>2</sub>	1.041	0.896
C <sub>2</sub>	1.048	0.974
N <sub>2</sub>	1.016	1.023
HC≡CH	0.998	1.001

(b) full-valence CASSCF

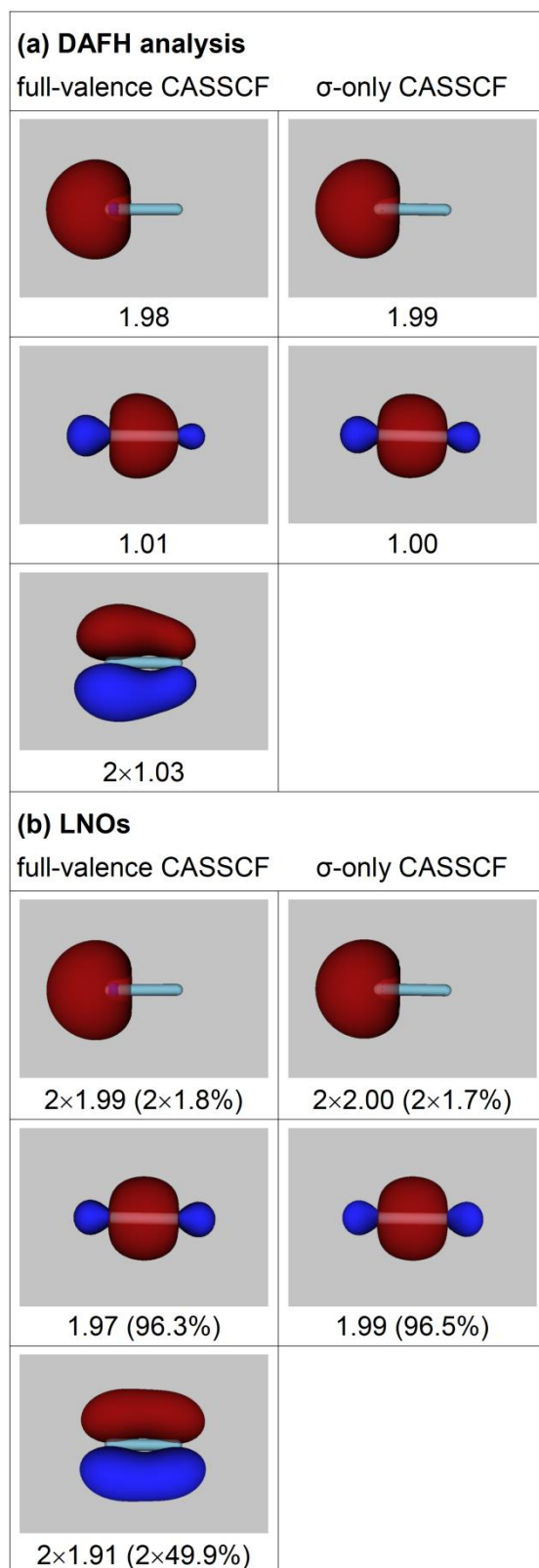
Molecule	W-M	W-M( $\sigma$ )	W-M( $\pi$ )	CBO	CBO( $\sigma$ )	CBO( $\pi$ )	$n(\pi)$
C <sub>2</sub>	2.729	1.070	1.658	2.825	1.030	1.795	4.007
N <sub>2</sub>	2.698	0.993	1.705	2.841	1.009	1.831	4.014
HC≡CH	2.758	0.977	1.781	2.854	0.988	1.866	4.001

(c) CASSCF(8,16)

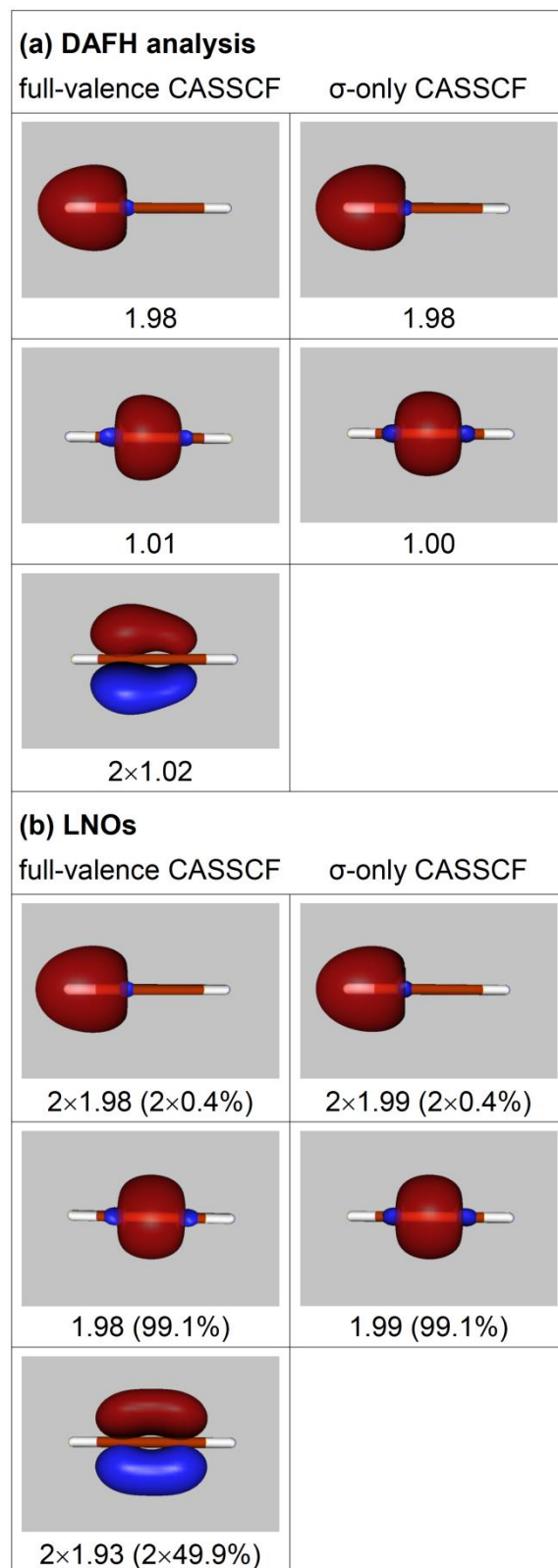
Molecule	W-M	W-M( $\sigma$ )	W-M( $\pi$ )	CBO	CBO( $\sigma$ )	CBO( $\pi$ )	$n(\pi)$
C <sub>2</sub>	2.706	1.054	1.652	2.791	1.014	1.777	4.004

DAFH functions and LNOs for the ground state of N<sub>2</sub> are depicted in Figure S1, together with the corresponding occupation numbers, for different CASSCF constructions. The quantities in brackets signify the relative contributions to the overall CBO( $\sigma$ ) and CBO( $\pi$ ) values. The corresponding results for the ground state of HC≡CH are presented in Figure S2. For both of these molecules, the various results are in accord with intuitive expectations for more or less ordinary triple bonds. The same cannot be said for the comparable calculations for C<sub>2</sub>.

**Figure S1.** DAFH functions and LNOs for  $N_2$ , together with occupation numbers. The quantities in brackets signify the relative contributions to the overall  $CBO(\sigma)$  and  $CBO(\pi)$  values. ( $R=1.2425 \text{ \AA}$ , cc-pVQZ basis.)



**Figure S2.** DAFH functions and LNOs for HC≡CH (analyzed as two CH domains), together with occupation numbers. Quantities in brackets signify relative contributions to CBO( $\sigma$ ) and CBO( $\pi$ ) values. ( $R_{CC}=1.2425 \text{ \AA}$ ,  $R_{CH}=1.06 \text{ \AA}$ , cc-pVQZ basis.)



### ***S-C: Geometry dependence of the bonding in the ground state of C<sub>2</sub>***

CASSCF(8,16) energies (cc-pVQZ basis) for the  $X^1\Sigma_g^+$  ground state of C<sub>2</sub> are reported in Table S4 for a range of geometries from 1 Å to 3.25 Å. Various numerical quantities that were obtained from these wavefunctions are collected in Tables S5 and S6 (see also Figure 5) for geometries from  $R_e$  to 3.25 Å.

**Table S4.** CASSCF(8,16) energies for the  $X^1\Sigma_g^+$  ground state of C<sub>2</sub> (cc-pVQZ basis).

$R / \text{Å}$	$E/\text{hartree}$
1.00	-75.560317
1.05	-75.623742
1.10	-75.665432
1.15	-75.690538
1.20	-75.703083
1.2425	-75.706245
1.30	-75.702420
1.35	-75.693632
1.40	-75.681392
1.45	-75.666942
1.475	-75.659232
1.50	-75.651441
1.55	-75.637282
1.60	-75.625672
1.65	-75.615207
1.70	-75.605035
1.75	-75.595004
1.80	-75.585160
1.85	-75.575586
1.90	-75.566363
2.00	-75.549231
2.10	-75.534141
2.25	-75.515718
2.50	-75.495962
2.75	-75.486551
3.00	-75.482538
3.25	-75.480804

**Table S5.** Various numerical results for the  $X^1\Sigma_g^+$  ground state of  $C_2$  calculated using the CASSCF(8,16) construction (cc-pVQZ basis).

$R/\text{\AA}$	W-M	W-M( $\sigma$ )	W-M( $\pi$ )	CBO	CBO( $\sigma$ )	CBO( $\pi$ )	$n(\pi)$
1.2425	2.706	1.054	1.652	2.791	1.014	1.777	4.004
1.30	2.645	1.039	1.607	2.749	0.999	1.750	3.999
1.35	2.582	1.019	1.563	2.700	0.978	1.723	3.992
1.40	2.508	0.993	1.515	2.639	0.949	1.690	3.979
1.45	2.416	0.961	1.456	2.557	0.913	1.644	3.951
1.475	2.356	0.940	1.415	2.494	0.888	1.607	3.918
1.50	2.252	0.912	1.340	2.360	0.842	1.518	3.819
1.55	1.889	0.853	1.036	1.728	0.656	1.072	3.219
1.60	1.667	0.857	0.811	1.480	0.743	0.737	2.678
1.65	1.577	0.871	0.706	1.430	0.825	0.604	2.438
1.70	1.524	0.877	0.647	1.410	0.865	0.545	2.331
1.75	1.480	0.877	0.603	1.394	0.884	0.510	2.274
1.80	1.437	0.873	0.565	1.379	0.893	0.487	2.241
1.85	1.394	0.864	0.529	1.363	0.895	0.468	2.218
1.90	1.347	0.852	0.495	1.345	0.894	0.452	2.203
2.00	1.246	0.820	0.426	1.030	0.882	0.147	2.181
2.10	1.132	0.776	0.356	0.985	0.861	0.124	2.168
2.25	0.937	0.685	0.252	0.903	0.815	0.088	2.153
2.50	0.597	0.485	0.112	0.297	0.258	0.038	2.138
2.75	0.341	0.296	0.044	0.176	0.162	0.014	2.127
3.00	0.194	0.174	0.020	0.103	0.097	0.006	2.120
3.25	0.114	0.103	0.010	0.061	0.058	0.003	2.116



**Table S6.** Values of  $\tilde{n}_p$  (see Equation (5)) for  $\text{C}_2$  ( $X^1\Sigma_g^+$ ) (CASSCF(8,16)/cc-pVQZ).

$R/\text{\AA}$	first $\sigma$	'bond-like' $\sigma$	first $\pi$
1.2425	1.458	2.000	2.000
1.30	1.439	2.000	2.000
1.35	1.419	2.000	2.000
1.40	1.395	2.000	2.000
1.45	1.365	2.000	2.000
1.475	1.346	2.000	2.000
1.50	1.309	2.000	2.000
1.55	1.108	2.000	2.000
1.60	1.026	2.000	2.000
1.65	1.014	2.000	2.000
1.70	1.012	2.000	2.000
1.75	1.011	2.000	2.000
1.80	1.011	2.000	2.000
1.85	1.011	2.000	2.000
1.90	1.011	2.000	2.000
2.00	1.011	2.000	1.341
2.10	1.011	2.000	1.275
2.25	1.010	2.000	1.183
2.50	1.010	1.298	1.074
2.75	1.008	1.157	1.026
3.00	1.006	1.084	1.011
3.25	1.004	1.047	1.005