# Local spins: improved Hilbert-space analysis 

Eloy Ramos-Cordoba, ${ }^{a}$ Eduard Matito, ${ }^{a}$ Pedro Salvador* ${ }^{a}$ and István Mayer ${ }^{b}$

Received 29th May 2012, Accepted 17th September 2012
DOI: 10.1039/c2cp42513k

The decomposition of $\left\langle\hat{S}^{2}\right\rangle$ for a general wave function has been carried out in the framework of the Hilbert-space analysis. The one and two-center components fulfill all physical requirements imposed to date. An inherent ambiguity of the Hilbert-space decomposition of a two-electron quantity, in particular using a Mulliken-type scheme, is also discussed in detail. The formalism of effective atomic densities has allowed us to derive in a simple manner appropriate expressions for the decomposition of $\left\langle\hat{S}^{2}\right\rangle$ in the framework of Hilbert space analysis that are consistent with Mulliken population analysis and related quantities. Using a particular mapping we have derived the Hilbert-space expressions also in the framework of Löwdin population analysis in a straightforward manner. The numerical results obtained with the latter formalism have proved to be more robust and reliable.

## 1 Introduction

The concept of local spin emerges in a quite natural fashion when describing the electronic structure of systems with diradical character such as non-Kekulé molecules or transition state structures of chemical reactions. Heisenberg Hamiltonian models also invoke the concept of local spin in order to assess the nature of spin-spin interactions between magnetic centers. Often, the spin properties of a molecule can be characterized by the spin density. There are, however, cases where the overall system is a singlet (where there is no spin density), but for which the existence of some local spin is assumed. In the last few years there has been a growing interest in recovering local spins from the analysis of the wave function of ab initio calculations. ${ }^{1-15}$ Different schemes have been proposed in the literature, most of which are rooted in the decomposition of the expectation value of the spin squared operator into atomic and diatomic contributions, for both single-determinant and correlated wave functions. Because the partitioning of the single physical quantity $\left\langle\hat{S}^{2}\right\rangle$, which in the case of singlet wave functions is zero, into components is not unique, a number of physical requirements ${ }^{4,9,15}$ have been introduced.
(i) One should get no spins whatever for covalent systems described by a closed-shell RHF wave function using doublyfilled orbitals.
(ii) If the wave function is properly dissociating, then the asymptotic values of the atomic spins obtained for the atoms

[^0]at large distances should coincide with the corresponding values of the free atoms.
(iii) In an open-shell system the overall $\left\langle\hat{S}^{2}\right\rangle$ does not depend on the actual $\hat{S}_{z}$ projection of the electronic state (multiplet) considered, so one may request to have $\left\langle\hat{S}^{2}\right\rangle$ components that do not depend on $\hat{S}_{z}$ either.
(iv) No two-center terms should appear in the case of singleelectron systems (or ROHF systems with a single unpaired electron).

In a previous paper ${ }^{15}$ we showed that the following general expression

$$
\begin{align*}
\left\langle\hat{S}^{2}\right\rangle= & \frac{3}{4} \int u\left(\vec{r}_{1}\right) \mathrm{d} \vec{r}_{1}+\frac{1}{2} \iint\left[\Gamma\left(\vec{r}_{1}, \vec{r}_{2}\right)-\frac{1}{2} \rho^{\mathrm{s}}\left(\vec{r}_{1} ; \vec{r}_{2}\right) \rho^{\mathrm{s}}\left(\vec{r}_{2} ; \vec{r}_{1}\right)\right] \mathrm{d} \vec{r}_{1} \mathrm{~d} \vec{r}_{2} \\
& -\frac{1}{2} \iint\left[\Gamma\left(\vec{r}_{1}, \vec{r}_{2} ; \vec{r}_{2}, \vec{r}_{1}\right)-\frac{1}{2} \rho^{\mathrm{s}}\left(\vec{r}_{1} ; \vec{r}_{1}\right) \rho^{\mathrm{s}}\left(\vec{r}_{2} ; \vec{r}_{2}\right)\right] \mathrm{d} \vec{r}_{1} \mathrm{~d} \vec{r}_{2} \tag{1}
\end{align*}
$$

is the natural starting point to derive atomic and diatomic components of $\left\langle\hat{S}^{2}\right\rangle$ that satisfy requirements (i) to (iv). This equation is written in terms of the density of effectively unpaired electrons, $u(\vec{r})$, defined by Takatsuka et al. ${ }^{16}$ as

$$
\begin{equation*}
u(\vec{r})=2 \rho(\vec{r})-\int \rho\left(\vec{r} ; \vec{r}^{\prime}\right) \rho\left(\vec{r}^{\prime} ; \vec{r}\right) \mathrm{d} \vec{r}^{\prime} \tag{2}
\end{equation*}
$$

the spin-density matrix

$$
\begin{equation*}
\rho^{\mathrm{s}}\left(\vec{r} ; \vec{r}^{\prime}\right)=\rho^{\alpha}\left(\vec{r} ; \vec{r}^{\prime}\right)-\rho^{\beta}\left(\vec{r} ; \vec{r}^{\prime}\right), \tag{3}
\end{equation*}
$$

and the spin-less cumulant of the second order density matrix, $\Gamma\left(\vec{r}_{1}, \vec{r}_{2} ; \vec{r}_{1}{ }_{1}, \vec{r}^{\prime}{ }_{2}\right)$, which vanishes for single-determinant wave functions and can be defined as the sum of the usual (spindependent) cumulants as

$$
\begin{equation*}
\Gamma\left(\vec{r}_{1}, \vec{r}_{2} ; \vec{r}_{1}^{\prime}, \vec{r}_{2}^{\prime}\right)=\sum_{\sigma, \sigma^{\prime}} \Gamma^{\sigma \sigma^{\prime} \sigma \sigma^{\prime}}\left(\vec{r}_{1}, \vec{r}_{2} ; \vec{r}_{1}^{\prime}, \vec{r}_{2}^{\prime}\right) \tag{4}
\end{equation*}
$$

where

$$
\begin{align*}
\Gamma^{\sigma \sigma^{\prime} \sigma \sigma^{\prime}}\left(\vec{r}_{1}, \vec{r}_{2} ; \vec{r}_{1}^{\prime}, \vec{r}_{2}^{\prime}\right)= & \rho_{2}^{\sigma \sigma^{\prime} \sigma \sigma^{\prime}}\left(\vec{r}_{1}, \vec{r}_{2} ; \vec{r}_{1}^{\prime}, \vec{r}_{2}^{\prime}\right) \\
& -\rho^{\sigma}\left(\vec{r}_{1}, \vec{r}_{1}^{\prime}\right) \rho^{\sigma^{\prime}}\left(\vec{r}_{2}, \vec{r}_{2}^{\prime}\right)  \tag{5}\\
& -\delta_{\sigma \sigma^{\prime}}\left[\rho^{\sigma}\left(\vec{r}_{1}, \vec{r}_{2}^{\prime}\right) \rho^{\sigma}\left(\vec{r}_{2}, \vec{r}_{1}^{\prime}\right)\right] .
\end{align*}
$$

In ref. 15 we obtained one and two-center contributions for a general wave function in the framework of the 3D-space analysis, i.e., for "fuzzy atoms" ${ }^{17}$ and Bader's atomic domains. ${ }^{18}$

In this paper we wish to undertake the decomposition in the context of the so-called "Hilbert-space analysis". ${ }^{19}$ The motivation is twofold: first, aside its conceptual relevance, the Hilbert-space decomposition does not require atomic numerical integrations, in contrast to the 3D-space analysis; thus it is exact (is free of the numerical errors of that integration). Also, the significant reduction in the computational cost of the decomposition may be relevant for very large systems, especially as compared to the 3D-space methods with complicated atomic basins of Bader's analysis. However, there is an apparent ambiguity in decomposing two-electron quantities in the framework of Hilbert-space analysis (in particular for the Mulliken-type scheme), which to date has not received due attention. In this paper we also wish to analyze in more detail this problem, which is particularly relevant in the case of the decomposition of $\left\langle\hat{S}^{2}\right\rangle$. That ambiguity will be exposed in the next section. Then, we will briefly describe the formalism of effective atomic densities, ${ }^{20}$ which will allow us to derive in a straightforward manner the most appropriate expressions for the decomposition of $\left\langle\hat{S}^{2}\right\rangle$ in the framework of Hilbert space analysis. Finally, some numerical results at the correlated level will be presented and discussed.

## 2 Alternative summation schemes in the Hilbert-space analysis

The decomposition of physical quantities into atomic and diatomic contributions is rooted on the identification of an atom within the molecule. Since practical quantum chemistry mostly uses atom-centered basis sets, the atom may be identified with its nucleus and the subspace spawned by the set of atomic basis functions centered on it. The simplest example of application of such Hilbert-space analysis is Mulliken population analysis, ${ }^{21}$ perhaps the most familiar method to determine the number of electrons associated with an atom. Mulliken's gross population of atom A is defined as:

$$
\begin{equation*}
N_{\mathrm{A}}=\sum_{\mu \in \mathrm{A}} \sum_{\nu} D_{\mu \nu} S_{\nu \mu}=\sum_{\mu \in \mathrm{A}}(\boldsymbol{D} \boldsymbol{S})_{\mu \mu}, \tag{6}
\end{equation*}
$$

where the notation $\mu \in \mathrm{A}$ indicates that the summation runs over all atomic basis functions centered on atom A. We recall in this context that matrix $\boldsymbol{D S}$ is the proper finite basis representation of the first-order density matrix if an overlapping basis set $(\boldsymbol{S} \neq \boldsymbol{I})$ is used. ${ }^{22}$

In a similar manner, the Mayer-Wiberg (closed-shell for simplicity) bond order, ${ }^{23} B_{\mathrm{AB}}$, between atoms A and B is defined as

$$
\begin{equation*}
B_{\mathrm{AB}}=\sum_{\mu \in \mathrm{A}} \sum_{\sigma \in \mathrm{B}}(\boldsymbol{D} \boldsymbol{S})_{\mu \sigma}(\boldsymbol{D} \boldsymbol{S})_{\sigma \mu} \tag{7}
\end{equation*}
$$

Inspecting the expression in eqn (6), one can see that the overlap integrals enter it in a somewhat non-symmetric manner: one of the subscripts $(\mu)$ is serving for subdividing the quantity into atomic contributions, while another $(\nu)$ is a "dummy" index, for which summation over the whole basis is performed-it is used to form the matrix-product $\boldsymbol{D S}$. This difference may be connected with the fact that for overlapping basis sets matrix $\boldsymbol{D S}$ is twice the projection matrix performing the projection of any vector $\boldsymbol{d}$ of LCAO coefficients on the subspace of the occupied molecular orbitals as $\boldsymbol{D S S} \boldsymbol{r}^{22}$ The same distinction appears also in eqn (7) of the Mayer-Wiberg bond order. In the case of real orbitals, one could get exactly the same Mulliken atomic populations also in the form $\sum_{\mu \in \mathrm{A}}(\boldsymbol{S D})_{\mu \mu}$, i.e., by using matrix $\boldsymbol{S D}$ which performs the analogous projection of the row-vectors $\boldsymbol{d}^{\dagger}$ as $\boldsymbol{d}^{\dagger} \boldsymbol{S D}$. While in the first case the systematization of the terms according to the individual atoms corresponds to the subscript coming from the "ket" part of the overlap integral, in the second one it corresponds to the subscript coming from its "bra" part. It seems logical to stick to one of these possibilities (we prefer the first one), and use it in all types of analyses. Thus the splitting of the terms in the expression of the bond order index eqn (7) corresponds to the subscripts of the overlap integrals coming from the "kets".

In principle, if an expression contains products with two overlap matrices, then a subdivision into atomic and diatomic contributions by taking one subscript from "bra" and another from "ket" is also possible. In the case of the bond order, that leads to a modified definition of the bond order index ${ }^{24}$ as

$$
\begin{equation*}
B_{\mathrm{AB}}^{\prime}=\sum_{\mu \in \mathrm{A}} \sum_{\nu \in \mathrm{B}}(\boldsymbol{S} \boldsymbol{D} \boldsymbol{S})_{\mu \nu} \boldsymbol{D}_{\nu \mu} \tag{8}
\end{equation*}
$$

As the bond order is a component of the integral of the exchange density, formally both definitions could be acceptable: they represent different decompositions of that integral into a sum of one- and two-center contributions. However, there is a serious argument favoring the definition of eqn (7). The modified, non-symmetric, definition eqn (8) gives results that are much less "chemical" than those given by the original one: it cannot, for instance, recover the integer values for firstrow diatomics (e.g., 3 for $\mathrm{N}_{2}$ ) if a minimal basis set is used, as does the original definition of eqn (7). Another argument against such type of "bra"-"ket" mixing is the high degree or arbitrariness that would be introduced in the case of e.g. Generalized Population Analysis, ${ }^{25}$ typically used to detect patterns of multicenter bonding, where the expressions may contain three, four or more overlap matrices. Furthermore, it has been shown ${ }^{26}$ that one can introduce a particular mapping (see Appendix) between the atomic overlap matrices of the atomic orbitals and the conventional overlap matrix that permits finding a one-to-one correspondence between the Hilbert-space and the more general 3D-space analyses expressions of quantities like bond orders, atomic valences or energy components. Such a general mapping is not possible for expressions involving subdivision of the terms according to both "bra" and "ket" subscripts.

These considerations are of interest in the present context because in the recent paper by Alcoba et al. ${ }^{9}$ a decomposition
of $\left\langle\hat{S}^{2}\right\rangle$ is performed in a manner that one index of the overlap matrix is assigned according to the term coming from the "bra" and another coming from the "ket". (These authors distinguish between them by using both subscripts and superscripts, which, however, do not represent covariant and contravariant indices.) Therefore, their decomposition is consistent only with the use of the alternative bond order formula eqn (8). $\dagger$
In the next section we will briefly describe the formalism of effective atomic densities, ${ }^{20}$ which will allow us to derive in a straightforward manner the most appropriate expressions for the decomposition of $\left\langle\hat{S}^{2}\right\rangle$ in the framework of Hilbert space analysis.

## 3 Effective atomic density matrices formalism

The formalism of the atomic and diatomic effective densities is based on the exact decomposition of one- and two-electron densities into components that can be considered their onecenter and one- and two-center contributions, respectively. These atomic and diatomic densities are identified with the contributions of each atom and pairs of atoms to the overall density, and can be used to derive in a common framework atomic populations, bond orders, atomic valences, molecular energy components, etc. for any kind of atom in molecule definition. In the simplest case, one can define the effective atomic contributions to the electron density, $\rho_{\mathrm{A}}(\vec{r})$, simply fulfilling $\rho(\vec{r}) \equiv \sum_{\mathrm{A}}^{N} \rho_{\mathrm{A}}(\vec{r})$. The integral over the whole space of this function for atom A quite naturally yields the electron population associated with the atom

$$
\begin{equation*}
\int \rho_{\mathrm{A}}(\vec{r}) \mathrm{d} \vec{r}=N_{\mathrm{A}} \tag{9}
\end{equation*}
$$

The actual numerical value depends upon how $\rho_{\mathrm{A}}(\vec{r})$ is defined. In the framework of 3D-space analysis, $\rho_{\mathrm{A}}(\vec{r})$ can be written in general as

$$
\begin{equation*}
\rho_{\mathrm{A}}(\vec{r})=w_{\mathrm{A}}(\vec{r}) \rho(\vec{r}) \tag{10}
\end{equation*}
$$

where $w_{\mathrm{A}}$ is a non-negative weight function defined for each atom and each point of the 3D space satisfying $\sum_{\mathrm{A}} w_{\mathrm{A}}(\vec{r})=1$. The actual definition of atom in the molecule ("fuzzy" or disjoint) is contained in the atomic weight functions.

In the case of Hilbert-space analysis, the effective atomic density can be most suitable written in terms of the matrix elements of the LCAO density matrix as

$$
\begin{equation*}
\rho_{\mathrm{A}}(\vec{r})=\sum_{\mu \in \mathrm{A}} \sum_{\nu} D_{\mu \nu} \chi_{\nu}^{*}(\vec{r}) \chi_{\mu}(\vec{r}) . \tag{11}
\end{equation*}
$$

It is trivial to see that the integration of eqn (11) yields Mulliken's gross population of atom A, in accord with eqn (6).

In a similar manner, by combining the appropriate effective atomic contributions of the first-order density matrix to build effective diatomic exchange densities:

$$
\begin{equation*}
\rho_{x}^{\mathrm{AB}}\left(\vec{r} ; \vec{r}^{\prime}\right)=\frac{1}{2}\left[\rho_{\mathrm{A}}\left(\vec{r} ; \vec{r}^{\prime}\right) \rho_{\mathrm{B}}\left(\vec{r}^{\prime} ; \vec{r}\right)+\rho_{\mathrm{B}}\left(\vec{r} ; \vec{r}^{\prime}\right) \rho_{\mathrm{A}}\left(\vec{r}^{\prime} ; \vec{r}\right)\right] \tag{12}
\end{equation*}
$$

[^1]where
\[

$$
\begin{equation*}
\rho_{\mathrm{A}}\left(\vec{r}, \vec{r}^{\prime}\right)=\sum_{\mu \in \mathrm{A}} \sum_{\nu} D_{\mu \nu} \chi_{\nu}^{*}(\vec{r}) \chi_{\mu}\left(\vec{r}^{\prime}\right) \tag{13}
\end{equation*}
$$

\]

is the atomic component of the first order spin-less density matrix $\rho\left(\vec{r}, \vec{r}^{\prime}\right)$. One can easily recover upon integration of eqn (12) the expression of eqn (7) for the Mayer-Wiberg (closed-shell) bond order

$$
\begin{align*}
\iint \rho_{x}^{\mathrm{AB}}\left(\vec{r}, \vec{r}^{\prime}\right) \mathrm{d} \vec{r} \mathrm{~d} \vec{r}^{\prime}= & \frac{1}{2} \sum_{\mu \in \mathrm{A}} \sum_{\nu} D_{\mu \nu} S_{\nu \sigma} \sum_{\sigma \in \mathrm{B}} \sum_{\lambda} D_{\sigma \lambda} S_{\lambda \mu} \\
& +\frac{1}{2} \sum_{\mu \in \mathrm{B}} \sum_{\nu} D_{\mu \nu} S_{\nu \sigma} \sum_{\sigma \in \mathrm{A}} \sum_{\lambda} D_{\sigma \lambda} S_{\lambda \mu} \\
= & \sum_{\mu \in \mathrm{A}} \sum_{\sigma \in \mathrm{B}}(\boldsymbol{D} \boldsymbol{S})_{\mu \sigma}(\boldsymbol{D} \boldsymbol{S})_{\sigma \mu}=B_{\mathrm{AB}} \tag{14}
\end{align*}
$$

and so forth.
One advantage of using the formalism of effective atomic densities is that one can switch from 3D-space to Hilbert-space formulae or vice versa simply by taking the appropriate form of the effective densities involved in the calculation. In a recent paper ${ }^{15}$ we have put forward an improved general formula for the decomposition of $\left\langle\hat{S}^{2}\right\rangle$ applicable for both single-determinant and correlated wave functions. The numerical implementation of the resulting one- and two-center components was originally carried out in the 3D-physical space. Here we will make use of the formalism of the atomic and diatomic effective matrices depicted above to derive in a simple manner the appropriate one and two-center components of $\left\langle\hat{S}^{2}\right\rangle$ in the framework of Hilbert-space analysis.
This exercise is of particular interest here because of the formal ambiguity affecting Hilbert-space decompositions in the selection of the indices put forward in the previous section. We most definitely recommend to stick to the assignment of subscripts that will be obtained here, which is consistent with both Mulliken population analysis and the original bond order definition eqn (7).

## 4 Decomposition of $\left\langle\tilde{S}^{\mathbf{2}}\right\rangle$

Within the formalism of the effective atomic densities, the respective one- and two-center contributions to $\left\langle\hat{S}^{2}\right\rangle$ can be formally written from the general expression of eqn (1) simply as

$$
\begin{align*}
\left\langle\hat{S}^{2}\right\rangle_{\mathrm{A}}= & \frac{3}{4} \int u_{\mathrm{A}}\left(\vec{r}_{1}\right) \mathrm{d} \vec{r}_{1}+\frac{1}{2} \iint\left[\Gamma_{\mathrm{AA}}\left(\vec{r}_{1}, \vec{r}_{2}\right)\right. \\
& \left.-\frac{1}{2} \rho_{\mathrm{A}}^{\mathrm{s}}\left(\vec{r}_{1} ; \vec{r}_{2}\right) \rho_{\mathrm{A}}^{\mathrm{s}}\left(\vec{r}_{2} ; \vec{r}_{1}\right)\right] \mathrm{d} \vec{r}_{1} \mathrm{~d} \vec{r}_{2}-\frac{1}{2} \iint\left[\Gamma_{\mathrm{AA}}\left(\vec{r}_{1}, \vec{r}_{2} ; \vec{r}_{2}, \vec{r}_{1}\right)\right. \\
& \left.-\frac{1}{2} \rho_{\mathrm{A}}^{\mathrm{s}}\left(\vec{r}_{1} ; \vec{r}_{1}\right) \rho_{\mathrm{A}}^{\mathrm{s}}\left(\vec{r}_{2} ; \vec{r}_{2}\right)\right] \mathrm{d} \vec{r}_{1} \mathrm{~d} \vec{r}_{2} \tag{15}
\end{align*}
$$

and

$$
\begin{align*}
\left\langle\hat{S}^{2}\right\rangle_{\mathrm{AB}}= & \frac{1}{2} \iint\left[\Gamma_{\mathrm{AB}}\left(\vec{r}_{1}, \vec{r}_{2}\right)-\frac{1}{2} \rho_{\mathrm{A}}^{\mathrm{s}}\left(\vec{r}_{1} ; \vec{r}_{2}\right) \rho_{\mathrm{B}}^{\mathrm{s}}\left(\vec{r}_{2} ; \vec{r}_{1}\right) \mathrm{d} \vec{r}_{1} \mathrm{~d} \vec{r}_{2}\right. \\
& -\frac{1}{2} \iint\left[\Gamma_{\mathrm{AB}}\left(\vec{r}_{1}, \vec{r}_{2} ; \vec{r}_{2}, \vec{r}_{1}\right)-\frac{1}{2} \rho_{\mathrm{A}}^{\mathrm{s}}\left(\vec{r}_{1} ; \vec{r}_{1}\right) \rho_{\mathrm{B}}^{\mathrm{s}}\left(\vec{r}_{2} ; \vec{r}_{2}\right)\right] \mathrm{d} \vec{r}_{1} \mathrm{~d} \vec{r}_{2}, \tag{16}
\end{align*}
$$

where the atomic (in the case of $u(\vec{r})$ and $\rho^{\mathrm{s}}\left(\vec{r}_{1}, \vec{r}^{\prime}\right)$ ) and diatomic (in the case of the cumulants, $\Gamma$ ) densities have been conveniently used, instead of their global counterparts in eqn (1).

In order to be consistent with the definitions of eqn (11) in the framework of Hilbert-space analysis, the effective atomic contributions of the density of effectively unpaired electrons and the spin density matrix must be taken as

$$
\begin{equation*}
\rho_{\mathrm{A}}^{\mathrm{s}}\left(\vec{r} ; \vec{r}^{\prime}\right)=\sum_{\mu \in \mathrm{A}} \sum_{\nu} P_{\mu \nu}^{\mathrm{s}} \chi_{\nu}^{*}(\vec{r}) \chi_{\mu}\left(\vec{r}^{\prime}\right) \tag{17}
\end{equation*}
$$

and

$$
\begin{equation*}
u_{\mathrm{A}}(\vec{r})=\sum_{\mu \in \mathrm{A}} \sum_{\nu}\left[2 D_{\mu \nu}-(\boldsymbol{D S D})_{\mu \nu}\right] \chi_{\nu}^{*}(\vec{r}) \chi_{\mu}(\vec{r}), \tag{18}
\end{equation*}
$$

where we have made use of eqn (2).
The spin-less cumulant, $\Gamma$, being a genuine two-electron quantity, consists of atomic (if $\mathrm{A}=\mathrm{B})$ and diatomic $(\mathrm{A} \neq \mathrm{B})$ contributions:

$$
\begin{equation*}
\Gamma_{\mathrm{AB}}\left(\vec{r}_{1}, \vec{r}_{2} ; \vec{r}_{1}^{\prime}, \vec{r}_{2}^{\prime}\right)=\sum_{\mu \in \mathrm{A}} \sum_{\sigma \in \mathrm{B}} \sum_{\nu, \lambda} \Gamma_{\mu \sigma \nu \lambda} \chi_{\nu}^{*}\left(\vec{r}_{1}\right) \chi_{\lambda}^{*}\left(\vec{r}_{2}\right) \chi_{\mu}\left(\vec{r}_{1}^{\prime}\right) \chi_{\sigma}\left(\vec{r}_{2}^{\prime}\right) \tag{19}
\end{equation*}
$$

where $\Gamma_{\mu \sigma \nu \lambda}$ are the corresponding matrix elements of the cumulant in the atomic orbital basis.

Substituting eqn (17)-(19) into (15) and integrating, one obtains, after some manipulations, the final expression for the atomic components of $\left\langle\hat{S}^{2}\right\rangle$

$$
\begin{align*}
\left\langle\hat{\boldsymbol{S}}^{2}\right\rangle_{\mathrm{A}}= & \frac{3}{4} \sum_{\mu \in \mathrm{A}}\left[2(\boldsymbol{D} \boldsymbol{S})_{\mu \mu}-(\boldsymbol{D} \boldsymbol{S} \boldsymbol{D} \boldsymbol{S})_{\mu \mu}\right]-\frac{1}{4} \sum_{\mu, \nu \in \mathrm{A}}\left(\boldsymbol{P}^{\mathrm{s}} \boldsymbol{S}\right)_{\mu \nu}\left(\boldsymbol{P}^{\mathrm{s}} \boldsymbol{S}\right)_{\nu \mu} \\
& +\frac{1}{4} \sum_{\mu, \nu \in \mathrm{A}}\left(\boldsymbol{P}^{\mathbf{s}} \boldsymbol{S}\right)_{\mu \mu}\left(\boldsymbol{P}^{\mathbf{s}} \boldsymbol{S}\right)_{\nu \nu} \\
& +\frac{1}{2} \sum_{\mu, \sigma \in \mathrm{A}} \sum_{\nu, \lambda}\left(\Gamma_{\mu \sigma \nu \lambda}-\Gamma_{\mu \sigma \lambda \nu}\right) S_{\lambda \sigma} S_{\nu \mu} . \tag{20}
\end{align*}
$$

Similarly, for the diatomic spin components one gets

$$
\begin{align*}
\left\langle\hat{S}^{2}\right\rangle_{\mathrm{AB}}= & -\frac{1}{4} \sum_{\mu \in \mathrm{A}} \sum_{\nu \in \mathrm{B}}\left(\boldsymbol{P}^{\mathrm{s}} \boldsymbol{S}\right)_{\mu \nu}\left(\boldsymbol{P}^{\mathrm{s}} \boldsymbol{S}\right)_{\nu \mu} \\
& +\frac{1}{4} \sum_{\mu \in \mathrm{A}} \sum_{\nu \in \mathrm{B}}\left(\boldsymbol{P}^{\mathrm{s}} \boldsymbol{S}\right)_{\mu \mu}\left(\boldsymbol{P}^{\mathrm{s}} \boldsymbol{S}\right)_{\nu \nu}  \tag{21}\\
& +\frac{1}{2} \sum_{\mu \in \mathrm{A}} \sum_{\sigma \in \mathrm{B}} \sum_{\nu, \lambda}\left(\Gamma_{\mu \sigma \nu \lambda}-\Gamma_{\mu \sigma \lambda \nu}\right) S_{\lambda \sigma} \boldsymbol{S}_{\nu \mu}
\end{align*}
$$

In the single-determinant case the cumulants vanish and these formulae reduce to those derived independently in ref. 27.

Eqn (13) and (14) of Alcoba et al. ${ }^{9}$ are similar to our eqn (20) and (21). In that paper, however, aside from the fact that the authors started from a formula different from eqn (1) - therefore the coefficients of the different terms of eqn (20) and (21) are different - the authors also chose a different convention in the treatment of the cumulant part: one index of the overlap matrix is assigned according to the term coming from the "bra" and another coming from the "ket". From now on we will refer to that different convention as formula with "interchanged" indices.

Hilbert-space analysis is not restricted to Mulliken's recipe. Another alternative scheme is Löwdin population analysis, ${ }^{28}$
in which the atomic orbitals are first transformed to an orthogonal basis. Even though it is less often used, Löwdin analysis typically exhibits less basis set effects than Mulliken's. Indeed, it is well-known that Mulliken-based analyses can yield meaningless results if combined with diffuse functions lacking marked atomic character. ${ }^{29}$ In the Löwdin basis the overlap matrix is a unit matrix and, as a consequence, the schemes with conventional and "interchanged" indices are equivalent.

## 5 Numerical results

We have written a program that performs the decomposition of $\left\langle\hat{S}^{2}\right\rangle$ described above in the framework of the Hilbert-space analysis for both Mulliken and Löwdin schemes, using the appropriate effective atomic overlap matrices outlined in the Appendix. Since Löwdin analysis is not strictly rotational invariant ${ }^{30}$ with Cartesian 6 d atomic orbitals, we recommend its use only with pure 5d orbitals. We have included results with 6 d functions only for comparison purposes. The first- and second-order density matrices have been obtained using a modified version of Gaussian-03 program suite ${ }^{31}$ and an auxiliary program ${ }^{32}$ that reads and processes CISD and CASSCF outputs. All calculations have been carried out with the geometrical structure of the molecules optimized at the current level of theory unless otherwise stated.

To assess the numerical effect on the use of the different summation schemes in the Hilbert-space analysis, we have studied the $\mathrm{H}_{2}$ molecule at the $\operatorname{CASSCF}(2,4)$ level for several basis sets. Table 1 gathers the local spin values on the H atom for the conventional $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{H}}$ and "interchanged" index conventions $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{H}}^{\text {int }}$ within Mulliken's scheme. The values for Löwdin $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{H}}^{\mathrm{L}}$ and 3 D -space analysis $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{H}}^{3 \mathrm{D}}$ (using Becke atoms) are also included for comparison.

The local spin values using eqn (20) are close to zero in all cases, in line with the physical expectations. The numbers exhibit reasonably small basis set dependence and are also very similar to those obtained in the framework of 3D-space analysis. Using the alternative formula with "interchanged" indices the values are somewhat too large, as compared with the "conventional" ones, and suffer from strong basis set effects, especially when combining two sets of diffuse functions and Cartesian 6d 10 f orbitals. With this extended basis set

Table 1 Atomic local spin values calculated at the $\operatorname{CASSCF}(2,4)$ level for the $\mathrm{H}_{2}$ molecule at interatomic distance $R_{\mathrm{H}-\mathrm{H}}=0.746 \AA$ for several basis sets

| Basis set | $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{H}}$ | $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{H}}^{\text {int }}$ | $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{H}}^{\mathrm{L}}$ | $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{H}}^{3 \mathrm{D}}$ |
| :--- | :---: | :---: | :---: | :---: |
| cc-pVDZ | 0.036 | 0.060 | 0.028 | 0.034 |
| cc-pVTZ | 0.035 | 0.051 | 0.025 | 0.034 |
| cc-pVQZ | 0.039 | 0.066 | 0.027 | 0.036 |
| cc-pVTZ (6d) | 0.035 | 0.052 | 0.027 | 0.034 |
| cc-pVQZ (6d 10f) | 0.043 | 0.089 | 0.028 | 0.036 |
| aug-cc-pVDZ | 0.038 | 0.069 | 0.025 | 0.034 |
| aug-cc-pVTZ | 0.020 | 0.031 | 0.026 | 0.036 |
| aug-cc-pVQZ | 0.035 | 0.044 | 0.026 | 0.036 |
| d-aug-cc-pVQZ | 0.045 | 0.130 | 0.024 | 0.034 |
| aug-cc-pVTZ (6d) | 0.046 | 0.115 | 0.027 | 0.036 |
| aug-cc-pVQZ (6d 10f) | 0.055 | 0.192 | 0.026 | 0.036 |
| d-aug-cc-pVQZ (6d 10f) | 0.095 | 0.891 | 0.023 | 0.034 |

even the $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{H}}$ value $(0.095)$ is considerably larger than the rest. The $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{H}}^{\mathrm{L}}$ and $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{H}}^{3 \mathrm{D}}$ values show virtually no basis set dependency. However, it seems that for the smaller basis sets Mulliken's values are closer to the 3D-space ones than Löwdin's.

The recommended Mulliken-type decomposition has also been applied to a series of singlet molecules and the results are presented in Table 2. The optimized geometries and the wavefunctions were obtained at the CISD $/ 6-31 \mathrm{G}^{* *}$ level of theory (with Cartesian 6 d functions). Note that for these systems the overall $\left\langle\hat{S}^{2}\right\rangle$ value is zero, but small local atomic spins can be induced by correlation fluctuations. One should only expect the presence of significant diatomic contributions in singlet systems if there would be any antiferromagnetic coupling that could be distinguished from covalent bonding.

As anticipated, the molecules of the $\mathrm{H}_{n} \mathrm{X}$ series $\left(\mathrm{HF}, \mathrm{H}_{2} \mathrm{O}\right.$, and $\mathrm{NH}_{3}$ ) show small values of local spin. The only systems with atomic spin contributions larger than 0.1 are homonuclear diatomic $\mathrm{Li}_{2}$ and $\mathrm{Be}_{2}$, with $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{Li}}=0.156$ and $\left\langle\hat{S}^{2}\right\rangle_{\text {Be }}=0.175$, respectively. These values are consistent with those obtained within the framework of 3D-space analysis. ${ }^{15}$ On the other hand, in the series of hydrocarbons the local spins on the C atoms reported here show relevant differences. Within the 3D-space formulation the atomic spin contributions were always below 0.1. ${ }^{15}$ In the Hilbert-space framework both $\mathrm{CH}_{4}\left(\left\langle\hat{S}^{2}\right\rangle_{\mathrm{C}}=0.320\right)$ and $\mathrm{C}_{2} \mathrm{H}_{6}\left(\left\langle\hat{S}^{2}\right\rangle_{\mathrm{C}}=0.199\right)$ present quite significant local spin on the C atoms. The most striking finding in Table 2 is the negative local spin on the C atom obtained for $\mathrm{C}_{2} \mathrm{H}_{2}$. Since there is no physical explanation for a negative value of $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{C}}$, we tried to understand this odd behavior. The local spin obtained at the same level of theory with the 3D-space formulation was $0.083,{ }^{15}$ which rules out the truncated CISD wave function as responsible for the spurious number. In order to check for basis set effects on the local spin for this system, we have computed the $\left\langle\hat{S}^{2}\right\rangle_{C}$ values at the CISD/6-31G** optimized geometry using several basis sets.

Table 2 CISD $/ 6-31 \mathrm{G}^{* *}$ atomic $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{A}}$ and diatomic $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{AB}}$ values for a set of singlet molecules at optimized geometries

| Molecule |  | $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{A}} /\left\langle\hat{S}^{2}\right\rangle_{\mathrm{AB}}$ | Molecule |  | $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{A}} /\left\langle\hat{S}^{2}\right\rangle_{\mathrm{AB}}$ |
| :--- | :--- | ---: | :--- | :--- | :---: |
| $\mathrm{H}_{2}$ | H | 0.036 | $\mathrm{C}_{2} \mathrm{H}_{6}$ | C | 0.199 |
|  | $\mathrm{H}-\mathrm{H}$ | -0.036 |  | H | 0.024 |
| $\mathrm{Li}_{2}$ | Li | 0.156 |  | $\mathrm{C}-\mathrm{C}$ | -0.122 |
|  | $\mathrm{Li}-\mathrm{Li}$ | -0.156 |  | $\mathrm{C}-\mathrm{H}$ | -0.069 |
| $\mathrm{Be}_{2}$ | Be | 0.175 |  | $\mathrm{C} \cdots \mathrm{H}$ | 0.034 |
|  | $\mathrm{Be}-\mathrm{Be}$ | -0.175 |  | $\mathrm{H}-\mathrm{H}$ | 0.018 |
| HF | H | 0.006 |  | $\mathrm{H} \cdots \mathrm{H}$ | -0.015 |
|  | F | 0.006 | $\mathrm{C}_{2} \mathrm{H}_{4}$ | C | 0.056 |
|  | $\mathrm{H}-\mathrm{F}$ | -0.006 |  | H | 0.024 |
| $\mathrm{H}_{2} \mathrm{O}$ | H | 0.013 |  | $\mathrm{C}-\mathrm{C}$ | -0.094 |
|  | O | 0.013 |  | $\mathrm{C}-\mathrm{H}$ | -0.036 |
|  | $\mathrm{O}-\mathrm{H}$ | -0.007 |  | $\mathrm{C} \cdots \mathrm{H}$ | 0.055 |
|  | $\mathrm{H} \cdots \mathrm{H}$ | -0.006 |  | $\mathrm{H}-\mathrm{H}$ | -0.002 |
| $\mathrm{NH}_{3}$ | N | 0.061 |  | $\mathrm{H} \cdots \mathrm{H}_{\mathrm{cis}}$ | -0.026 |
|  | H | 0.019 |  | $\mathrm{H} \cdots \mathrm{H}$ trans | -0.014 |
|  | $\mathrm{~N}-\mathrm{H}$ | -0.020 | $\mathrm{C}_{2} \mathrm{H}_{2}$ | C | -0.139 |
|  | $\mathrm{H} \cdots \mathrm{H}$ | 0.000 |  | H | 0.019 |
| $\mathrm{CH}_{4}$ | C | 0.320 |  | $\mathrm{C}-\mathrm{C}$ | 0.147 |
|  | H | 0.026 |  | $\mathrm{C}-\mathrm{H}$ | 0.048 |
|  | $\mathrm{C}-\mathrm{H}$ | -0.080 |  | $\mathrm{C} \cdots \mathrm{H}$ | -0.056 |
|  | $\mathrm{H} \cdots \mathrm{H}$ | 0.018 |  | $\mathrm{H} \cdots \mathrm{H}$ | -0.006 |

Table 3 Atomic $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{C}}$ components for acetylene molecule computed at the CISD level of theory with different basis sets

| Basis set | $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{C}}$ | $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{C}}^{\mathrm{L}}$ | $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{C}}^{3 \mathrm{D}}$ |
| :--- | ---: | :---: | :---: |
| STO-3G | 0.012 | 0.179 | 0.159 |
| 6-31G | -0.074 | 0.119 | 0.114 |
| 6-31G |  |  |  |
| 6-31G**(6d) | -0.136 | 0.082 | 0.084 |
| 6-311G | -0.139 | 0.080 | 0.083 |
| 6-311G | 0.030 | 0.119 | 0.113 |
| 6-311G | -0.118 | 0.085 | 0.085 |
| cc-pVDZ | -0 d 10f) | -0.143 | 0.091 |
| c-pVTZ | -0.162 | 0.078 | 0.084 |
| cc-pVTZ(6d 10f) | -0.271 | 0.087 | 0.085 |
| aug-cc-pVDZ | -0.146 | 0.104 | 0.079 |
| aug-cc-pVTZ | 0.729 | 0.078 | 0.087 |

The results are gathered in Table 3. The results using Löwdin and 3D-space schemes are also included. The $\left\langle\hat{S}^{2}\right\rangle_{C}$ is still negative for most basis sets, except for the STO-3G, 6-311G and aug-cc-pVDZ basis sets. Note the completely meaningless value of -2.926 obtained with Mulliken's scheme with the quite standard aug-cc-pVTZ basis set. $\ddagger$ Moreover, since negative values appear for both small and relatively large basis sets with and without polarization functions it is difficult to draw any general conclusion from the data. However, the local spin on C atoms using both Löwdin's and the 3D-space formulation is always small and positive, as it should be. One can see a systematic lowering of the value upon inclusion of polarization functions and no significant effects of the diffuse functions.

In this context it is worth noting that rather odd numbers have been obtained in the literature when combining Mulliken analysis techniques and genuine two-electron quantities from correlated wave functions. For instance, Vyboishchikov et al. ${ }^{33}$ found unphysically positive correlation contributions for diatomic energies at the CID level of theory. On the other hand, the bond order indices that in the correlated case make use of the actual pair density (the so-called delocalization index, $\mathrm{DI}^{34-39}$ ) have also been a matter of debate as, for the simplest case of $\mathrm{H}_{2}$ described with Weinbaum's classical correlated wave function, the DI gives just $0.39 .{ }^{38}$ One can conclude once again that in some cases one can get spurious results when decomposing quantities that explicitly include the second-order density matrix in the framework of Mulliken analysis. Fortunately, according to our experience, the pathological case of acetylene seems to be quite exceptional (e.g. no such problems occur for the isoelectronic $\mathrm{N}_{2}$ molecule). Nevertheless, the Hilbert-space results using Löwdin's scheme prove to be much more robust and reliable, especially for large basis sets.

Finally, we have also considered several radical (doublet) and diradical singlet molecules. In principle, the magnitude of the local spin values and the diatomic spin components, compared to the ideal values for localized spins, can help to quantify the diradical character of the molecule. For a system of two perfectly localized anti-parallel spins on centers A and B,
$\ddagger$ For this peculiar molecule, a non-nuclear attractor is found for the STO-3G and cc-pVTZ basis sets, having no apparent effect on the local spin values.

Table 4 Atomic $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{C}}$ and diatomic $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{CC}}$ components at the CASSCF $/ 6-31 \mathrm{G}^{* *}$ level of theory. Active spaces used are $(6,6)$ for benzene, $(7,7)$ for phenyl radical, $(8,8)$ for ortho-, meta- and parabenzyne and $(3,3)$ for the allyl radical

|  | $\mathrm{C}_{1}$ | $\mathrm{C}_{2}$ | $\mathrm{C}_{3}$ | $\mathrm{C}_{4}$ | $\mathrm{C}_{5}$ | $\mathrm{C}_{6}$ |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| Benzene |  |  |  |  |  |  |
| $\mathrm{C}_{1}$ | 0.114 | -0.101 | 0.079 | -0.069 |  |  |
| Phenyl radical |  |  |  |  |  |  |
| $\mathrm{C}_{1}$ | 0.968 | -0.145 | 0.100 | -0.103 |  |  |
| $\mathrm{C}_{2}$ |  | 0.129 | -0.102 | 0.079 | -0.071 | 0.078 |
| $\mathrm{C}_{3}$ |  |  | 0.127 | -0.102 | 0.080 |  |
| $\mathrm{C}_{4}$ |  |  |  | 0.118 |  |  |
| $o$-Benzyne |  |  |  |  |  |  |
| $\mathrm{C}_{1}$ | 0.324 | -0.304 | 0.086 | -0.075 | 0.080 | -0.110 |
| $\mathrm{C}_{3}$ |  |  | 0.130 | -0.107 | 0.079 |  |
| $\mathrm{C}_{4}$ |  |  |  | 0.115 | -0.092 |  |
| $m$-Benzyne | 0.540 | -0.101 | -0.344 | -0.056 | 0.069 | -0.104 |
| $\mathrm{C}_{1}$ | 0.133 |  | 0.066 | -0.065 |  |  |
| $\mathrm{C}_{2}$ |  |  |  | 0.116 | -0.097 | 0.074 |
| $\mathrm{C}_{4}$ |  |  |  |  | 0.121 |  |
| $\mathrm{C}_{5}$ |  |  |  |  |  |  |
| $p$-Benzyne | 0.962 | -0.171 | 0.125 | -0.862 |  |  |
| $\mathrm{C}_{1}$ | 0.145 | -0.106 |  | -0.072 | 0.078 |  |
| $\mathrm{C}_{2}$ |  |  |  |  |  |  |
| Allyl radical | 0.440 | -0.145 | 0.151 |  |  |  |
| $\mathrm{C}_{1}$ | 0.145 |  |  |  |  |  |
| $\mathrm{C}_{2}$ |  |  |  |  |  |  |

a proper (without spin contamination) wave function would yield local values of $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{A}}=\left\langle\hat{S}^{2}\right\rangle_{\mathrm{B}}=3 / 4$ and $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{AB}}=-3 / 4$, for an overall value of $\left\langle\hat{S}^{2}\right\rangle=0$, characteristic of a singlet. In the case of a molecular system with a singly localized unpaired electron on a center A, one can expect a local value of $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{A}}$ close to 3/4 (if this ideal system is described at the ROHF level of theory then $\left\langle\hat{S}^{2}\right\rangle_{\mathrm{A}} \equiv 3 / 4$ and all diatomic spin contributions exactly vanish). ${ }^{15}$

The local spin analysis for benzyl and allyl radicals and the set of diradical ortho-, meta- and para-benzyne molecules in the singlet electronic state are gathered in Table 4. Benzene molecule has also been included for comparison. The systems have been studied at the CASSCF/6-31G** level of theory with appropriate active spaces, i.e., full $\pi$-valence and appropriate $\sigma$ orbitals for the radical systems. We report only the results obtained for Mulliken-type analysis. Using Löwdin or 3D-space schemes does not change significantly the values for these systems.

Non-negligible local spin (0.114) is observed on the C atoms of benzene, due to the fluctuation of $\pi$-electrons induced by the electron correlation. The sign of the diatomic contributions is consistent with the chemical picture displayed in Fig. 1a. Such localization of the spins is also consistent with the observed decrease in the electronic aromaticity indices upon inclusion of electron correlation ${ }^{40}$ (note that a single-determinant restricted description of benzene gives identically zero local spins). Of course, the magnitude of the local spin and the diatomic spin contributions is very small compared to the ideal $3 / 4$ value, as one could anticipate for a genuine diamagnetic system.

In phenyl radical the unpaired electron is localized on a $\sigma$ orbital at/near $\mathrm{C}_{1}$ (in which the H atom is absent). The local spin in the remaining C atoms is slightly greater than that in the case of benzene, as well as the magnitude and sign pattern of the diatomic spin components. The local spin value of the






(g)

Fig. 1 Localized spins picture emerging from the local spin analysis. Circled arrows represent the $\sigma$ contributions, small arrows indicate the $\pi$ counterpart.
radical center ( 0.968 ) exceeds the ideal value of a singly localized electron (3/4). Furthermore, the $\left\langle\hat{S}^{2}\right\rangle_{12}$ value increases (in absolute value), as well as the local spin on $\mathrm{C}_{2}$ with respect to the benzene. Of course, one cannot expect these contributions to be fully additive as the shape of the $\pi$-orbitals in the phenyl radical system is modified by the presence of the unpaired $\sigma$ electron, and there can also be some $\sigma-\pi$ interplay. Nevertheless, one can conclude that there must be a parallel alignment of the local spin arising from the $\pi$ and $\sigma$ electrons in the radical center, as indicated in Fig. 1b.

The analysis of the allyl radical gives a completely different picture. In this case there are significant local spin contributions from all C atoms. The main local spin centers are $\mathrm{C}_{1}$ and $\mathrm{C}_{3}$ atoms, with a value of 0.440 , and the central $\mathrm{C}_{2}$ atom also contributes 0.145 . The diatomic terms reveal partial anti-ferromagnetic coupling $(-0.145)$ between $\mathrm{C}_{1}$ and $\mathrm{C}_{2}$, and a ferromagnetic one of similar magnitude ( 0.151 ) between $\mathrm{C}_{1}$ and $\mathrm{C}_{3}$. It is worth noting that the diatomic spin components would vanish for a ROHF description of this system. Thus, it is clear that in this case there is a significant interplay between the three $\pi$ electrons of the system, as depicted in Fig. 1g.

Among the set of diradical benzyne isomers, para-benzyne exhibits local spin features similar to the phenyl radical. The local spin on the radical centers (atoms $\mathrm{C}_{1}$ and $\mathrm{C}_{4}$ ) is 0.962 and the diatomic spin contribution $\left\langle\hat{S}^{2}\right\rangle_{14}=-0.862$ indicates antiferromagnetic coupling. The $\left\langle\hat{S}^{2}\right\rangle_{12}$ value is similar to that obtained for the phenyl radical, and the diatomic values involving the remaining atoms with contributions from the
$\pi$ orbitals are close to the case of benzene. The overall picture of this system is consistent with a $\sigma, \sigma$-diradical, as depicted in Fig. 1d.

The diradical character decreases in the case of ortho- and meta-benzyne. In the former, the local spin on atoms $\mathrm{C}_{1}$ and $\mathrm{C}_{2}$ is just 0.324 , substantially smaller than the ideal value of $3 / 4$ for a fully localized electron. Concomitantly, the diatomic $\left\langle\hat{S}^{2}\right\rangle_{12}$ value decreases to -0.304 . Since both values practically compensate each other, the local spin contribution from the $\sigma$ electrons is largely confined in these two atoms. Therefore, the small magnitude of the local spin indicates that the two electrons must exhibit some genuine pairing, leading to an increase in the bonding interaction between the two atoms relative to benzene. Indeed, the computed bond orders for ortho-benzyne and benzene are 2.12 and 1.38 , respectively. The diatomic terms involving the remaining atoms are again quite similar to the case of benzene. The negative $\left\langle\hat{S}^{2}\right\rangle_{16}$ value of -0.110 could have its origin merely in the $\pi$ electron system. Hence, it is difficult in this case to establish whether the picture of the system is the one displayed in Fig. 1c or the one with the $\sigma$ spins of atoms $C_{1}$ and $C_{2}$ interchanged. We have also performed the analysis for a $\operatorname{CASSCF}(2,2)$ wave function (at the same geometry) where only the $\sigma$ electrons are correlated. In this case the local spin on the $\mathrm{C}_{1}$ atom is just 0.271 . The estimate of $0.11-0.13$ contribution of the $\pi$ part for this system seems to point towards an anti-parallel arrangement of the $\sigma$ and $\pi$ electrons on the $\mathrm{C}_{1}$ and $\mathrm{C}_{2}$ atoms, as depicted in the figure.

The meta-benzyne molecule represents an intermediate situation between the ortho and para isomers. There are however some interesting features. First of all, the local spin on the $C_{1}$ and $C_{3}$ atoms amounts to an intermediate value of 0.540 , but the diatomic term is just -0.304 . Intuitively one could conclude that the contribution of the $\sigma$ to the local spin is not fully compensated between the two atoms, as in the case of ortho-benzyne. However, since the contribution of the $\pi$ electrons to the $\left\langle\hat{S}^{2}\right\rangle_{13}$ value is positive (indicating parallel alignment), it partially compensates for a larger negative contribution from the anti-parallel arrangement of the $\sigma$ electrons. In this sense it is noticeable that for this molecule the $\left\langle\hat{S}^{2}\right\rangle_{12}$ value $(-0.101)$ is exactly the same as in benzene, and also very similar to the $\left\langle\hat{S}^{2}\right\rangle_{16}$ one ( -0.104 ). However, due to symmetry, a parallel or anti-parallel alignment of the $\sigma$ and $\pi$ local spin contributions is not possible for both atoms $\mathrm{C}_{1}$ and $\mathrm{C}_{3}$, as shown in Fig. 1e and f. As these atoms are equivalent by symmetry, the picture of the localized spins of this system must be a combination of the two (equivalent) configurations. This suggests that the $\sigma$ and $\pi$ contributions to the local spin of this molecule could be additive to a large extent. Indeed, the values of the atomic and diatomic terms involving $\mathrm{C}_{2}, \mathrm{C}_{4}, \mathrm{C}_{5}$ and $\mathrm{C}_{6}$ centers are very similar to those of benzene. Also, the local spin on the $\mathrm{C}_{1}$ atom for a $\operatorname{CASSCF}(2,2)$ wave function is 0.469. After adding the estimate contribution of the $\pi$ part (ca. 0.1) one gets a value that is only slightly larger than the actual value of 0.540 .

In summary, the diradical character of the three isomers increases from ortho to para, in agreement with other analysis. ${ }^{41}$ The local spin analysis allows for a deeper insight into how the local spins are distributed in the centers and its magnitude.

## 6 Conclusions

We have carried out the decomposition of the expectation value of the spin operator for a general wave function in the framework of the Hilbert-space analysis that fulfills all the requirements imposed to date. We have shown that there is an ambiguity affecting Mulliken-type decompositions in the selection of the indices where the atoms are centered. We definitely recommend to stick to the assignment of subscripts that is consistent with Mulliken population analysis and the original Mayer-Wiberg bond orders. The results obtained are in good agreement with physical expectation and, in general, do not depend too much on the basis set. For the particular case of acetylene one can get spurious results when performing the decomposition of quantities that explicitly depend upon the second-order density matrix, even with small basis sets. We show that Löwdin's scheme is more robust and reliable in all cases.

## 7 Appendix

### 7.1 Mapping between 3D and Hilbert space analyses

One can introduce a mapping between the atomic overlap matrices (used in the framework of 3D-space analysis) and their Hilbert space analogues in order to establish a one-to-one correspondence between the expressions obtained for 3D-space and Hilbert-space analyses.

Let us consider for simplicity a closed-shell system with doubly occupied molecular orbitals. In the framework of 3D-space analysis, the gross atomic population of atom A is obtained as

$$
\begin{align*}
N_{\mathrm{A}} & =\int w_{\mathrm{A}}(\vec{r}) \rho(\vec{r}) \mathrm{d} \vec{r}=\sum_{\mu \nu} D_{\mu \nu} S_{\nu \mu}^{\mathrm{A}} \\
& =2 \sum_{\mu \nu} \sum_{i}^{\mathrm{occ}} c_{\nu i}^{*} S_{\nu \mu}^{\mathrm{A}} c_{\mu i}=2 \operatorname{tr}\left(\boldsymbol{C}^{+} \boldsymbol{S}^{\mathrm{A}} \boldsymbol{C}\right) \tag{22}
\end{align*}
$$

where

$$
\begin{equation*}
S_{\nu \mu}^{\mathrm{A}}=\int w_{\mathrm{A}}(\vec{r}) \chi_{\nu}^{*}(\vec{r}) \chi_{\mu}(\vec{r}) \mathrm{d} \vec{r} \tag{23}
\end{equation*}
$$

are the elements of the atomic overlap matrix $\boldsymbol{S}^{\mathrm{A}}$ and $\boldsymbol{C}$ is the matrix containing the orbital coefficients of the occupied molecular orbitals. The Mulliken gross population defined in eqn (6) can be also written in terms of the molecular orbital coefficients as

$$
\begin{equation*}
N_{\mathrm{A}}=2 \sum_{\mu \in \mathrm{A}} \sum_{\nu} \sum_{i}^{\mathrm{occ}} c_{\nu i}^{*} S_{\nu \mu} c_{\mu i}=2 \operatorname{tr}\left(\boldsymbol{C}^{+} \boldsymbol{S} \eta^{\mathrm{A}} \boldsymbol{C}\right) \tag{24}
\end{equation*}
$$

where $\eta^{\mathrm{A}}$ is a block-truncated unit matrix with all elements equal to zero except $\eta_{\mu \mu}^{\mathrm{A}}=1$ for $\mu \in \mathrm{A}$. Comparing eqn (22) and (24) one can put into correspondence with the atomic overlap matrix in atomic orbital basis $\boldsymbol{S}^{\mathrm{A}}$ the matrix product $\boldsymbol{S} \eta^{\mathrm{A}}$. It is easy to see that if one expresses one- and two-center terms of ref. 15 in the atomic orbital basis and replaces the matrix elements $S_{\mu \nu}^{\mathrm{A}}$ by the $\left[S_{\eta}{ }^{\mathrm{A}}\right]_{\mu \nu}$ ones, the one- and twocenter terms of eqn (20) and (21) can be recovered. Note that this Mulliken's effective atomic overlap matrix is non-symmetric, and this is the reason why the "interchanged" indices convention
mentioned above leads to different expressions for the oneand two-center components. In that scheme one essentially uses both $\boldsymbol{S} \eta^{\mathrm{A}}$ and $\eta^{\mathrm{A}} \boldsymbol{S}$ matrices. Thus, the one- and twocenter contributions of $\left\langle\hat{S}^{2}\right\rangle$ in the framework of Löwdin analysis can be easily derived if an analogous mapping could be established. In Löwdin population analysis one has

$$
\begin{equation*}
N_{\mathrm{A}}^{\mathrm{L}}=\sum_{\mu \in \mathrm{A}} \sum_{\nu \sigma} S_{\nu \mu}^{1 / 2} D_{\mu \sigma} S_{\sigma \nu}^{1 / 2} \tag{25}
\end{equation*}
$$

which can also be expressed in terms of the MO coefficients as

$$
\begin{equation*}
N_{\mathrm{A}}^{\mathrm{L}}=2 \sum_{\mu \in \mathrm{A}} \sum_{\nu \sigma} \sum_{i}^{\mathrm{occ}} c_{\sigma i}^{*} S_{\sigma \mu}^{1 / 2} \boldsymbol{S}_{\mu \nu}^{1 / 2} c_{\nu i}=2 \operatorname{tr}\left(\boldsymbol{C}^{+} \boldsymbol{S}^{1 / 2} \eta^{\mathrm{A}} \boldsymbol{S}^{1 / 2} \boldsymbol{C}\right) \tag{26}
\end{equation*}
$$

Comparing this expression with eqn (22) and (24) it is easy to identify $\boldsymbol{S}^{1 / 2} \eta^{\mathrm{A}} \boldsymbol{S}^{1 / 2}$ as the appropriate Löwdin's effective atomic overlap matrix. Note that in this case, the atomic overlap matrix is symmetric, which means that in the framework of Löwdin analysis the conventional and "interchanged" index schemes are equivalent.

## Acknowledgements

Financial help has been furnished by the Spanish MICINN Projects No. CTQ2011-23441/BQU, CTQ2011-23156/BQU and Acción Complementaria del MCI (PCI2006-A7-0631). Financial support from MICINN and the FEDER fund (European Fund for Regional Development) was also provided by grant UNGI08-4E-003. E.R.-C. acknowledges support from the Spanish FPU program (Grant No. AP2008-01231). E.M. acknowledges financial support from the EU under a Marie Curie Career Integration grant (PCI09-GA-2011-294240) and the Beatriu de Pinós program from AGAUR for the postdoctoral grant (BP_B_00236). I.M. acknowledges partial financial support from the Hungarian Scientific Research Fund (grant OTKA 71816).

## References

1 A. Clark and E. Davidson, J. Chem. Phys., 2001, 115, 7382-7392.
2 A. Clark and E. Davidson, Mol. Phys., 2002, 100, 373-383.
3 A. Clark and E. Davidson, J. Phys. Chem. A, 2002, 106, 6890-6896.
4 I. Mayer, Chem. Phys. Lett., 2007, 440, 357-359.
5 D. Alcoba, L. Lain, A. Torre and R. Bochicchio, Chem. Phys. Lett., 2009, 470, 136-139.
6 I. Mayer, Chem. Phys. Lett., 2009, 478, 323-326.
7 I. Mayer and E. Matito, Phys. Chem. Chem. Phys., 2010, 12, 11308-11314.
8 D. Alcoba, A. Torre, L. Lain and R. Bochicchio, Chem. Phys. Lett., 2011, 504, 11308-11314.
9 D. R. Alcoba, A. Torre, L. Lain and R. C. Bochicchio, J. Chem. Theory Comput., 2011, 7, 3560-3566.
10 M. Reiher, Faraday Discuss., 2006, 135, 97-124.
11 C. Herrmann, M. Reiher and B. Hess, J. Chem. Phys., 2005, 122, 034102.

12 A. V. Luzanov and O. V. Prezhdo, Mol. Phys., 2007, 105, 2879-2891.
13 M. Podewitz, C. Herrmann, A. Malassa, M. Westerhausen and M. Reiher, Chem. Phys. Lett., 2008, 451, 301-308.

14 C. Herrmann, L. Yu and M. Reiher, J. Comput. Chem., 2006, 27, 1223-1239.
15 E. Ramos-Cordoba, E. Matito, I. Mayer and P. Salvador, J. Chem. Theory Comput., 2012, 8, 1270-1279.
16 K. Takatsuka, T. Fueno and K. Yamaguchi, Theor. Chim. Acta, 1978, 48, 175-183.
17 I. Mayer and P. Salvador, Chem. Phys. Lett., 2004, 383, 368-375.
18 R. F. W. Bader, Atoms in Molecules: A Quantum Theory, Oxford Univ. Press, Oxford, 1990.
19 I. Mayer, J. Comput. Chem., 2007, 28, 204-221.
20 S. F. Vyboishchikov, P. Salvador and M. Duran, J. Chem. Phys., 2005, 122, 244110-244123.
21 R. S. Mulliken, J. Chem. Phys., 1955, 23, 1833-1841.
22 I. Mayer, THEOCHEM, 1992, 255, 1-7.
23 I. Mayer, Chem. Phys. Lett., 1983, 97, 270-274.
24 H. Sato and S. Sakaki, J. Phys. Chem. B, 2007, 111, 672.
25 R. Ponec and D. L. Cooper, Int. J. Quantum Chem., 2004, 97, 1002-1011.
26 J. G. Angyan, M. Loos and I. Mayer, J. Phys. Chem., 1994, 98, 5244-5248.
27 I. Mayer, Chem. Phys. Lett., 2012, 539-540, 172-174.
28 P.-O. Löwdin, Adv. Quantum Chem., 1970, 5, 185.
29 J. Baker, Theor. Chim. Acta, 1985, 68, 221.
30 I. Mayer, Chem. Phys. Lett., 2004, 393, 209.
31 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, Gaussian 03, Revision C.02, Gaussian, Inc., Pittsburgh, PA, 2003.
32 E. Matito and F. Feixas, DMN program, University of Girona (Spain) and University of Szczecin, Poland, 2009.
33 S. Vyboishchikov and P. Salvador, Chem. Phys. Lett., 2006, 430, 204-209.
34 R. Bader and M. Stephens, Chem. Phys. Lett., 1974, 26, 445-449.
35 R. Bader and M. Stephens, J. Am. Chem. Soc., 1975, 97, 7391-7399.
36 J. Ángyán, E. Rosta and P. Surján, Chem. Phys. Lett., 1999, 299, $1-8$.
37 M. Giambiagi, M. Giambiagi and F. Jorge, Theor. Chim. Acta, 1985, 68, 337-341.
38 I. Mayer, Int. J. Quantum Chem., 1986, 29, 477-483.
39 X. Fradera, M. Austen and R. Bader, J. Phys. Chem. A, 1999, 103, 304-314.
40 E. Matito, M. Solà, P. Salvador and M. Duran, Faraday Discuss., 2007, 135, 325-345.
41 A. Clark and E. Davidson, J. Am. Chem. Soc., 2001, 123, 10691-10698.


[^0]:    ${ }^{a}$ Department of Chemistry and Institute of Computational Chemistry, University of Girona, 17071 Girona, Spain.
    E-mail: Pedro.Salvador@udg.edu
    ${ }^{b}$ Institute of Organic Chemistry, Research Centre for Natural Sciences, Hungarian Academy of Sciences, P.O. Box 17, H-1525 Budapest, Hungary

[^1]:    $\dagger$ It has recently been discovered that this type of decomposition had also been used by some of us ${ }^{7}$ in the decomposition of $\left\langle\hat{S}^{2}\right\rangle$ from a formula different from eqn (1) as a result of a programming error: two subscripts have been interchanged by a mistake in the treatments of the cumulants.

