Reduced densities and cumulants, bond indices and natural adaptive orbitals

ELECTRONIC SUPPLEMENTARY INFORMATION

Electron counting in position space: from quantum fragments to Lewis structures to multicenter bonds

by

Ángel Martín Pendás* and Evelio Francisco*

Departamento de Química Física y Analítica. Facultad de Química. Universidad de Oviedo. 33006-Oviedo. Spain.

Email: ampendas@uniovi.es
Email: evelio@uniovi.es

Contents

1. Calculation of spinless \( n^{th} \)-order reduced densities
2. Cumulant densities from reduced densities
3. Bond indices and natural adaptive orbitals
4. References
1 Calculation of spinless $n^{\text{th}}$-order reduced densities

The spinless $n^{\text{th}}$-order reduced density ($n$RD) is defined here as

$$\rho^n(r_1 \ldots r_n) = \binom{N}{n} n! \int \Psi^* \Psi \, d\mathbf{x}_{i>n} \, d\sigma_{i\leq n}$$  \hspace{1cm} (1)

where $\Psi(1, N)$ is a $N-$ electron wavefunction, $\mathbf{x} \equiv r\sigma$ is a spin($\sigma$)-spatial($r$) coordinate, and $d\mathbf{x}_{i>n}$ and $d\sigma_{i\leq n}$ mean integration over $x_{n+1} \ldots x_N$ and $\sigma_1 \ldots \sigma_n$, respectively. We will develop an useful expression for $\rho^n$ in case that $\Psi$ is a multideterminant wavefunction

$$\Psi(1, N) = \sum_K C_K \Psi_K(1, N),$$ \hspace{1cm} (2)

where each determinant $\Psi_K(1, N)$ is made of real and orthonormal spin-orbitals $\phi_i$:

$$\Psi_K(1, N) = \frac{1}{\sqrt{N!}} \det |\phi_{r_1}(1) \ldots \phi_{r_N}(N)|$$ \hspace{1cm} (3)

Then, Eq. 1 becomes

$$\rho^n(r_1 \ldots r_n) = \sum_{K,L} C_K C_L \rho^n_{KL}(r_1 \ldots r_n), \quad \text{with}$$ \hspace{1cm} (4)

$$\rho^n_{KL}(r_1 \ldots r_n) = \frac{1}{(N-n)!} \int \Psi_K \Psi_L \, d\mathbf{x}_{i>n} \, d\sigma_{i\leq n}.$$ \hspace{1cm} (5)

We will summarize now the main steps necessary to write $\rho^n(r_1 \ldots r_n)$ in the form

$$\rho^n(r_1 \ldots r_n) = \sum_{a_1 \ldots a_n b_1 \ldots b_n} c_{a_1 \ldots a_n}^{b_1 \ldots b_n} \prod_k^n \phi_{a_k}(r_k) \phi_{b_k}(r_k).$$ \hspace{1cm} (6)

A $r_i > 0$ in Eq. 3 denotes a spin-orbital with a spin function $\alpha$, and $-r_i < 0$ refers to the spin-orbital with the same spatial part and spin function $\beta$. Every $r_i$ satisfies $1 \leq |r_i| \leq M$, where $M$ is the total number of molecular orbitals (MO). We will follow the nomenclature of Löwdin[1] and rename the spin-orbitals ($\phi_{r_1} \ldots \phi_{r_N}$) and ($\phi_{s_1} \ldots \phi_{s_N}$) (the latter are the spin-orbitals defining $\Psi_L(1, N)$) as $(u_1 \ldots u_N)$ and $(v_1 \ldots v_N)$, respectively. The overlap between the Slater determinants $U = \det |u_1 \ldots u_N|$ and $V = \det |v_1 \ldots v_N|$ is given by $[1] \langle U \mid V \rangle = N! D_{UV} = N! \det |d_{uv}(kl)|$, where $d_{uv}(kl) = \langle u_k \mid v_l \rangle$. For the present case, with an orthonormal MO basis, $d_{uv}(kl)$ is either 1 or 0 depending on whether $u_k = v_l$ or $u_k \neq v_l$, respectively. If the spin-orbitals $(v_1 \ldots v_N)$ in $V$ are re-ordered such that $D_{UV}$ is diagonal, the coefficient $C_L$ of the Slater determinant remains unchanged or changes its sign depending if the number of transpositions needed to put these spin-orbitals back to their original situation ($R$) is even or odd, respectively. Now, $[u_1(1) \ldots u_N(N)]$ is expanded in terms of its first $n$ rows$[1]$

$$[u_1(1) \ldots u_N(N)] = \sum_k \det |u_{k_1} \ldots u_{k_n}(n)| \times \det_u(n | k).$$ \hspace{1cm} (7)

In this equation, $k$ runs over all possible ordered sets $k_1 < k_2 < \ldots < k_n$ and $\det_u(n | k)$, that only depends on the coordinates of electrons $n+1$ to $N$, is the determinant obtained
by eliminating the rows 1...n and the columns $k_1...k_n$ from $U$. If Eq. 7 for $U$ and the analogous one for $V$ are put in Eq. 5, and coordinates $x_{n+1}...x_N$ are integrated, we obtain

$$
\rho_{KL}^n(r_1...r_n) = \int \sum_{k,l} |U_k| |V_l| D_{UV}(k|l) \, d\sigma_{i\leq n},
$$

where $|U_k| = \det|u_{k_1}(1)...u_{k_n}(n)|$, $|V_l| = \det|v_{l_1}(1)...v_{l_n}(n)|$, and $D_{UV}(k|l)$ is the minor of order $(N-n)$ built by deleting the rows $k_1...k_n$ and the columns $l_1...l_n$ from $D_{UV}$. Since $D_{UV}$ is already a diagonal determinant with only 1’s and 0’s in the diagonal, each $D_{UV}(k|l)$ in Eq. 8 can only be 1 or 0. $D_{UV}(k|l)$ will thus vanish for $k \neq l$ regardless $K = L$ or $K \neq L$. When $K = L$, $D_{UV} = \det |I_N| = 1$, where $I_N$ is the $(N \times N)$ unit matrix, and all the $D_{UV}(k|l)$ are 1. Hence,

$$
\rho_{KK}^n(r_1...r_n) = \int \sum_k |U_k| |V_k| = \sum_k \sum_{p,Q} (-1)^{P+Q} \prod_{i=1}^n \left( \int u_{p_i}(x_i)v_{q_i}(x_i) \, d\sigma \right)
$$

where $P$ and $Q$ run over the $n!$ permutations of the indices contained in $k$, and $(p_1...p_n)$ and $(q_1...q_n)$ are the permutation of $k$ associated to $P$ and $Q$, respectively. The product over $i$ only survives if $u_{p_i}$ and $v_{q_i}$ are both either $\alpha$ or $\beta$ spin-orbitals. When this condition is met, calling $a_i = |p_i|$ and $b_i = |q_i|$, for $K = L$, the coefficient $c_{a_1a_2...a_n}^{b_1b_2...b_n}$ in Eq. 6 has to be increased by $C_K^2(-1)^{P+Q+R}$.

We consider now the case where $\Psi_K$ and $\Psi_L$ differ in $d$ spin-orbitals. When $d > n$ the determinant $D_{UV}$ (after re-ordering the columns of $V$ such that $D_{UV}$ be diagonal) will contain more than $n$ 0’s in the diagonal. Consequently, all the minors $D_{UV}(k|k)$ obtained by deleting $n$ rows and (the same) $n$ columns from $D_{UV}$ will still contain one or more 0’s in the diagonal and will be 0. In other words, the product $\Psi_K^*\Psi_L$ will not contribute to $\rho^n$ if both determinants differ in more than $n$ spin-orbitals. The particular cases with $n = 1$ and $n = 2$ are very well known.[2] On the other hand, since a determinant remains unchanged if the same series of rows and columns exchanges are performed, we can always put $D_{UV}$ into a form such that the $d \leq n$ spin-orbitals which are different in $U$ and $V$ appear in the first $d$ positions. After this has been done, $D_{UV}(k|k)$ will vanish unless $k_1 = 1$, $k_2 = 2$, ..., $k_d = d$. As a consequence, $\rho_{KK}^n$ for $K \neq L$ is also given by Eq. 9 with the particularity that the first $d$ $k_i$’s are fixed to 1, 2, ..., albeit $P$ and $Q$ still run over the $n!$ permutations of the indices 1, 2, ..., $d$, $k_{d+1},...k_n$, and $p_1...p_n$ and $q_1...q_n$ are the permutations of 1, 2, ..., $d$, $k_{d+1},...k_n$ associated to $P$ and $Q$, respectively. The rest of arguments are analogous to those given after Eq. 9, with the difference that $c_{a_1a_2...a_n}^{b_1b_2...b_n}$ has to be increased by $C_KC_L(-1)^{P+Q+R}$ instead of $C_K^2(-1)^{P+Q+R}$, as in the $K = L$ case.

Since we are using real spin-orbitals, $\phi_{a_k}(r_k)\phi_{b_k}(r_k) = \phi_{b_k}(r_k)\phi_{a_k}(r_k)$, so that each pair of indices $(a_k, b_k)$ can be condensed to a single index $i_k$, defined as $i_k = \max(a_k, b_k) \times (\max(a_k, b_k) - 1) + \min(a_k, b_k)$, and $\rho^n$ written also in the form

$$
\rho^n(r_1,...,r_n) = \sum_{i_1i_2...i_n} C_{i_1i_2...i_n} \prod_k \varphi_{i_k}(r_k),
$$

where the products of pairs of MOs are stored in the order $\varphi_1 = \phi_1\phi_1$, $\varphi_2 = \phi_2\phi_1$, $\varphi_3 = \phi_2\phi_2$, etc. As a consequence of the electron indistinguishability, all $C_{i_1i_2...i_n}$’s differing
only in the order of \(i_1,i_2,\ldots,i_n\) are equal, so that only the those coefficients with \(i_1 \geq i_2 \geq \ldots \geq i_n\) need be stored. \(\rho^{n-1}\) can be obtained from \(\rho^n\) by integrating the over \(r_n:\)

\[
\rho^{n-1}(r_1,\ldots,r_{n-1}) = \frac{1}{N - n + 1} \int \rho^n(r_1,\ldots,r_n) dr_n, \tag{11}
\]

\[
= \frac{1}{N - n + 1} \sum_{i_1,i_2,\ldots,i_n} C_{i_1,i_2,\ldots,i_n} \varphi_{i_1}(r_1) \varphi_{i_2}(r_2) \cdots I_{i_n} \tag{12}
\]

where \(I_{i_n} = \int \varphi_{i_n}(r_n) dr_n = 1\) for \(i_n = 1,3,6,\ldots\) and \(I_{i_n} = 0\) otherwise.

\section{Cumulant densities from reduced densities}

Equation 10 for \(n = 1,2,\) and 3 adopts the form (we suppress the upper limit \(n \times (n+1)/2\) in all the following summation symbols):

\[
\rho^1(r_1) = \rho(r_1) = \sum_i C_i \varphi_i(r_1) \tag{13}
\]

\[
\rho^2(r_1,r_2) = \sum_{ij} C_{ij} \varphi_i(r_1) \varphi_j(r_2) \tag{14}
\]

\[
\rho^3(r_1,r_2,r_3) = \sum_{ijk} C_{ijk} \varphi_i(r_1) \varphi_j(r_2) \varphi_k(r_3), \tag{15}
\]

with \(i_j = C_{ji}\) and \(C_{ijk} = C_{ikj} = C_{jki} = C_{kij} = C_{kji}.\) The \(n^{th}\)-order cumulant density \((nCD), \rho^n_c,\) is that part of \(\rho^n\) that can not be obtained in terms of \(\rho^m's\) with \(m < n.\) It can be obtained following the procedure described in the Appendix of Ref. [3]. The explicit formulas for the first 3 cumulants are

\[
\rho^1_c(r_1) = \rho(r_1) \tag{16}
\]

\[
\rho^2_c(r_1,r_2) = \rho(r_1) \rho(r_2) - \rho^2(r_1,r_2) \tag{17}
\]

\[
\rho^3_c(r_1,r_2,r_3) = \rho(r_1) \rho(r_2) \rho(r_3) - \frac{1}{2} \hat{S} \rho(r_1) \rho^2(r_2,r_3) + \frac{1}{2} \rho^3(r_1,r_2,r_3) \tag{18}
\]

where \(\hat{S} \rho(r_1) \rho^2(r_2,r_3) = \rho(r_1) \rho^2(r_2,r_3) + \rho(r_2) \rho^2(r_1,r_3) + \rho(r_3) \rho^2(r_1,r_2)\). The nCDs for \(n = 1-9\) are collected in Table 1. Each term \(\rho^p_1 \rho^{p_2}_2 \cdots\) in this table must be actually understood as (the absence of a subindex in \(\rho^n\) means that its value is 1):

\[
\rho^p_1 \rho^{p_2}_2 \cdots \rightarrow \hat{S} \rho^{p_1}_1 \rho^{p_2}_2 \cdots \rho^{p_2}_2 \cdots, \tag{19}
\]

where \(\hat{S}\) is a symmetrizing operator. For instance, \(\rho^2_1 \rho^2_2\) that appears for \(n = 4\) must be replaced by \(\rho^1_1 \rho^1_2 \rho^2_3 \rho_4 + \rho^1_1 \rho^1_3 \rho^2_2 \rho_4 + \rho^1_1 \rho^1_4 \rho^2_2 \rho_3 + \rho^1_2 \rho^1_3 \rho^2_1 \rho_4 + \rho^1_2 \rho^1_4 \rho^2_1 \rho_3 + \rho^1_4 \rho^1_3 \rho^2_1 \rho_2).\)

Using Eqs. 13–15, the cumulants 16–18 can be recast as

\[
\rho^1_c(r_1) = \sum_i D_i \varphi_i(r_1) \tag{20}
\]

\[
\rho^2_c(r_1,r_2) = \sum_{ij} D_{ij} \varphi_i(r_1) \varphi_j(r_2) \tag{21}
\]

\[
\rho^3_c(r_1,r_2,r_3) = \sum_{ijk} D_{ijk} \varphi_i(r_1) \varphi_j(r_2) \varphi_k(r_3) \tag{22}
\]
Table 1: Cumulant of orders 1 \ldots 9. The number of terms appears in parenthesis.

<table>
<thead>
<tr>
<th>$n$</th>
<th>Cumulant of orders 1 \ldots $n$</th>
<th>Number of terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\rho^1$</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>$\rho_2^1$, $\rho^2$</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>$\rho_3^1$, $\rho_3^2$, $\rho^3$</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>$\rho_4^1$, $\rho_4^2$, $\rho_4^3$, $\rho^4$</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>$\rho_5^1$, $\rho_5^2$, $\rho_5^3$, $\rho_5^4$, $\rho^5$</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>$\rho_6^1$, $\rho_6^2$, $\rho_6^3$, $\rho_6^4$, $\rho_6^5$, $\rho^6$</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>$\rho_7^1$, $\rho_7^2$, $\rho_7^3$, $\rho_7^4$, $\rho_7^5$, $\rho_7^6$, $\rho^7$</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>$\rho_8^1$, $\rho_8^2$, $\rho_8^3$, $\rho_8^4$, $\rho_8^5$, $\rho_8^6$, $\rho_8^7$, $\rho^8$</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>$\rho_9^1$, $\rho_9^2$, $\rho_9^3$, $\rho_9^4$, $\rho_9^5$, $\rho_9^6$, $\rho_9^7$, $\rho_9^8$, $\rho^9$</td>
<td>9</td>
</tr>
</tbody>
</table>
with
\[ D_i = C_i \] (23)
\[ D_{ij} = C_i C_j - C_{ij} \] (24)
\[ D_{ijk} = C_i C_j C_k + [C_{ijk} - C_i C_{jk} - C_j C_{ik} - C_k C_{ij}] / 2 \] (25)

The symmetries of the \( D_{i_1 i_2 \ldots i_n} \) and \( C_{i_1 i_2 \ldots i_n} \) coefficients are the same. The cumulant \( \rho^{n-1}_c \) can be obtained from \( \rho^n_c \) by integrating over \( r_n \):
\[ \rho^{n-1}_c(r_1, \ldots, r_{n-1}) = \int \rho^n_c(r_1, r_2, \ldots, r_n) dr_n. \] (26)

Integrating all the electrons:
\[ \int \rho^n_c(r_1, \ldots, r_n) dr_1 \ldots dr_n = N. \] (27)

This equation can be immediately derived by applying \( n \) times Eq. 26. In effect, after applying \( n-1 \) times Eq. 26 we have
\[ \rho^1_c(r_1) = \rho(r_1) = \int \rho^2_c(r_1, r_2) dr_2, \] (28)
and integrating one last time
\[ \int \rho(r_1) dr_1 = N. \] (29)

3 Bond indices and natural adaptive orbitals

The extensivity of CDs allows for any \( \rho^n_c \) to be obtained from the \( m \) CDs with \( m > n \) (see Eq. 26). We can use this property to obtain a one-basin partition of \( \rho(r) = \rho^1_c(r) \) from the exchange correlation (xc) cumulant, \( \rho_{xc}(r_1, r_2) = \rho^2_c(r_1, r_2) \)
\[ \rho^1_c(r) = \sum_a \int_{\Omega_a} dr_2 \rho^2_c(r, r_2) = \sum_a \rho^1_a(r) \equiv \sum_a \rho^a(r), \] (30)
where \( \bigcup_{i=a}^{m} \Omega_a = \mathcal{R}^3 \). The \( \rho^a(r) \)'s are exactly equivalent to the charge weighted domain averaged Fermi Holes (DAFH) [4, 5] that we have explored in the past [6, 7]. Similarly, \( \rho^n_c(r) \) allows for a \((n-1)\)-basin partition of \( \rho^1_c(r) \):
\[ \rho^1_c(r) = \sum_{ab} \rho^{ab}(r) = \sum_{ab} \int_{\Omega_a} dr_2 \int_{\Omega_b} dr_3 \rho^3_c(r, r_2, r_3), \] (31)
\[ \rho^1_c(r) = \sum_{abc} \rho^{abc}(r) = \sum_{abc} \int_{\Omega_a} dr_2 \int_{\Omega_b} dr_3 \int_{\Omega_c} dr_4 \rho^4_c(r, r_2, r_3, r_4), \] (32)
\[ \cdots \]
\[ \rho^1_c(r) = \sum_{ab \ldots n-1} \rho^{ab \ldots n-1}(r) = \sum_{ab \ldots n-1} \int_{\Omega_a} dr_2 \int_{\Omega_b} dr_3 \ldots \int_{\Omega_{n-1}} dr_n \rho^4_c(r, r_2, \ldots, r_n). \] (33)
This partition of \( \rho_c^2(\mathbf{r}) \) into basins, pairs of basins, etc, can be extended to higher CDs:

\[
\rho_c^2(r_1, r_2) = \sum_a \int_{\Omega_a} dr_3 \rho_c^3(r_1, r_2, r_3)
\]

\[
\rho_c^2(r_1, r_2) = \sum_{ab} \int_{\Omega_a} dr_3 \int_{\Omega_b} dr_4 \rho_c^4(r_1, r_2, r_3, r_4),
\]

\[
\rho_c^2(r_1, r_2) = \sum_{ab...n-2} \int_{\Omega_a} dr_3 \int_{\Omega_b} dr_4 \ldots \int_{\Omega_{n-2}} dr_n \rho_c^n(r_1, r_2, \ldots, r_n).
\]

If instead of integrating all but electron 1 in Eqs. 30-32 to obtain a generalized \( n \)-basin density, all the electrons are integrated, the result is a scalar depending only on the definition of the \( \Omega_i \) basins:

\[
\langle N_a \rangle = \int_{\Omega_a} dr_1 \rho_c^1(r_1),
\]

\[
\langle N_{ab} \rangle = \int_{\Omega_a} dr_1 \int_{\Omega_b} dr_2 \rho_c^2(r_1, r_2),
\]

\[
\langle N_{abc} \rangle = \int_{\Omega_a} dr_1 \int_{\Omega_b} dr_2 \int_{\Omega_c} dr_3 \rho_c^3(r_1, r_2, r_3),
\]

\[
\langle N_{ab...n} \rangle = \int_{\Omega_a} dr_1 \int_{\Omega_b} dr_2 \ldots \int_{\Omega_n} dr_n \rho_c^n(r_1, \ldots, r_n).
\]

Since \( \rho_c^1(\mathbf{r}) = \rho(\mathbf{r}) \) and \( \rho_{ac}(r_1, r_2) = \rho_c^2(r_1, r_2) \), then \( \langle N_a \rangle \) and \( 2 \langle N_{ab} \rangle (a \neq b) \) are the average electronic population of basin \( \Omega_a \) and the delocalization index (DI) between \( \Omega_a \) and \( \Omega_b \), respectively. In general, \( n!\langle N_{ab...} \rangle \) can be identified with a \( n \)-center bonding or delocalization index (DI). It is very easy to show by using Eq. 26 that each generalized density \( \rho_{abc...}(\mathbf{r}) \) is normalized to the \( n \)-center DI of the same order, i.e.

\[
\int \rho_c(r)dr = \langle N_a \rangle \quad \quad (41)
\]

\[
\int \rho_{ab}(r)dr = \langle N_{ab} \rangle \quad \quad (42)
\]

\[
\int \rho_{abc}(r)dr = \langle N_{abc} \rangle \quad \quad (43)
\]

\[
\int \rho_{ab...n}(r)dr = \langle N_{ab...n} \rangle. \quad \quad (44)
\]

Each \( n \)-center DI can be recovered from the \( (n + 1) \)-center DIs, \( N = \sum_a N_a, \langle N_a \rangle = \sum_b \langle N_{ab} \rangle, \langle N_{ab} \rangle = \sum_c \langle N_{abc} \rangle, \ldots \), and the total number of electron, \( N \), can be partitioned into basins, pairs of basins, etc, \( N = \sum_a \langle N_a \rangle = \sum_{ab} \langle N_{ab} \rangle = \sum_{abc} \langle N_{abc} \rangle \ldots \).

The \( n \)-center generalized densities \( \rho_{abc...}(\mathbf{r}) \) admit the unified expression

\[
\rho_{abc...}(\mathbf{r}) = \phi(\mathbf{r})D_{abc...}\phi(\mathbf{r})^\dagger, \quad \quad (45)
\]

where the row \( \phi = (\phi_1, \phi_2, \ldots, \phi_M) \) represents the set of all the MOs of the system and \( D_{abc...} \) is a symmetric matrix. The diagonalization of \( \mathbf{D}^a \) (corresponding to taking \( n = 1 \))
leads to the DAFH eigenvectors ($\psi_i^a$) and eigenvalues ($n_i^a$), yielding $\rho^a = \sum n_i^a (\psi_i^a)^2$. This diagonalization has been used many times by Robert Ponec and other authors to analyze chemical bonding issues in relatively large sets of molecules. The $\psi_i^a$’s are called domain natural orbitals (DNO) and may be either localized or delocalized, with only the latter contributing significantly to bonding [6]. Similar diagonalizations of further order CDs may be even more important. The two-center partition of $\rho$ coming from $\rho_3^c$ (Eq. 31) will give rise, after diagonalizing $D_{ab}$, to orbitals $\psi_{ab}^i$ which describe the two-center bonds between two given basins. Each of these orbitals will contribute additively to the bond order with the quantity $n_{ab}^i$, the eigenvalue associated to $\psi_{ab}^i$. In a similar way, we can decompose 3-, 4- and, in general, $n$-center bonds. We call these effective one-electron functions Natural Adaptive Orbitals (NAdOs) [8]. In all the cases, the $n$-center DI is given as the sum of all eigenvalues of these NAdOs, $\langle N_{abc...} \rangle = \sum_i n_{abc...}^i$. As the diagonalization leaves invariant the trace of a matrix we also have $\langle N_{abc...} \rangle = \text{tr} D_{abc...}$. A deeper insight on the meaning of the NAdOs is given in Reference [8].

References