Hierarchies of quantum chemical descriptors induced by statistical analyses of domain occupation number operators

Guillaume Acke, Stijn De Baerdemacker, Ángel Martín Pendás, Patrick Bultinck

Article Type:

Advanced Review

Abstract

As the approximations to the wave functions governing quantum chemical systems are becoming more and more complex, it is becoming increasingly important to devise descriptors that condense information and in this way provide conceptual building blocks that can be used for understanding the practical results of those approximations. Quantum chemical descriptors that are able to capture the statistical signatures of quantum chemical interactions have been shown to lead to such insightful results. Central to an understanding of these descriptors is the concept of a 'domain occupation number operator', which allows us to treat the so-called real space and Hilbert space partitionings on the same footing. We will show that many of the existing descriptors can be expressed as the (central) densities and density cumulants associated with these domain operator. We will also show that these densities can be obtained by successive differentiations of generating functions, effectively establishing a hierarchy behind these domain associated densities. Not only do the resulting hierarchies indicate how many of the previously reported descriptors are related, they also show which areas have not yet been explored.

1

GRAPHICAL TABLE OF CONTENTS

electron density

(cDFT, ELF, ELI, DORI)

$$\mathcal{C}[g, \boldsymbol{t}] = \ln\left(\mathcal{M}[g, \boldsymbol{t}]\right)$$

domain population

$\frac{\partial}{\partial g(\boldsymbol{r}_1)}$ $\frac{\partial}{\partial t_a}$

exchange-correlation hole $\kappa^{(1)}(\boldsymbol{r}_1) = \rho^{(1)}(\boldsymbol{r}_1)$ $\kappa^{(1)}(\Omega_a) = \rho^{(1)}(\Omega_a)$ (de)localization index

$\frac{\partial}{\partial g(\boldsymbol{r}_2)}$ $\frac{\partial}{\partial t_a}$ $\frac{\partial}{\partial g(\boldsymbol{r}_1)}$ $\frac{\partial}{\partial t_b}$

$\kappa^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2)$ $\kappa^{(2)}(\boldsymbol{r}_1, \Omega_a)$ $\kappa^{(2)}(\Omega_a, \Omega_b)$

$\frac{\partial^{N-2}}{\partial t_l \ldots \partial t_k}$

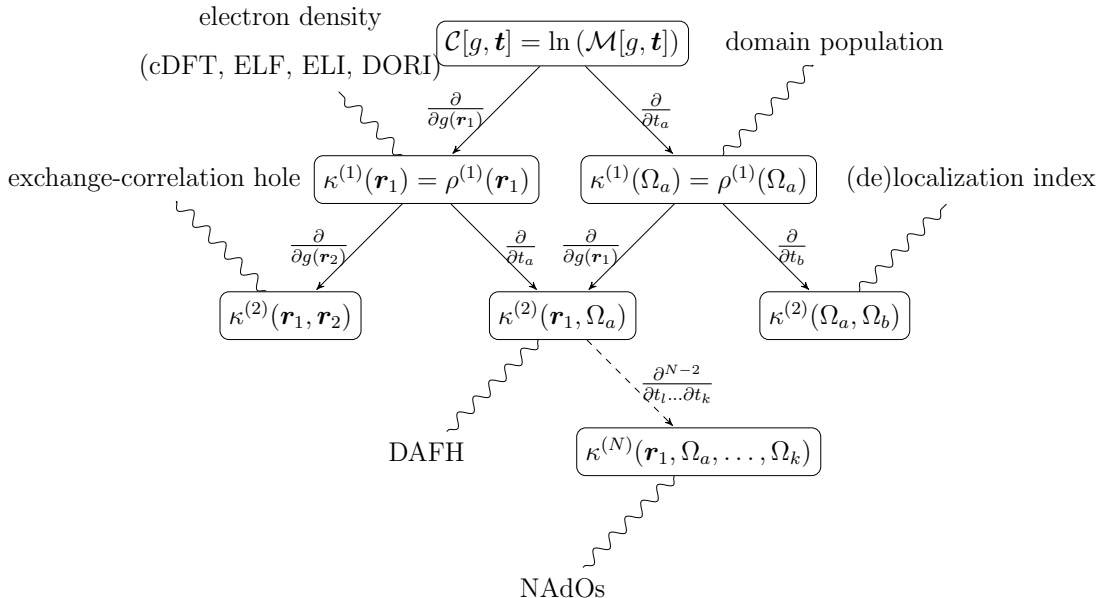DAFH $\kappa^{(N)}(\boldsymbol{r}_1, \Omega_a, \ldots, \Omega_k)$

NAdOs

Figure 1: The statistics of (domain) occupation numbers operators induces a hierarchy of quantum chemical descriptors. Not only does the resulting hierarchy indicate how many of the previously reported chemical indices are related, they also show which descriptors have not yet been fully explored.

INTRODUCTION

The field of quantum chemistry continues to make important advances in devising increasingly accurate approximations to the wave functions of increasingly larger systems. These approximations are aimed at modeling as accurately as possible the joint behavior of all electrons. However, as these approximations become increasingly complex, it has become increasingly difficult to gain insight into the wave functions that arise from actual computations. As the approximate wave functions are constantly being loaded with more and more information, it becomes increasingly important to devise descriptors that condense information and in this way provide conceptual building blocks that can be used in "chemical" reasoning.

In this review we will demonstrate that a large part of the existing quantum chemical descriptors are based on the concepts of 'domains' and 'occupation numbers'. We define a

'domain' as a collection of orbitals. This definition will allow us to describe those domains defined in position space (also called 'real space'), those defined on non-orthogonal orbitals (leading to e.g. Mulliken descriptions) and those defined in terms orthogonal orbitals (for e.g. the abstract Hubbard model) on the same footing. 'Occupation numbers' will indicate to what extent states are filled by particles of a quantum-chemical system consisting of many identical particles. By extending the occupation number to the domain occupation numbers, we can describe to what extent domains are filled with particles. By treating the resulting occupation as a random variable, we can use generating functions to generate (central) 'moments' and 'cumulants'. It are these statistical quantities that provide insightful interpretations of wave functions. As a result, the generating functions establish a natural hierarchy or framework into which many reported quantum chemical descriptors fit. Furthermore, this framework provides a structure that offers a consistent interpretation while also pointing to those descriptors that have not yet been explored.

ORBITAL AND DOMAIN OCCUPATION

Many quantum chemical descriptors are based on the premise that chemical knowledge is encoded in the extent to which domains are filled by electrons. Among others, this is the basis behind a plethora of population analysis methods that will be discussed below. In order to elucidate the framework behind these descriptors, we first have to introduce operators that probe the occupation of domains. For this, we have to start from operators that probe the occupation of a single orbital. These operators are most efficiently formulated in terms of second quantized operators. We have provided a short introduction to the necessary second quantized machinery in the appendix.

Orbital occupation

The occupation number of a single orbital indicates to what extent that orbital is filled by particles of the quantum chemical system consisting of many identical particles. In an orthonormal spin-orbital basis, the second quantized occupation-number operator $\hat{n}(I)$

associated with the occupation of the spin-orbital ϕ_I is given by

$$\hat{n}(I) := \hat{\phi}_I^+ \hat{\phi}_I^- \ . \tag{1}$$

By acting on the antisymmetric many-particle state $|\phi_{k_1} \ldots \phi_{k_N}\rangle = |\boldsymbol{k}\rangle$, with $\boldsymbol{k} = \{k_1, \ldots, k_N\}$ a unique set of spin-orbital indices, the occupation-number operator $\hat{n}(I)$ determines to what extent the orbital ϕ_I is filled in $|\phi_{k_1} \ldots \phi_{k_N}\rangle$

$$
\begin{aligned}
\hat{n}(I) |\phi_{k_1} \ldots \phi_{k_N}\rangle &= \hat{\phi}_I^+ \hat{\phi}_I^- |\phi_{k_1} \ldots \phi_{k_N}\rangle \\
&= n(I) |\phi_{k_1} \ldots \phi_{k_N}\rangle \\
&= \begin{cases} 0 \, |\phi_{k_1} \ldots \phi_{k_N}\rangle & \text{if } I \text{ not in } \boldsymbol{k} \\ 1 \, |\phi_{k_1} \ldots \phi_{k_N}\rangle & \text{if } I \text{ in } \boldsymbol{k} \end{cases} .
\end{aligned} \tag{2}
$$

where $n(I)$ is the eigenvalue associated with the eigenvector $|\phi_{k_1} \ldots \phi_{k_N}\rangle$. In the orthonormal spin-position orbital basis $|\boldsymbol{x}\rangle = |\boldsymbol{r}, \sigma\rangle$, the occupation-number operator can be written in terms of field operators

$$\hat{n}(\boldsymbol{x}) = \hat{\psi}^+(\boldsymbol{x})\hat{\psi}^-(\boldsymbol{x}) \ . \tag{3}$$

In a non-orthogonal basis χ_I, the occupation-number operator can be written as (Surján, 1989)

$$\hat{n}(\chi_I) = \hat{\chi}_I^+ \hat{\varphi}_I^- = \sum_J \hat{\varphi}_I^+ \hat{\varphi}_J^- \boldsymbol{\Sigma}_{IJ} \ , \tag{4}$$

where $\boldsymbol{\Sigma}_{IJ}$ is the overlap matrix for that non-orthogonal spin-orbital basis and φ_I is the bi-orthogonal basis for which the usual anticommutation rules are valid

$$\left[\hat{\chi}_I^+, \hat{\varphi}_I^-\right]_+ = \delta_{IJ} \ . \tag{5}$$

Domain occupation

If we consider a domain to be a weighted collection of orbitals, then we can construct an associated domain occupation-number operator as

$$\hat{n}(\Omega_a) = \sum_I w_{\Omega_a}(I)\hat{n}(I) \ , \tag{6}$$

where $w_{\Omega_a}(I)$ is the weight factor of domain Ω_a associated with spin-orbital I. This approach is a second quantized application of the 'atomic decomposition of the identity', which was

4

introduced by Mayer (Mayer & Hamza, 2005; Mayer, 2016). In spin-position space, this definition reduces to

$$\hat{n}(\Omega_a) = \int w_{\Omega_a}(\boldsymbol{x})\hat{n}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \ , \tag{7}$$

and for a non-orthogonal spin-orbital basis to (Mayer, 1983, 1985)

$$\hat{n}(\Omega_a) = \sum_I w_{\Omega_a}(\chi_I)\hat{n}(\chi_I) = \sum_{IJ} w_{\Omega_a}(\chi_I)\hat{\chi}_I^+ \hat{\varphi}_I^- = \sum_{IJ} w_{\Omega_a}(\chi_I)\hat{\varphi}_I^+ \hat{\varphi}_J^- \boldsymbol{\Sigma}_{IJ} \ . \tag{8}$$

In practice, the domain itself is mostly defined on the spin-independent part of the basis. In this case, the domain occupation-number operator $\hat{n}(\Omega)$ associated with a domain Ω is given by a weighted summation over all spin-independent occupation-number operators associated with a spin-independent single particle basis $\{|i\rangle\}$

$$\hat{n}(\Omega) = \sum_i w_\Omega(i)\hat{n}(i) = \sum_i w_\Omega(i)\left[\hat{n}(i,\alpha) + \hat{n}(i,\beta)\right] \ . \tag{9}$$

If we take multiple domains into consideration, we can define domain operators associated with a d domain *partition* $\boldsymbol{\Omega} = (\Omega_1, \ldots, \Omega_d)$, which consists of d domains

$$\forall a \leq d : \hat{n}(\Omega_a) = \sum_i w_{\Omega_a}(i)\hat{n}(i) \ , \tag{10}$$

where the following relations are obeyed

$$\sum_{a=1}^d w_{\Omega_a}(i) = 1 \tag{11}$$

$$w_{\Omega_a}(i) \geq 0 \ . \tag{12}$$

As such, a domain partition ensures that every orbital is included in at least one domain. We note that even if we only take one domain Ω in account, we are in fact describing a domain partition as we also implicitly consider the complement $\overline{\Omega}$ of the domain Ω.

An important subclass of these domain operators are those for which the weights are equal to Dirac delta type functions, such as

$$\hat{n}(\Omega_a) = \sum_I \delta(I \in \Omega_a)\hat{n}(I) \ . \tag{13}$$

An application of these Dirac delta type domains is that of the non-orthogonal Mulliken partitioning scheme. In this scheme, a domain is associated with a given nucleus A and

contains all spin-orbitals centered on that nucleus

$$\hat{n}(\Omega_A) = \sum_{IJ} \delta(I \in \Omega_A)\hat{n}(\chi_I)\mathbf{\Sigma}_{IJ} \ . \tag{14}$$

In spin-position space, Dirac delta type weight functions $w_{\Omega_a}(i) = \delta(i \in \Omega_a)$ lead to the domain operators proposed by Ziesche (Ziesche, 2000)

$$\hat{n}(\Omega) = \sum_{\sigma} \int \delta(\boldsymbol{r} \in \Omega)\hat{\psi}^+(\boldsymbol{r},\sigma)\hat{\psi}^-(\boldsymbol{r},\sigma)\mathrm{d}\boldsymbol{r} \ , \tag{15}$$

which are the second quantized analogs of the population operators as introduced by Diner and Claverie (Diner & Claverie, 1975). The domains obtained by topological analysis of the electron density (the 'Quantum Theory of Atoms in Molecules') (R. F. Bader, 1990; Popelier, 2000; Matta & Boyd, 2007) belong to this category, as do the domains obtained from a topological analysis of any other quantum chemical scalar field such as the 'Electron Localization Function' (Silvi, 2004).

Other partitions are obtained by optimizing the shape of domains with respect to a certain objective function. In the 'loge' theory, developed by Daudel and coworkers (Aslangul, Constanciel, Daudel, & Kottis, 1972), that objective function is a missing information function (Aslangul, Constanciel, Daudel, Esnault, & Ludena, 1974). In the theory of 'Maximum Probability Domains', devised by Savin (Savin, 2001), the resulting domains maximize the probability to find ν and only ν electrons inside themselves.

A domain partition that deviates from these Dirac delta type weights is the Hirshfeld domain partition (Hirshfeld, 1977; Bultinck, Van Alsenoy, Ayers, & Carbó-Dorca, 2007). In this partition, the weights $w_{\Omega_a}(\boldsymbol{r})$ are determined by reference to a promolecule constructed from superimposing the isolated density functions ρ_A^{pro} of the M atoms that comprise the molecule

$$w_{\Omega_a}(\boldsymbol{r}) = \frac{\rho_A^{\mathrm{pro}}(\boldsymbol{r})}{\sum_M \rho_M^{\mathrm{pro}}(\boldsymbol{r})} \ . \tag{16}$$

As in principle the weight function is non-zero over the entire position space, such domain partitions are also called 'fuzzy'.

ORBITAL AND DOMAIN DENSITIES

Under the influence of the molecular wave function, the domains formulated above act as 'Open Quantum Systems' (Breuer, Petruccione, et al., 2002; A. M. Pendás & Francisco, 2018b). These open systems interact with an external environment, which can be its complement or a collection of other domains. As the wave function is in general not an eigenfunction of the corresponding domain occupation number operator, this forces the occupation to fluctuate randomly. As such, we have to resort to a statistical analysis of this occupation number operator. In this section, we focus on the average values of these operators, as they will determine the resulting average occupations or densities.

Orbital densities

In the theory of statistics, the first moment (also called the average or the expectation value) $\langle X \rangle$ of a stochastic variable X is given by

$$\langle X \rangle := \int x P(x) \mathrm{d}x \ , \tag{17}$$

and the d-th order moment by

$$\langle X^d \rangle := \int x^d P(x) \mathrm{d}x \ , \tag{18}$$

where $P(X)$ is the continuous probability distribution associated with the stochastic variable X.

The first moment $\langle \hat{n}(I) \rangle$ of the occupation-number operator $\hat{n}(I) = \hat{\phi}_I^+ \hat{\phi}_I^-$ is also called the '1st order reduced density' $\rho^{(1)}(I)$ or 'electron density' of the orbital $|\phi_I\rangle$. Not only does this descriptor form the main crux behind Density Functional Theory (Parr & Yang, 1989; Engel & Dreizler, 2011), its response to changes in number of electrons and/or external potential forms the basis of Conceptual Density Functional Theory (Geerlings, De Proft, & Langenaeker, 2003) and its analysis in terms of itself and its gradient leads to descriptors such as the Electron Localization Function (Savin, Nesper, Wengert, & Fässler, 1997; Silvi & Savin, 1994), the Electron Localizability Indicator (Kohout, 2004), the Non-Covalent Interactions index (Johnson et al., 2010) and the Density Overlap Region Indicator (De Silva & Corminboeuf, 2014).

As shown by Kong and Valeev (Kong & Valeev, 2011) and Hanauer and Köhn (Hanauer & Köhn, 2012), we can assign an interpretation to these moments of occupation-number operators, by first considering the moment of the orbital occupation-number operator $\hat{n}(I)$ for a configuration interaction expansion $|\Psi\rangle = \sum_{\boldsymbol{k}} c_{\boldsymbol{k}} |\boldsymbol{k}\rangle$, where $\boldsymbol{k}$ denotes an antisymmetric many-particle state. The electron density of $|\phi_I\rangle$ is then given by

$$\rho^{(1)}(I) = \langle\hat{n}(I)\rangle = \sum_{\boldsymbol{k},\boldsymbol{l}} c_{\boldsymbol{k}}^* c_{\boldsymbol{l}} \langle\boldsymbol{k}\,|\,\hat{n}(I)\,|\,\boldsymbol{l}\rangle = \sum_{\boldsymbol{k}} |c_{\boldsymbol{k}}|^2 \langle\boldsymbol{k}\,|\,\hat{n}(I)\,|\,\boldsymbol{k}\rangle = \sum_{\boldsymbol{k}} k_I |c_{\boldsymbol{k}}|^2 . \tag{19}$$

As $|c_{\boldsymbol{k}}|^2$ is the statistical weight or probability of the single Slater determinant $|\boldsymbol{k}\rangle$ (with $\sum_{\boldsymbol{k}} |c_{\boldsymbol{k}}|^2 = 1$), $\rho^{(1)}(I)$ represents the average occupation-number of orbital ϕ_I. As pointed out by Hanauer (Hanauer & Köhn, 2012), this is equivalent to regarding the occupation numbers k_I as stochastic variables, which can only take two values, one or zero. We can attach the same interpretation to $\rho^{(1)}(\boldsymbol{r})$: for the orbital $|\boldsymbol{r}\rangle$, $\rho^{(1)}(\boldsymbol{r})$ represents the average occupation of that position-space orbital. The corresponding d-th moment of the occupation-number operator can be used to construct the d-th order reduced densities

$$\begin{aligned}
\rho^{(d)}(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_d) &= \left\langle \hat{\mathcal{N}}\hat{n}(\boldsymbol{r}_1)\ldots\hat{n}(\boldsymbol{r}_d) \right\rangle \\
&= \left\langle \hat{\mathcal{N}} \prod_{i=1}^{d} \hat{n}(\boldsymbol{r}_i) \right\rangle \\
&= \left\langle \hat{\psi}^+(\boldsymbol{r}_1)\ldots\hat{\psi}^+(\boldsymbol{r}_d)\hat{\psi}^-(\boldsymbol{r}_d)\ldots\hat{\psi}^-(\boldsymbol{r}_1) \right\rangle ,
\end{aligned} \tag{20}$$

where we have used the operator $\hat{\mathcal{N}}$ ensure normal order (Ziesche, 2000) (see appendix for more details). This normal operator ensures that products of d occupation-number operators actually describe simultaneous occupation of d orbitals. Indeed, this normal-ordering operator allows us to reorder the respective creation and annihilation operators to the form needed by the d-th order density. At the same time, this normal ordering also removes same site elements

$$\hat{\mathcal{N}}\hat{n}(\boldsymbol{r})\hat{n}(\boldsymbol{r}) |\Psi\rangle = \hat{\psi}^+(\boldsymbol{r})\hat{\psi}^+(\boldsymbol{r})\hat{\psi}^-(\boldsymbol{r})\hat{\psi}^-(\boldsymbol{r}) |\Psi\rangle = 0 , \tag{21}$$

instead of allowing for effective removal of occupation-number operators due to their idempotency

$$\hat{n}(\boldsymbol{r})\hat{n}(\boldsymbol{r}) |\Psi\rangle = \hat{n}(\boldsymbol{r}) |\Psi\rangle . \tag{22}$$

As such, the respective d-th order reduced densities describe the average occupation-number of a set of d orbitals (Kong & Valeev, 2011)

$$\rho^{(d)}(I, \ldots, L) = \left\langle \hat{\mathcal{N}} \underbrace{\hat{n}(I) \ldots \hat{n}(L)}_{d \text{ terms}} \right\rangle = \sum_{\boldsymbol{k}} \underbrace{k_I \ldots k_L}_{d \text{ terms}} |c_{\boldsymbol{k}}|^2 . \tag{23}$$

We note that this normal operator often is neglected (Diner & Claverie, 1975; Silvi, 2004). This introduces spurious single orbital occupation number operators $\hat{n}(\boldsymbol{r}_1)$

$$\hat{n}(\boldsymbol{r}_1)\hat{n}(\boldsymbol{r}_2) = \hat{\mathcal{N}}\hat{n}(\boldsymbol{r}_1)\hat{n}(\boldsymbol{r}_2) + \delta(\boldsymbol{r}_1 - \boldsymbol{r}_2)\hat{n}(\boldsymbol{r}_1) , \tag{24}$$

for which one has to correct if one is to describe occupancy statistics. The far-reaching conceptual problems associated with introducing such 'unphysical self-pairing' have been discussed in detail by Kutzelnigg (Kutzelnigg, 2003).

Starting from the N-th order reduced density $\rho^{(N)}$, we can obtain lower order densities by taking normalized partial traces

$$\begin{smallmatrix} p \\ q \end{smallmatrix}\!\downarrow := \frac{(N-p)!}{(N-q)!} \text{tr}_{q+1,\ldots,p} , \tag{25}$$

such that

$$\begin{smallmatrix} p \\ q \end{smallmatrix}\!\downarrow\rho^{(p)}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_q, \ldots, \boldsymbol{r}_p) = \rho^{(q)}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_q) . \tag{26}$$

This corresponds to the McWeeny normalization convention (McWeeny, 1960), where each d-th order reduced density is normalized to $\frac{N!}{(N-d)!}$. Note the contrast with the Löwdin normalization convention (Löwdin, 1955c, 1955b, 1955a), where the p-th order density matrix is normalized to the number of unique combinations $\binom{N}{p}$. The latter convention leads to an easier notation for energetic expectation values, but is less suited for elucidating the underlying statistics. Hence, we will exclusively use the McWeeny normalization convention in this review.

Domain densities

In the case of the domain occupation-number operator $\hat{n}(\Omega_a)$, the first moment $\rho^{(1)}(\Omega_a)$ is given by

$$\rho^{(1)}(\Omega_a) = \langle \hat{n}(\Omega_a) \rangle = \int w_{\Omega_a}(\boldsymbol{r})\rho^{(1)}(\boldsymbol{r})\mathrm{d}\boldsymbol{r} , \tag{27}$$

and is equal to the average number of electrons in the domain Ω_a (also called its population). The second moments $\rho^{(2)}(\Omega_a, \Omega_a)$ and $\rho^{(2)}(\Omega_a, \Omega_b)$ are then given by

$$\rho^{(2)}(\Omega_a, \Omega_a) = \left\langle \hat{\mathcal{N}} \hat{n}(\Omega_a)^2 \right\rangle = \int \int w_{\Omega_a}(\boldsymbol{r}_1) w_{\Omega_a}(\boldsymbol{r}_2) \rho^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) \mathrm{d}\boldsymbol{r}_1 \mathrm{d}\boldsymbol{r}_2 \tag{28}$$

$$\rho^{(2)}(\Omega_a, \Omega_b) = \left\langle \hat{\mathcal{N}} \hat{n}(\Omega_a) \hat{n}(\Omega_b) \right\rangle = \int \int w_{\Omega_a}(\boldsymbol{r}_1) w_{\Omega_b}(\boldsymbol{r}_2) \rho^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) \mathrm{d}\boldsymbol{r}_1 \mathrm{d}\boldsymbol{r}_2 \,, \tag{29}$$

where $\rho^{(2)}(\Omega_a, \Omega_a)$ is the average number of electron pairs in Ω_a and $\rho^{(2)}(\Omega_a, \Omega_b)$ is the average number of pairs that can be formed between the electrons in Ω_a and Ω_b (R. Bader & Stephens, 1975).

We note that there exists a widespread confusion between between reduced densities and probabilities, which leads to questionable interpretations when generalized to domains. Indeed, McWeeny (McWeeny, 1960) states that

> "$\rho^{(1)}(\boldsymbol{x}_1)$ is the probability of finding a particle with variables in the range $\mathrm{d}\boldsymbol{x}_1$ at point $\boldsymbol{x}_1$ in configuration space."

although he immediately adds

> "It should be noted that $\rho^{(1)}(\boldsymbol{x}_1)$ integrates to N (not 1), and that it is therefore the 'number density'."

As pointed out by Diner and Claverie (Claverie & Diner, 1976) and restated by Savin (Savin, 2004) many years later, this distinction becomes particularly relevant when considering probabilities defined over domains:

> "The belief that the probability of finding one electron in a given region Ω, $p(1, N-1)$, is given by the population of this region, or the average number of electrons in Ω, $\rho^{(1)}(\Omega)$ is wrong, however, when the total number of electrons in the system, $N > 1$. This can be seen most easily when considering the case when Ω corresponds to the whole space yielding $\int \rho^{(1)}(\boldsymbol{r}) \mathrm{d}\boldsymbol{r} = N$; this is impossible as the probability must lie between 0 and 1."

Quantities $p^{(d)}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_d)$ that integrate to one *can* be obtained by forcing normalization upon $\rho^{(d)}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_d)$

$$p^{(d)}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_d) = \frac{(N-d)!}{N!} \rho^{(d)}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_d) . \tag{30}$$

These normalized quantities should then be interpreted as the probability that *at least d* electrons are at the positions $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_d$ and *not* as the probability that d electrons are at the positions $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_d$. However, in most cases these normalization factor is neglected. As such, contrary to widespread perception,

$$\rho^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \rho^{(1)}(\boldsymbol{r}_1)\rho^{(1)}(\boldsymbol{r}_2) \tag{31}$$

does *not* indicate that the can constituent electrons be considered as statistically independent (Kutzelnigg, 2003). Furthermore, as we will show later, even if this normalization is taken into account, the statistics of these normalized quantities are frequently dealt with too freely, which results in questionable interpretations.

Reduced density generating function

We can define the moment generating function $\mathcal{M}(t)$ of the stochastic variable X as the moment of the stochastic variable e^{tX} with $t \in \mathbb{R}$

$$\mathcal{M}(t) := \langle e^{tX} \rangle = \left\langle 1 + tX + \frac{t^2 X^2}{2!} + \ldots \right\rangle . \tag{32}$$

Since the generating function gathers all moments as coefficients of a series expansion, all orders of moments can be obtained by successive differentiation of this generating function

$$\langle X^n \rangle = \lim_{t \to 0} \frac{\mathrm{d}^n \mathcal{M}(t)}{\mathrm{d}t^n} . \tag{33}$$

The variable t functions as a 'probe' variable, in the sense that this variable only appears as a dummy variable for the differentiation and does not appear in the final expression for the moments, which is obtained formally in the limit of t goes to zero.

By considering occupation numbers as random variables (Hanauer & Köhn, 2012), we can construct an associated generating function. If we take d domains into account, this

11

generating function is given by (Ziesche, 2000)

$$\mathcal{M}[g, \boldsymbol{t}] := \left\langle \hat{\mathcal{N}} \exp \left(\int g(\boldsymbol{r}) \hat{n}(\boldsymbol{r}) \mathrm{d}\boldsymbol{r} + \sum_{a=1}^{d} t_a \hat{n}(\Omega_a) \right) \right\rangle . \tag{34}$$

As such, the partial derivative with respect to $g(\boldsymbol{r}_1)$ leads to the first-order reduced density at $|\boldsymbol{r}_1\rangle$

$$\lim_{\substack{g \to 0 \\ \boldsymbol{t} \to \boldsymbol{0}}} \frac{\partial \mathcal{M}[g, \boldsymbol{t}]}{\partial g(\boldsymbol{r}_1)} = \lim_{\substack{g \to 0 \\ \boldsymbol{t} \to \boldsymbol{0}}} \left\langle \hat{\mathcal{N}} \left(\int \hat{n}(\boldsymbol{r}) \delta(\boldsymbol{r} - \boldsymbol{r}_1) \mathrm{d}\boldsymbol{r} \right) \exp \left(\int g(\boldsymbol{r}) \hat{n}(\boldsymbol{r}) \mathrm{d}\boldsymbol{r} + \sum_{a=1}^{d} t_a \hat{n}(\Omega_a) \right) \right\rangle$$

$$= \left\langle \hat{\mathcal{N}} \hat{n}(\boldsymbol{r}_1) \right\rangle = \rho^{(1)}(\boldsymbol{r}_1) , \tag{35}$$

where $g \to 0$ is used to denote $\forall \boldsymbol{r} : g(\boldsymbol{r}) \to 0$. In general, we can obtain the d-th order reduced density $\rho^{(d)}$ by successive differentiation

$$\rho^{(d)}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_d) = \lim_{\substack{g \to 0 \\ \boldsymbol{t} \to \boldsymbol{0}}} \frac{\partial^d \mathcal{M}[g, \boldsymbol{t}]}{\partial g(\boldsymbol{r}_d) \ldots \partial g(\boldsymbol{r}_1)} . \tag{36}$$

In the case of domains, taking the partial derivative with respect to t_a leads to moments associated with domains. As such, the first partial derivative leads to the population of domain Ω_a

$$\lim_{\substack{g \to 0 \\ \boldsymbol{t} \to \boldsymbol{0}}} \frac{\partial \mathcal{M}[g, \boldsymbol{t}]}{\partial t_a} = \lim_{\substack{g \to 0 \\ \boldsymbol{t} \to \boldsymbol{0}}} \left\langle \hat{\mathcal{N}} \hat{n}(\Omega_a) \right\rangle = \rho^{(1)}(\Omega_a) . \tag{37}$$

Taking two partial derivatives leads to average number of pairs inside domains

$$\lim_{\substack{g \to 0 \\ \boldsymbol{t} \to \boldsymbol{0}}} \frac{\partial^2 \mathcal{M}[g, \boldsymbol{t}]}{\partial t_a^2} = \left\langle \hat{\mathcal{N}} \hat{n}(\Omega_a)^2 \right\rangle = \rho^{(2)}(\Omega_a, \Omega_a) \tag{38}$$

$$\lim_{\substack{g \to 0 \\ \boldsymbol{t} \to \boldsymbol{0}}} \frac{\partial^2 \mathcal{M}[g, \boldsymbol{t}]}{\partial t_b \partial t_a} = \left\langle \hat{\mathcal{N}} \hat{n}(\Omega_a) \hat{n}(\Omega_b) \right\rangle = \rho^{(2)}(\Omega_a, \Omega_b) . \tag{39}$$

In general, we can obtain the d-th order reduced density $\rho^{(d)}$ over domains by successive differentiation

$$\rho^{(d)}(\Omega_1, \ldots, \Omega_d) = \lim_{\substack{g \to 0 \\ \boldsymbol{t} \to \boldsymbol{0}}} \frac{\partial^d \mathcal{M}[g, \boldsymbol{t}]}{\partial t_1 \ldots \partial t_d} . \tag{40}$$

We can also define an average over k-sites and l-domains by taking the following partial derivatives

$$\rho^{(d)}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_k, \Omega_1, \ldots, \Omega_d) = \lim_{\substack{g \to 0 \\ \boldsymbol{t} \to \boldsymbol{0}}} \frac{\partial^{k+l} \mathcal{M}[g, \boldsymbol{t}]}{\partial g(\boldsymbol{r}_1) \ldots \partial g(\boldsymbol{r}_k) \partial t_1 \ldots \partial t_l} . \tag{41}$$

In position space, these quantities have been coined the 'Coarse-Grained Densities' (A. M. Pendás, Francisco, & Blanco, 2007b), where the indices that are not associated with domains are also called 'free'.

We again reiterate the importance of including normal ordering into the definitions of the respective moments. If we do not use the $\hat{\mathcal{N}}$ operator to project out the diagonal elements (Silvi, 2004; Poater, Duran, Solá, & Silvi, 2005), we are not always describing situations where we are investigating the same number of sites. Indeed, the corresponding moment generating function is then given by

$$\mathcal{M}_{\hat{\mathcal{N}}}[g, \boldsymbol{t}] = \left\langle \exp\left(\int g(\boldsymbol{r})\hat{n}(\boldsymbol{r})\mathrm{d}\boldsymbol{r} + \sum_{a=1}^{d} t_a \hat{n}(\Omega_a)\right)\right\rangle , \tag{42}$$

so the second order domain averaged density becomes

$$\lim_{\substack{g\to 0 \\ \boldsymbol{t}\to \boldsymbol{0}}} \frac{\partial^2 \mathcal{M}_{\hat{\mathcal{N}}}[g, \boldsymbol{t}]}{\partial t_a \partial t_b} = \langle \hat{n}(\Omega_a)\hat{n}(\Omega_b)\rangle = \int\int w_{\Omega_a}(\boldsymbol{x}_1)w_{\Omega_b}(\boldsymbol{x}_2)\langle \hat{n}(\boldsymbol{x}_1)\hat{n}(\boldsymbol{x}_2)\rangle\,\mathrm{d}\boldsymbol{x}_1\mathrm{d}\boldsymbol{x}_2$$

$$= \int\int w_{\Omega_a}(\boldsymbol{x}_1)w_{\Omega_b}(\boldsymbol{x}_2)\left\langle \hat{\psi}^+(\boldsymbol{x}_1)\hat{\psi}^-(\boldsymbol{x}_1)\hat{\psi}^+(\boldsymbol{x}_2)\hat{\psi}^-(\boldsymbol{x}_2)\right\rangle\,\mathrm{d}\boldsymbol{x}_1\mathrm{d}\boldsymbol{x}_2$$

$$= \int\int w_{\Omega_a}(\boldsymbol{x}_1)w_{\Omega_b}(\boldsymbol{x}_2)$$

$$\left\langle \hat{\psi}^+(\boldsymbol{x}_1)\hat{\psi}^+(\boldsymbol{x}_2)\hat{\psi}^-(\boldsymbol{x}_2)\hat{\psi}^-(\boldsymbol{x}_2) - \delta(\boldsymbol{x}_1 - \boldsymbol{x}_2)\hat{\psi}^+(\boldsymbol{x}_1)\hat{\psi}^-(\boldsymbol{x}_2)\right\rangle\,\mathrm{d}\boldsymbol{x}_1\mathrm{d}\boldsymbol{x}_2$$

$$= \lim_{\substack{g\to 0 \\ \boldsymbol{t}\to \boldsymbol{0}}} \frac{\partial^2 \mathcal{M}[\boldsymbol{t}]}{\partial t_a \partial t_b} - \int w_{\Omega_a}(\boldsymbol{x}_1)w_{\Omega_b}(\boldsymbol{x}_1)\rho^{(1)}(\boldsymbol{x}_1)\mathrm{d}\boldsymbol{x}_1 . \tag{43}$$

Only in the case where an idempotent (Dirac delta type) weight factor is used, does this reduce to the form proposed by Silvi (Silvi, 2004)

$$\langle \hat{n}(\Omega_a)\hat{n}(\Omega_b)\rangle = \left\langle \hat{\mathcal{N}}\hat{n}(\Omega_a)\hat{n}(\Omega_b)\right\rangle + \langle \delta_{ab}\hat{n}(\Omega_a)\rangle . \tag{44}$$

From a statistical point of view, the addition of this spurious one-site operator has little to do with fixing two site occupations and is a direct consequence of not imposing normal ordering on the operator string.

As all moments can be generated from $\mathcal{M}[g, \boldsymbol{t}]$, this generating function is forms the basis for many quantum chemical descriptors. Successive differentiations of the reduced density generating function generate a hierarchy of descriptors that characterize the average occupations of domains and orbitals (see Figure 2). On the bottom layer of this hierarchy
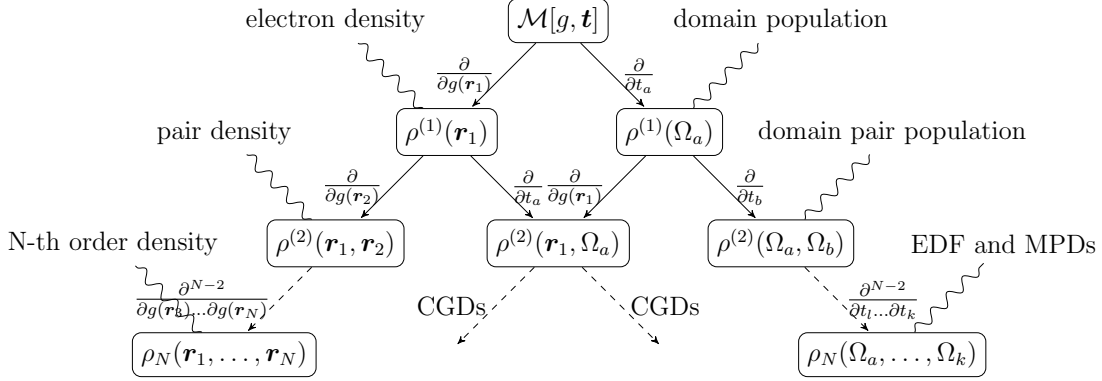
13

Figure 2: Overview of descriptors derived from the reduced density generating functional $\mathcal{M}[p, \boldsymbol{d}]$. Arrows indicate successive differentiations of that generating function.

we can find quantities where there are as many occupation number operators as there are electrons. As such, these quantities can be normalized to probabilities of finding all electrons at certain positions or domains, whereas quantities on a higher level can only be interpreted in terms of finding *at least* a given number of electrons. As the former descriptors have known significant advancements in recent years, we will discuss these probabilities further in the next section.

DOMAIN PROBABILITIES

In the previous section, we defined the probability of finding *at least* d electrons in a domain. In this section, we define the probability of finding *exactly* d electrons in a domain or a partition of electrons $(\nu_1, \ldots, \nu_d)$ over a partition of domains. Such partitions can be specified as

$$[\boldsymbol{\nu}, \boldsymbol{\Omega}] = [(\nu_1, \ldots, \nu_d), (\Omega_1, \ldots, \Omega_d)] \ . \tag{45}$$

The collection of all probabilities $p[(\nu_1, \ldots, \nu_d), \boldsymbol{\Omega}]$ associated with a domain partition $\boldsymbol{\Omega}$ has been called the 'Electron number Distribution Function' or EDF. This EDF has a rich history in conceptual quantum chemistry. Already in the fifties, Daudel and coworkers tried to 'localize' electrons - that are physically inherently delocalized - by deforming domains (Daudel, Brion, & Odiot, 1955; Daudel, Bader, Stephens, & Borrett, 1974). Daudel argued that in view of the success of Lewis's model, there should be some molecule-dependent 'best'

decomposition in regions (also called 'loges'), where the resulting regions would have a high degree of similarity with the structures that followed from Lewis's theory. He proposed a minimization of the missing information function I by deforming the domain Ω (and necessarily, its complement $\overline{\Omega}$)

$$I(\nu, \Omega) = -\sum_{\nu} p[(\nu, N - \nu), (\Omega, \overline{\Omega})] \ln p[(\nu, N - \nu), (\Omega, \overline{\Omega})] \ , \qquad (46)$$

The best 'loges', meaning those domains that minimize the missing information function, exhibit a high degree of similarity with the structures that follow from Lewis's theory (Daudel et al., 1955). Unfortunately, these partitions are very difficult to obtain computationally (Aslangul et al., 1972) and the method remains applicable only to the smallest of model systems.

However, many years later Savin (Savin, 2001) was able to show that chemical interpretations can be given to a *single* domain whose shape has been optimized so that the probability of finding ν *and only* ν electrons inside it (with ν typically chosen to be equal to two, as a reference to the electron pair proposed by Lewis) is *maximal.* The resulting domains were coined 'Maximum Probability Domains' or MPDs (Savin, 2001). Thanks to further theoretical work by Cancès (Cancès, Keriven, Lodier, & Savin, 2004), obtaining MPDs in position space for general quantum chemical methods became a possibility. Unfortunately, the development of actual implementations has been severely hampered by the sheer complexity of the shape optimization algorithms needed for such a program. Despite numerous efforts (Scemama, Caffarel, & Savin, 2007; Causa & Savin, 2011; Causà & Savin, 2011; Lopes Jr, Braïda, Causa, & Savin, 2012; Causa, DAmore, Garzillo, Gentile, & Savin, 2013; Menéndez & Pendás, 2014; Menéndez, Pendás, Braïda, & Savin, 2014; Agostini, Ciccotti, Savin, & Vuilleumier, 2015; Causà, D'Amore, Gentile, Menéndez, & Calatayud, 2015; Acke et al., 2016; Turek, Braïda, & De Proft, 2017), the question of whether MPDs provide a mathematical bridge between chemical concepts and quantum mechanics remains unanswered to this day.

Simultaneously, the interpretative power behind the EDF for a *given* set of domains $\mathbf{\Omega}$ has been used to provide an probabilistic view of chemical bonds in position space (Savin, 2001; Cancès et al., 2004; Savin, 2005; Gallegos, Carbó-Dorca, Lodier, Cancès, & Savin,

2005; Bochicchio, Ponec, Lain, & Torre, 2000; Bochicchio, Ponec, Torre, & Lain, 2001; A. M. Pendás et al., 2007b; Francisco, Pendás, & Blanco, 2007; A. M. Pendás, Francisco, & Blanco, 2007c, 2007a; Á. M. Pendás, Francisco, & Blanco, 2007; Francisco, Pendás, & Blanco, 2008, 2011; Francisco & Martín Pendás, 2014). This view is based on 'resonance structures', where each resonance structure is a localization of N electrons in position space domains. As such, concepts such as electron pair shairing, polarity, charge transfer, and multiple bonding have been shown to be recovered from the statistical properties of $p[\boldsymbol{\nu}, \boldsymbol{\Omega}]$. However, we must stress that this approach uses 'precomputed' domains and in no way optimises those domains to obtain optimal probabilistic descriptions.

Probability operators

We can define domain probabilities as those probabilities that can be obtained by associating *all* electrons with certain domains. If we focus exclusively on domains, the generating function $\mathcal{M}[g, \boldsymbol{t}]$ reduces to

$$\mathcal{M}[\boldsymbol{t}] := \left\langle \hat{\mathcal{N}} \exp \left(\sum_{a=1}^{d} t_a \hat{n}(\Omega_a) \right) \right\rangle . \tag{47}$$

The probabilities detailed above are then obtained by selecting only those terms containing N occupation-number operators from the generating function $\mathcal{M}[\boldsymbol{t}]$

$$\left\langle \hat{\mathcal{N}}_N \exp \left(\sum_{a=1}^{d} t_a \hat{n}(\Omega_a) \right) \right\rangle = \sum_{\substack{\nu_1, \ldots, \nu_d \\ \sum\limits_{a=1}^{d} \nu_a = N}} \left\langle \hat{\mathcal{N}} \prod_{a=1}^{d} \frac{t^{\nu_a}}{\nu_a!} \hat{n}(\Omega_a)^{\nu_a} \right\rangle$$

$$= \sum_{\nu \vdash N} \left\langle \hat{\mathcal{N}} \prod_{a=1}^{d} \frac{t^{\nu_a}}{\nu_a!} \hat{n}(\Omega_a)^{\nu_a} \right\rangle , \tag{48}$$

where we have used the $\nu \vdash N$ as a shorthand for the summation over all partitions of N electrons over the domains (Andrews, 1998). Based on the form of the generating function, we can premise that the probability operators $\hat{p}[(\nu_1, \ldots, \nu_d), \boldsymbol{\Omega}]$ associated with a domain partition $\boldsymbol{\Omega} = (\Omega_1, \ldots, \Omega_d)$ are given by

$$\hat{p}[(\nu_1, \ldots, \nu_d), \boldsymbol{\Omega}] := \hat{\mathcal{N}} \prod_{a=1}^{d} \frac{\hat{n}(\Omega_a)^{\nu_a}}{\nu_a!} , \tag{49}$$

and reformulate the moment generating function as

$$\left\langle \hat{\mathcal{N}}_N \exp\left(\sum_{a=1}^{d} t_a \hat{n}(\Omega_a)\right)\right\rangle = \sum_{\nu \vdash N} \langle \hat{p}\left[(\nu_1,\ldots,\nu_d),\boldsymbol{\Omega}\right]\rangle\, t_1^{\nu_1}\ldots t_d^{\nu_d}\ . \tag{50}$$

If we use the following notation for a multinomial coefficient

$$\binom{N}{\nu_1\ldots\nu_d} = \frac{N!}{\nu_1!\ldots\nu_d!}\ , \tag{51}$$

we can show that the moments of the probability operators $\hat{p}\left[(\nu_1,\ldots,\nu_d),\boldsymbol{\Omega}\right]$ are equivalent to the original probability expressions as used by Bader (R. Bader & Stephens, 1975) and Savin (Savin, 2001)

$$
\begin{aligned}
\langle \hat{p}\left[(\nu_1,\ldots,\nu_d),\boldsymbol{\Omega}\right]\rangle &= \left\langle \hat{\mathcal{N}} \prod_{a=1}^{d} \frac{\hat{n}(\Omega_a)^{\nu_a}}{\nu_a!}\right\rangle \\
&= \frac{1}{\nu_1!\ldots\nu_d!} \underbrace{\int\ldots\int}_{N} w_{\Omega_1}(\boldsymbol{x}_1)\ldots w_{\Omega_1}(\boldsymbol{x}_{\nu_1})\ldots \left\langle \hat{\mathcal{N}}\hat{n}(\boldsymbol{x}_1)\ldots\hat{n}(\boldsymbol{x}_N)\right\rangle \mathrm{d}\boldsymbol{x}_1\ldots\mathrm{d}\boldsymbol{x}_N \\
&= \binom{N}{\nu_1\ldots\nu_d} \underbrace{\int\ldots\int}_{N} w_{\Omega_1}(\boldsymbol{x}_1)\ldots w_{\Omega_1}(\boldsymbol{x}_{\nu_1})\ldots |\Psi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)|^2\, \mathrm{d}\boldsymbol{x}_1\ldots\mathrm{d}\boldsymbol{x}_N \\
&= p\left[(\nu_1,\ldots,\nu_d),\boldsymbol{\Omega}\right]\ . \tag{52}
\end{aligned}
$$

As such, our interpretation of $\hat{p}\left[(\nu_1,\ldots,\nu_d),\boldsymbol{\Omega}\right]$ as a domain probability operator is correct.

We note that, as we have defined the probability operators solely in terms of occupation-number operators, the framework derived above is valid for any underlying one-electron spin-basis, with the generating function $\mathcal{M}[\boldsymbol{t}]$ representing a compact codification of the information contained in the EDF. As such, the EDF framework is not restricted to the most commonly used position space perspective and can also be applied to e.g. the abstract site basis of the Hubbard model (Acke et al., 2016).

How to calculate domain probabilities

In spin-position space, the domain associated number operator (for a domain defined in position space) is given by

$$\hat{n}(\Omega_a) = \int w_{\Omega_a}(\boldsymbol{r})\hat{\psi}^+(\boldsymbol{x})\hat{\psi}^-(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \sum_{IJ}\left(\int \phi_I^*(\boldsymbol{x})w_{\Omega_a}(\boldsymbol{r})\phi_J(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\right)\hat{\phi}_I^+\hat{\phi}_J^-\ . \tag{53}$$

If we define the spin-orbital overlap matrix $\boldsymbol{\Sigma}^{\Omega_a}$ over the domain Ω_a as

$$\Sigma_{IJ}^{\Omega_a} := \int \phi_I^*(\boldsymbol{x}) w_{\Omega_a}(\boldsymbol{r}) \phi_J(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \ , \tag{54}$$

and a matrix of generalized occupation-number operators $\hat{\boldsymbol{n}}$ as

$$\hat{\boldsymbol{n}}_{IJ} = \hat{\phi}_J^+ \hat{\phi}_I^- \ , \tag{55}$$

then the domain associated number operators $\hat{n}(\Omega_a)$ can be expressed as

$$\hat{n}(\Omega_a) = \mathrm{tr}\left(\boldsymbol{\Sigma}^{\Omega_a} \hat{\boldsymbol{n}}\right) \ . \tag{56}$$

As such, we obtain the following form of the probability generating function

$$\sum_{\nu \vdash N} p\left[(\nu_1, \ldots, \nu_d), \boldsymbol{\Omega}\right] t_1^{\nu_1} \ldots t_d^{\nu_d} = \left\langle \hat{\mathcal{N}}_N \exp\left[\mathrm{tr}\left(\left(\sum_{a=1}^d t_a \boldsymbol{\Sigma}^{\Omega_a}\right) \hat{\boldsymbol{n}}\right)\right] \right\rangle \ . \tag{57}$$

Since the trace is linear and

$$\exp\left(\mathrm{tr}(\boldsymbol{X})\right) = \det\left(\exp(\boldsymbol{X})\right) \ , \tag{58}$$

we obtain that

$$
\begin{aligned}
\sum_{\nu \vdash N} p\left[(\nu_1, \ldots, \nu_d), \boldsymbol{\Omega}\right] t_1^{\nu_1} \ldots t_d^{\nu_d} &= \left\langle \hat{\mathcal{N}}_N \det\left[\exp\left(\left(\sum_{a=1}^d t_a \boldsymbol{\Sigma}^{\Omega_a}\right) \hat{\boldsymbol{n}}\right)\right] \right\rangle \\
&= \left\langle \hat{\mathcal{N}}_N \det\left[\sum_{k=0}^\infty \frac{1}{k!} \left(\sum_{a=1}^d t_a \boldsymbol{\Sigma}^{\Omega_a}\right)^k \hat{\boldsymbol{n}}^k\right] \right\rangle \\
&= \left\langle \hat{\mathcal{N}}_N \det\left[\boldsymbol{I} + \left(\sum_{a=1}^d t_a \boldsymbol{\Sigma}^{\Omega_a}\right) \hat{\boldsymbol{n}}\right] \right\rangle \ ,
\end{aligned}
\tag{59}
$$

where in the last step we have used the fact that the normal ordering destroys all operators in which powers of a certain occupation-number operator occur. The moment of the matrix $\hat{\boldsymbol{n}}$ is determined by which orbitals are occupied in the respective Slater determinants in which the wave function can be expanded. As such, if we want to compute the moment between two Slater determinants $|\boldsymbol{k}\rangle$ and $|\boldsymbol{l}\rangle$, the above expression reduces to

$$\sum_{\nu \vdash N} p^{\boldsymbol{kl}}\left[(\nu_1, \ldots, \nu_d), \boldsymbol{\Omega}\right] t_1^{\nu_1} \ldots t_d^{\nu_d} = \det\left(\sum_{a=1}^d t_a \boldsymbol{\Sigma}'^{\boldsymbol{kl}, \Omega_a}\right) \ , \tag{60}$$

where Σ'^{kl,Ω_a} contains only those domain overlaps between the orbitals that are occupied in the respective Slater determinants. Although this connection has previously been determined by an explicit Slater determinant decomposition in first quantization (Francisco et al., 2007, 2008; Francisco, Pendás, & Costales, 2014), the current derivation clearly illustrates the power of the generating function perspective. Indeed, the complicated explicit decomposition can be replaced by the easier second quantized algebraic rules, allowing one to focus more on the statistical ideas behind such a derivation, as will be shown in the case for two domains.

The special case of two domains

If one considers only one domain Ω, one is also implicitly taking its complement $\overline{\Omega}$ into account. As this situation allows for far-reaching simplifications, we focus on the theory behind partitions consisting of two domains in this section. If we have only two domains Ω_a and Ω_b, then necessarily Ω_b is the complement of Ω_a and the following relation is valid

$$\hat{n}(\Omega_a) = \hat{n} - \hat{n}(\Omega_b) \; . \tag{61}$$

If we presume that the wave function is an eigenfunction of the number operator with eigenvalue N

$$\hat{n} \left| \Psi \right\rangle = N \left| \Psi \right\rangle \; , \tag{62}$$

we can focus solely on the action of one domain associated number-operator, say $\hat{n}(\Omega_a)$, which we will write as $\hat{n}(\Omega)$ (with $\hat{n}(\Omega_b)$ then equal to $\hat{n}(\overline{\Omega})$). In this case, the generating function is given by

$$\hat{\mathcal{N}}_N \exp\left[t_\Omega \hat{n}_\Omega + t_{\overline{\Omega}} \hat{n}_{\overline{\Omega}}\right] = \sum_{\nu=0}^{N} \hat{p}\left[(\nu, N-\nu), (\Omega, \overline{\Omega})\right] t_\Omega^\nu t_{\overline{\Omega}}^{N-\nu} \; , \tag{63}$$

with

$$\hat{p}\left[(\nu, N-\nu), (\Omega, \overline{\Omega})\right] = \frac{1}{\nu!(N-\nu)!}\hat{\mathcal{N}}\hat{n}(\Omega)^\nu \hat{n}(\overline{\Omega})^{N-\nu} \; . \tag{64}$$

The m-th order domain condensed density matrix can be expressed in terms of the associated probability distribution as (Ziesche, 2000)

$$\rho^{(m)}(\Omega, \ldots, \Omega) = \sum_{\nu=m}^{N} \frac{\nu!}{(\nu-m)!} p\left[(\nu, N-\nu), (\Omega, \overline{\Omega})\right] \; . \tag{65}$$

In turn, the ν-th probability can be expressed in terms of the domain condensed density matrices (Ziesche, 2000)

$$
\begin{aligned}
p\left[(\nu, N-\nu),(\Omega,\overline{\Omega})\right] &= \left\langle \hat{p}\left[(\nu, N-\nu),(\Omega,\overline{\Omega})\right]\right\rangle \\
&= \left\langle \frac{1}{\nu!(N-\nu)!}\hat{\mathcal{N}}\hat{n}(\Omega)^{\nu}\left(\hat{n}-\hat{n}(\Omega)\right)^{N-\nu}\right\rangle \\
&= \left\langle \frac{1}{\nu!(N-\nu)!}\hat{\mathcal{N}}\sum_{k=0}^{N-\nu}\binom{N-\nu}{k}(-1)^{k}\hat{n}(\Omega)^{k+\nu}\hat{n}^{N-(k+\nu)}\right\rangle \\
&= \left\langle \frac{1}{\nu!}\hat{\mathcal{N}}\sum_{k=0}^{N-\nu}\frac{1}{k!(N-(k+\nu))!}(-1)^{k}\hat{n}(\Omega)^{k+\nu}\hat{n}^{N-(k+\nu)}\right\rangle \\
&= \left\langle \hat{\mathcal{N}}\sum_{k=\nu}^{N}\frac{1}{\nu!(k-\nu)!}(-1)^{k-\nu}\hat{n}(\Omega)^{k}\frac{1}{(N-k)!}\hat{n}^{N-k}\right\rangle \\
&= \left\langle \hat{\mathcal{N}}\sum_{k=\nu}^{N}\frac{1}{\nu!(k-\nu)!}(-1)^{k-\nu}\hat{n}(\Omega)^{k}\right\rangle \\
&= \sum_{k=\nu}^{N}\frac{(-1)^{k-\nu}}{\nu!(k-\nu)!}\rho^{(k)}(\Omega,\dots,\Omega)\ .
\end{aligned}
\tag{66}
$$

This equation shows how one can formally obtain $p\left[(\nu, N-\nu),(\Omega,\overline{\Omega})\right]$ by "eliminating those contributions that can be attributed to intrusion of more than ν particles in the domain Ω" (Savin, 2004). It was this connection that allowed Savin to indicate the feasibility of the concept of MPDs, by performing probability calculations for atoms (Savin, 2001).

However, note that the above relationship was already derived in the 70's in the context of fluctuation theorems for statistical mechanical purposes (equation 13b in (Vezzetti, 1975)). This derivation was based on the fact that the probability that a domain of "space" contains exactly ν particles can be expressed in terms of distribution functions (the reduced densities, also called Janossy probability densities (Daley & Vere-Jones, 2002)). A few years later, Truskett (Truskett, Torquato, & Debenedetti, 1998) realized that the underlying framework is essentially that of a 'stochastic spatial point process' (in our case 'stochastic site process' would be a better designation). A point process is a type of stochastic process, for which any one realization consists of a set of elements in the relevant space (in our case, the space of sites) (Daley & Vere-Jones, 2002, 2008). This interpretation will have important consequences for the single determinant case.

If there are only two domains Ω and $\overline{\Omega}$ and the one-electron spin-basis is orthonormal,

20

then

$$\mathbf{\Sigma}^{\Omega} + \mathbf{\Sigma}^{\overline{\Omega}} = \mathbf{I} \ , \tag{67}$$

from which follows that

$$\mathbf{\Sigma}^{\overline{\Omega}} = \mathbf{I} - \mathbf{\Sigma}^{\Omega} \ . \tag{68}$$

This connection will again allow to express all probabilities in terms of quantities that are only related to the domain Ω. Again, we can express the domain associated number operators in the one-electron spin-basis

$$\hat{n}(\Omega) = \int w_{\Omega}(\boldsymbol{r})\hat{\psi}^{+}(\boldsymbol{x})\hat{\psi}^{-}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \sum_{IJ}^{N} \left(\int \phi_I^*(\boldsymbol{x})w_{\Omega}(\boldsymbol{x})\phi_J(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \right) \hat{\phi}_I^+ \hat{\phi}_J^-$$

$$= \sum_{IJ}^{N} \hat{\phi}_I^+ \mathbf{\Sigma}_{IJ}^{\Omega} \hat{\phi}_J^- \ , \tag{69}$$

$$\hat{n}(\overline{\Omega}) = \sum_{IJ}^{N} \hat{\phi}_I^+ \left(\mathbf{I} - \mathbf{\Sigma}_{IJ}^{\Omega} \right) \hat{\phi}_J^- \ . \tag{70}$$

Given that $\mathbf{\Sigma}^{\Omega}$ is Hermitian, we can diagonalize the matrix

$$\mathbf{\Sigma}^{\Omega} = \boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^{\dagger} \tag{71}$$

and use the eigenvectors to rotate the creation and annihilation operators unitarily, both for the domain Ω

$$\hat{n}(\Omega) = \sum_{IJ}^{N} \hat{\phi}_I^+ (\boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^{\intercal})_{IJ}\hat{\phi}_J^- = \sum_{I=1}^{N} \lambda_I \hat{\tilde{\phi}}_I^+ \hat{\tilde{\phi}}_I^- = \sum_{I=1}^{N} \lambda_I \hat{n}(\tilde{I}) \ , \tag{72}$$

and its complement $\overline{\Omega}$

$$\hat{n}(\overline{\Omega}) = \sum_{I=1}^{N} (1 - \lambda_I)\hat{n}(\tilde{I}) \ , \tag{73}$$

where we have used the notation $\tilde{I}$ to distinguish between the original basis and the rotated basis.

Hence, the probability generating function for two domains becomes

$$\sum_{\nu=0}^{N} p\left[(\nu, N - \nu), (\Omega, \overline{\Omega})\right] t_{\Omega}^{\nu} t_{\overline{\Omega}}^{N-\nu} = \left\langle \hat{\mathcal{N}}_N \exp \left(\sum_{I=1}^{N} [t_{\Omega}\lambda_I + t_{\overline{\Omega}}(1 - \lambda_I)]\, \hat{n}(\tilde{I}) \right) \right\rangle \ . \tag{74}$$

21

Because the $\hat{\mathcal{N}}_N$ operator automatically destroys operator strings with an occupation number operator count different than N, we can simplify this expression by setting $t_{\overline{\Omega}} = 1$ and $t_\Omega = t$

$$\sum_{\nu=0}^N p\left[(\nu, N-\nu),(\Omega,\overline{\Omega})\right] t^\nu = \left\langle \hat{\mathcal{N}}_N \exp\left(\sum_{I=1}^N [t\lambda_I + (1-\lambda_I)]\hat{n}(\tilde{I})\right)\right\rangle$$

$$= \left\langle \hat{\mathcal{N}}_N \prod_{I=1}^N \exp\left([(t-1)\lambda_I + 1]\hat{n}(\tilde{I})\right)\right\rangle$$

$$= \left\langle \hat{\mathcal{N}} \prod_{I=1}^N \left([(t-1)\lambda_I + 1]\hat{n}(\tilde{I})\right)\right\rangle \tag{75}$$

In the last expression, we have used the fact that only first-order terms of the exponential can survive due to the action of $\hat{\mathcal{N}}_N$ (since this operator removes diagonal elements and only retains terms with in total N occupation-number operators). As the occupation numbers of the molecular orbitals of a single Slater determinant are invariant under unitary rotations, we obtain the following for a single Slater determinant

$$\sum_{\nu=0}^N p\left[(\nu, N-\nu),(\Omega,\overline{\Omega})\right] t^\nu = \prod_{I=1}^N [(t-1)\lambda_I + 1]$$

$$= \det[\boldsymbol{I} + (t-1)\boldsymbol{\Lambda}]$$

$$= \det(\boldsymbol{X})\det[\boldsymbol{I} + (t-1)\boldsymbol{\Lambda}]\det(\boldsymbol{X}^\dagger)$$

$$= \det\left[\boldsymbol{I} + (t-1)\boldsymbol{\Sigma}^\Omega\right] . \tag{76}$$

This result is equivalent to what was obtained by an explicit determinant decomposition by Cancès (Cancès et al., 2004). Again, in our approach no such decomposition is needed and the link with the original work of Ziesche (Ziesche, 2000) that allowed Savin to propose the theory of MPDs (Savin, 2001) is clear.

Note that we can interpret the associated eigenvalues by recognizing the domains as 'Open Quantum Systems' (Breuer et al., 2002) and adapting the terminology of the 'Density Matrix Embedding Theory' or DMET (Knizia & Chan, 2013). In this theory, for a single Slater determinant, the overlap matrix $\boldsymbol{\Sigma}^\Omega$, which is defined as the overlap of the orbitals *projected onto the sites of the fragment* Ω, is diagonalized. The eigenvectors of $\boldsymbol{\Sigma}^\Omega$ define a rotation of the orbitals that divides them in three sets, depending on their eigenvalues. Eigenvalues of one and zero are associated with "pure fragment orbitals" (in our case "pure domain")

and "pure environment" (in our case "pure complementary domain") orbitals respectively. Eigenvalues in between zero and one are called "entangled orbitals" (in our case "shared orbitals").

In the paper by Cancès mentioned above (Cancès et al., 2004), it was also noted that the probabilities can be expressed as

$$p\left[(\nu, N - \nu), (\Omega, \overline{\Omega})\right] = a_\nu^N \ , \tag{77}$$

where the a_ν^N are defined by the recursion

$$\begin{cases} a_0^0 = 1 \\ a_0^k = \alpha_k a_0^{k-1} \\ a_j^k = \beta_k a_{j-1}^{k-1} + \alpha_k a_j^{k-1}, 1 \le j < k - 1 \\ a_k^k = \beta_k a_{k-1}^{k-1} \end{cases} \tag{78}$$

with

$$\alpha_k = 1 - \lambda_k \tag{79}$$

$$\beta_k = \lambda_k \ . \tag{80}$$

On the one hand, this allows for an efficient computation of the probabilities in the case of a single Slater determinant and two domains. On the other hand, this recursive scheme indicates that in the case of a single determinant wave function, the underlying framework is that of a determinantal point process (Macchi, 1975; Soshnikov, 2000). In a determinantal point process, the distribution of the number of points is equal to the distribution of the sum of independent Bernoulli(λ_j) random variables (a Poisson binomial distribution), where $0 < \lambda_j \le 1$ are the nonzero eigenvalues of the "kernel" of the determinantal process (in our case $\boldsymbol{\Sigma}^\Omega$) (Gottlieb, 2005).

Note that the interpretation as a determinantal point process is different from the interpretation of (Francisco, Pendás, & Blanco, 2009), where each electron is considered statistically independent from the others, lying either inside Ω or its complement with respective probabilities λ_k or $1 - \lambda_k$. The latter interpretation is at odds with the standard interpretation of statistical independence as a factorization of the probability density $|\Psi(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N)|^2$,

which can never be achieved when using a Slater determinant (Kutzelnigg, 2003). Indeed, according to the interpretation presented above, the statistical independence is related to the process of counting the number of particles in a domain, and not to the behavior of the particles themselves in spin-position space.

ORBITAL AND DOMAIN DENSITY FLUCTUATIONS

In the previous section, we showed how the moment generating function can be used to generate moments of the (domain) occupation number operator. A 'central moment' is the moment of a random variable about its mean. We can center the density operator $\hat{n}(\boldsymbol{r})$ around its mean $\langle \hat{n}(\boldsymbol{r}) \rangle$ by defining the density fluctuation operator $\delta\hat{n}(\boldsymbol{r})$ (Yamasaki & Goddard, 1998; Yamasaki, Mainz, & Goddard, 2000; Fulde, 2012)

$$\delta\hat{n}(\boldsymbol{r}) := \hat{n}(\boldsymbol{r}) - \langle \hat{n}(\boldsymbol{r}) \rangle \ . \tag{81}$$

In analogy with the moment generating function $\mathcal{M}[g, \boldsymbol{t}]$, we can construct a 'central moment' generating function $\mathcal{M}'[g, \boldsymbol{t}]$

$$\mathcal{M}'[g, \boldsymbol{t}] := \left\langle \hat{\mathcal{N}} \exp\left(\int g(\boldsymbol{r})\delta\hat{n}(\boldsymbol{r})\mathrm{d}\boldsymbol{r} + \sum_{a=1}^{d} t_a \delta\hat{n}(\Omega_a) \right) \right\rangle , \tag{82}$$

where

$$\delta\hat{n}(\Omega_a) := \hat{n}(\Omega_a) - \langle \hat{n}(\Omega_a) \rangle \ . \tag{83}$$

The first three moments of $\mathcal{M}'$ with respect to $g(\boldsymbol{r})$ are given by

$$\lim_{\substack{g\to 0 \\ \boldsymbol{t}\to\boldsymbol{0}}} \frac{\partial \mathcal{M}'[g, \boldsymbol{r}]}{\partial g(\boldsymbol{r}_1)} = \langle \delta\hat{n}(\boldsymbol{r}) \rangle = \langle \hat{n}(\boldsymbol{r}) \rangle - \langle \hat{n}(\boldsymbol{r}) \rangle = 0 = \rho'^{(1)}(\boldsymbol{r}_1) \tag{84}$$

$$\lim_{\substack{g\to 0 \\ \boldsymbol{t}\to\boldsymbol{0}}} \frac{\partial \mathcal{M}'[g, \boldsymbol{r}]}{\partial g(\boldsymbol{r}_1)\partial g(\boldsymbol{r}_2)} = \left\langle \hat{\mathcal{N}} \delta\hat{n}(\boldsymbol{r}_1)\delta\hat{n}(\boldsymbol{r}_2) \right\rangle = \left\langle \hat{\mathcal{N}} (\hat{n}(\boldsymbol{r}_1) - \langle \hat{n}(\boldsymbol{r}_1) \rangle)(\hat{n}(\boldsymbol{r}_2) - \langle \hat{n}(\boldsymbol{r}_2) \rangle) \right\rangle$$

$$= \left\langle \hat{\mathcal{N}} \hat{n}(\boldsymbol{r}_1)\hat{n}(\boldsymbol{r}_2) \right\rangle - \langle \hat{n}(\boldsymbol{r}_1) \rangle \langle \hat{n}(\boldsymbol{r}_2) \rangle$$

$$= \rho^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) - \rho^{(1)}(\boldsymbol{r}_1)\rho^{(1)}(\boldsymbol{r}_2) = \rho'^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) \tag{85}$$

$$\rho'^{(3)}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3) = \rho^{(3)}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3) - \hat{S}\rho^{(1)}(\boldsymbol{r}_1)\rho^{(2)}(\boldsymbol{r}_2, \boldsymbol{r}_3) + 2\rho^{(1)}(\boldsymbol{r}_1)\rho^{(1)}(\boldsymbol{r}_2)\rho^{(1)}(\boldsymbol{r}_3) \ , \tag{86}$$

where $\hat{S}$ is a symmetrizer (Ziesche, 2000). $\rho'^{(2)}$ is also called the exchange-correlation hole (Slater, 1951) and has been studied extensively in Density Functional Theory (Parr & Yang, 1989).

While the reduced densities describe the average occupation numbers of the relevant spin-position space orbitals, the reduced central densities describe the statistical correlation between the fluctuations of the occupation numbers. For instance, the second-order reduced central density $\kappa^{(2)}(I, J)$ defined in the spin-orbital basis ϕ_I describes the covariance between the orbital occupation numbers $n(I)$ and $n(J)$ and can be equated with $\text{cov}(n(I), n(J))$ (Kong & Valeev, 2011). Due to the fact that the above derivations are solely based on the concept of occupation-number operators, independent of the underlying orbital basis, a similar statement is valid for the position basis.

We note that the central moments treat occupied and non-occupied sites on the same footing (Hanauer & Köhn, 2012): *all* fluctuations of the site occupations are taken into account. This implies that the central density $\rho'^{(2)}(I, J)$ does not really describe "the hole associated with a reference electron in I", but rather describes all correlations between the occupation-number fluctuations that occur at the given orbitals. If we assume that the wave function can be written as $|\Psi\rangle = \sum_{\boldsymbol{k}} c_{\boldsymbol{k}} |\boldsymbol{k}\rangle$, we can see this more clearly as follows

$$
\begin{aligned}
\rho'^{(2)}(I, J) &= \left\langle \hat{\mathcal{N}} \delta\hat{n}(I)\delta\hat{n}(J) \right\rangle \\
&= \left\langle \hat{\mathcal{N}}\hat{n}(I)\hat{n}(J) \right\rangle - \langle\hat{n}(I)\rangle \langle\hat{n}(J)\rangle \\
&= \sum_{k} |c_{\boldsymbol{k}}|^2 \left(\left\langle \boldsymbol{k} \left| \hat{\mathcal{N}}\hat{n}(I)\hat{n}(J) \right| \boldsymbol{k} \right\rangle - \langle\hat{n}(I)\rangle \langle\hat{n}(J)\rangle \right).
\end{aligned}
\tag{87}
$$

Hence, even an ONV $|\boldsymbol{l}\rangle$ in which neither I nor J are occupied (i.e. there is no 'reference electron' as such) contributes a term $-|c_l|^2 \rho^{(1)}(I)\rho^{(1)}(J)$ to the central density.

The second-order domain central densities $\rho'^{(2)}(\Omega_a, \Omega_a)$ and $\rho'^{(2)}(\Omega_a, \Omega_b)$ are given by

$$
\rho'^{(2)}(\Omega_a, \Omega_a) = \rho^{(2)}(\Omega_a, \Omega_a) - \rho^{(1)}(\Omega_a)\rho^{(1)}(\Omega_a)
\tag{88}
$$

$$
\rho'^{(2)}(\Omega_a, \Omega_b) = \rho^{(2)}(\Omega_a, \Omega_b) - \rho^{(1)}(\Omega_a)\rho^{(1)}(\Omega_b)
\tag{89}
$$

and are also called the 'localization' and the 'delocalization' indices respectively (R. F. Bader & Stephens, 1974; Angyan, Loos, & Mayer, 1994; Fradera, Austen, & Bader, 1999; Fradera,

Poater, Simon, Duran, & Solà, 2002; Baranov & Kohout, 2011). By interpreting the domain indices as indices to a matrix $\boldsymbol{\rho}_2'(\Omega_a, \Omega_b)$, we obtain the 'Localization-Delocalization Matrix' or LDM of Matta et al. (Matta, Sumar, Cook, & Ayers, 2016). Further derivatives lead to M-th order central densities, which also have been called 'multicenter bonding' indices. The idea of multicenter bonding indices was first proposed by Giambiagi et al. (Giambiagi, de Giambiagi, & Mundim, 1990). They have been used extensively to quantify aromaticity (Ponec, Bultinck, & Saliner, 2005; Bultinck, Ponec, & Van Damme, 2005; Feixas, Matito, Poater, & Solà, 2015).

Multiple descriptors that 'mix' the orbital and domain perspective have already received significant attention from the conceptual quantum chemistry community. As such, the second-order orbital and domain central density $\rho'^{(2)}(\boldsymbol{r}, \Omega_a)$

$$\rho'^{(2)}(\boldsymbol{r}, \Omega_a) = \rho^{(2)}(\boldsymbol{r}, \Omega_a) - \rho^{(1)}(\boldsymbol{r})\rho^{(1)}(\Omega_a) \tag{90}$$

is also called the 'Domain Averaged Fermi Hole' or DAFH (Ponec, 1997, 1998; Poater, Sola, Duran, & Fradera, 2002; Ponec, Cooper, & Savin, 2008; Cooper & Ponec, 2008; Ponec & Feixas, 2009; Bultinck, Van Neck, Acke, & Ayers, 2012). In position space, $\hat{n}(\Omega)$ can be expressed in terms of orbitals that have been unitarily rotated by the eigenvectors of the domain overlap matrix $\boldsymbol{S}^\Omega$. As such, for a single Slater determinant wave function, we can reduce the second-order central density $\rho'^{(2)}(\boldsymbol{r}, \Omega)$ to

$$\begin{aligned}
\rho'^{(2)}(\boldsymbol{r}, \Omega) &= \left\langle \hat{\mathcal{N}}\hat{n}(\boldsymbol{r})\hat{n}(\Omega) \right\rangle - \langle \hat{n}(\boldsymbol{r}) \rangle \langle \hat{n}(\Omega) \rangle \\
&= \sum_{ijk} \lambda_k \tilde{\phi}_i^*(\boldsymbol{r})\tilde{\phi}_j(\boldsymbol{r}) \left[\left\langle \hat{\tilde{\phi}}_i^+ \hat{\tilde{\phi}}_k^+ \hat{\tilde{\phi}}_k^- \hat{\tilde{\phi}}_j^- \right\rangle - \left\langle \hat{\tilde{\phi}}_i^+ \hat{\tilde{\phi}}_j^- \right\rangle \left\langle \hat{\tilde{\phi}}_k^+ \hat{\tilde{\phi}}_k^- \right\rangle \right] \\
&= \sum_{ijk} \lambda_k \tilde{\phi}_i^*(\boldsymbol{r})\tilde{\phi}_j(\boldsymbol{r}) \left[\delta_{ij} - \delta_{ik}\delta_{jk} - \delta_{ij} \right] \\
&= -\sum_k \lambda_k \left| \tilde{\phi}_k(\boldsymbol{r}) \right|^2 ,
\end{aligned} \tag{91}$$

where we have used the fact that $\hat{n}(\boldsymbol{r})$ can be expressed in the basis of the rotated orbitals $\left\{ \tilde{\phi}_j \right\}$ as

$$\hat{n}(\boldsymbol{r}) = \sum_{ij} \tilde{\phi}_i^*(\boldsymbol{r})\tilde{\phi}_j(\boldsymbol{r})\hat{\tilde{\phi}}_i^+ \hat{\tilde{\phi}}_j^- , \tag{92}$$

and the fact that the occupation-numbers of a single Slater determinant are invariant under unitary rotations. As such, for a single Slater determinant wave function, the orbitals that
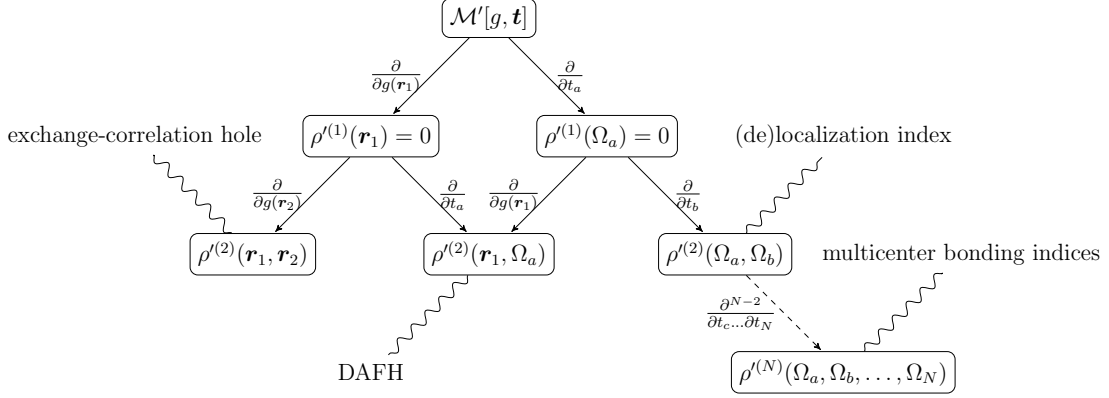
Figure 3: Overview of descriptors derived from the reduced central density generating functional $\mathcal{M}'[g, \boldsymbol{t}]$. Arrows indicate successive differentiations of that generating function.

have been rotated according to the domain orbital overlap matrix $\boldsymbol{S}^\Omega$ are also called the 'DAFH orbitals' or 'Domain Natural Orbitals' and the eigenvalues of $\boldsymbol{S}^\Omega$ the 'DAFH orbital occupation number' (Francisco et al., 2009).

Furthermore, if we define the first-order conditional density $\rho^{(1)}(\boldsymbol{r}_1|\boldsymbol{r}_2)$ as

$$\rho^{(1)}(\boldsymbol{r}_1|\boldsymbol{r}_2) = \frac{\rho^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2)}{\rho^{(1)}(\boldsymbol{r}_2)} \ , \tag{93}$$

we can express $\rho'^{(2)}(\boldsymbol{r}, \Omega)$ for multi-determinant wave function as

$$\begin{aligned}
\rho'^{(2)}(\boldsymbol{r}, \Omega) &= \left\langle \hat{\mathcal{N}} \hat{n}(\boldsymbol{r}) \hat{n}(\Omega) \right\rangle - \rho^{(1)}(\boldsymbol{r}) \left\langle \hat{n}(\Omega) \right\rangle \\
&= \sum_{j=1}^{M} \lambda_j \left[\left\langle \hat{\mathcal{N}} \hat{n}(\boldsymbol{r}) \hat{n}(\tilde{j}) \right\rangle - \rho^{(1)}(\boldsymbol{r}) \left\langle \hat{n}(\tilde{j}) \right\rangle \right] \\
&= \sum_{j=1}^{M} \lambda_j \rho^{(1)}(\tilde{j}) \left[\frac{\rho^{(2)}(\boldsymbol{r}, \tilde{j})}{\rho^{(1)}(\tilde{j})} - \rho^{(1)}(\boldsymbol{r}) \right] \\
&= -\sum_{j=1}^{M} \lambda_j \rho^{(1)}(\tilde{j}) \left[\rho^{(1)}(\boldsymbol{r}) - \rho^{(1)}(\boldsymbol{r}|\tilde{j}) \right] \ , \tag{94}
\end{aligned}$$

This is consistent with the interpretation of (Francisco et al., 2014): $\rho'^{(2)}(\boldsymbol{r}, \Omega)$ provides a decomposition of the excluded density lying in the region over which the central moment is calculated.

We note that most of the hierarchy generated from $\mathcal{M}'[g, \boldsymbol{t}]$ (see Figure 3) is still left largely unexplored. In view of the insight that has already been gleaned from those descriptors that already have been received attention from the community, we believe that many

more important chemical insights can be revealed using currently unexplored descriptors in this hierarchy. However, we note that one has to be careful in assigning additional chemical interpretations, such as indicating 'bonding', to these central densities. For example, according to Mayer (Mayer, 2007), the second-order central density is inadequate for defining bond orders as it leads to results that are not chemical. Bochicchio et al. (Bochicchio, Lain, & Torre, 2003) point out that bond orders that are defined on the fluctuation of electron populations should be restricted to noncorrelated wave functions. In further support of these findings, we have recently shown that increased covalent contribution lead to reductions in second-order domain central density (Acke & Bultinck, 2018; A. M. Pendás & Francisco, 2018a). As such, many of the works using delocalization indices that venture into additional chemical interpretations are worth revisiting.

ORBITAL AND DOMAIN DENSITY CUMULANTS

In the previous section, we showed how the moment generating function can be used to generate (central) moments of the (domain) occupation number operator. Next to moments, we can also generate cumulants from a related generated function. In classical statistics, the cumulants $\langle X^n \rangle_c$ of a stochastic variable X can be obtained through successive differentiation of the cumulant-generating function $\mathcal{C}(t)$, which is defined as the logarithm of the moment generating function $\mathcal{M}(t)$

$$\mathcal{C}(t) = \ln \mathcal{M}(t) = \ln \left\langle e^{tX} \right\rangle . \tag{95}$$

For instance, the second cumulant of X is given by

$$
\begin{aligned}
\langle X^2 \rangle_c &= \lim_{t \to 0} \frac{\mathrm{d}^2 \mathcal{C}(t)}{\mathrm{d}t^2} \\
&= \lim_{t \to 0} \frac{\mathrm{d}^2 \ln \mathcal{M}(t)}{\mathrm{d}t^2} \\
&= \lim_{t \to 0} \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{1}{M(t)} \frac{\mathrm{d}M(t)}{\mathrm{d}t} \right) \\
&= \lim_{t \to 0} \left(-\frac{1}{\mathcal{M}(t)^2} \frac{\mathrm{d}\mathcal{M}(t)}{\mathrm{d}t} \frac{\mathrm{d}\mathcal{M}(t)}{\mathrm{d}t} + \frac{1}{\mathcal{M}(t)} \frac{\mathrm{d}^2 \mathcal{M}(t)}{\mathrm{d}t^2} \right) \\
&= \langle X^2 \rangle - \langle X \rangle^2 ,
\end{aligned}
\tag{96}
$$

and is also called the variance of X or $\mathrm{var}(X)$. We can also generate multi-variable cumulants from $\mathcal{C}(t_X, t_Y)$, which depend on the multi-variable moment generating function $\mathcal{M}(t_X, t_Y)$

$$\mathcal{M}(t_X, t_Y) = \left\langle e^{t_X X + t_Y Y} \right\rangle \ . \tag{97}$$

As such, the second 'mixed' cumulant of X and Y is given by

$$\begin{aligned}
\mathcal{C}(t_X, t_Y) &= \lim_{t_X, t_Y \to 0} \frac{\mathrm{d}^2 \mathcal{C}(t_X, t_Y)}{\mathrm{d}t_X \mathrm{d}t_Y} \\
&= \lim_{t_X, t_Y \to 0} \left(-\frac{1}{\mathcal{M}(t_X, t_Y)^2} \frac{\mathrm{d}\mathcal{M}(t_X, t_Y)}{\mathrm{d}t_X} \frac{\mathrm{d}\mathcal{M}(t_X, t_Y)}{\mathrm{d}t_Y} + \frac{1}{\mathcal{M}(t_X, t_Y)} \frac{\mathrm{d}^2 \mathcal{M}(t_X, t_Y)}{\mathrm{d}t_X \mathrm{d}t_Y} \right) \\
&= \langle XY \rangle - \langle X \rangle \langle Y \rangle \ ,
\end{aligned} \tag{98}$$

and is also called the covariance of X and Y or $\mathrm{cov}(X, Y)$. A related quantity is the Pearson correlation coefficient $\rho_{X,Y}$

$$\rho_{X,Y} = \frac{\mathrm{cov}(X, Y)}{\sqrt{\mathrm{var}(X)} \sqrt{\mathrm{var}(Y)}} \ . \tag{99}$$

Here, $\rho_{X,Y} = 1$ or $\rho_{X,Y} = -1$ indicates perfect positive or negative statistical correlation, while $\rho_{X,Y} = 0$ indicates an absence of statistical correlation between X and Y (which does not necessarily equate to the statistical independence of X and Y).

As was initially shown by Kubo (Kubo, 1962), we can construct a quantum statistical cumulant generating function $\mathcal{C}\left[g, \boldsymbol{t}\right]$ from the moment generating function

$$\mathcal{M}[g, \boldsymbol{t}] := \left\langle \hat{\mathcal{N}} \exp \left(\int g(\boldsymbol{r}) \hat{n}(\boldsymbol{r}) \mathrm{d}\boldsymbol{r} + \sum_{a=1}^{d} t_a \hat{n}(\Omega_a) \right) \right\rangle \ , \tag{100}$$

as follows

$$\mathcal{C}\left[g, \boldsymbol{t}\right] = \ln\left(\mathcal{M}\left[g, \boldsymbol{t}\right]\right) \ . \tag{101}$$

As such, the first order cumulant $k^{(1)}(\boldsymbol{r}_1)$ for the position space orbital basis is equal to the first moment

$$\kappa^{(1)}(\boldsymbol{r}_1) = \rho^{(1)}(\boldsymbol{r}_1) \ , \tag{102}$$

while the second and third order cumulants $\kappa^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2)$ and $\kappa^{(3)}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3)$ are equal to the central moments

$$\kappa^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \rho^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) - \rho^{(1)}(\boldsymbol{r}_1)\rho^{(1)}(\boldsymbol{r}_2) = \rho'^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) \tag{103}$$

$$\kappa^{(3)}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3) = \rho^{(3)}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3) - \hat{S}\rho^{(1)}(\boldsymbol{r}_1)\rho^{(2)}(\boldsymbol{r}_2, \boldsymbol{r}_3) + \frac{1}{2}\rho^{(1)}(\boldsymbol{r}_1)\rho^{(1)}(\boldsymbol{r}_2)\rho^{(1)}(\boldsymbol{r}_3) = \rho'^{(3)}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3) ,$$
$$(104)$$

with $\hat{S}$ a symmetrizer (Ziesche, 2000). Although fourth and higher order cumulants are not equal to (central) moments, they can be constructed from them. Many theoretical problems are simpler to tackle when formulated in terms of cumulants. In particular, the cumulant of two random variables gives an indication of their mutual statistical dependence, with completely independent variables resulting in a zero cumulant. However, as already stated in the section on reduced densities, a probabilistic interpretation of reduced densities can lead to an erroneous interpretation of the derived cumulants. As stated by (McWeeny, 1960):

> "To a first approximation it would be expected that the motion of one electron would be independent of the instantaneous positions of the others, though the form of $\rho^{(1)}$ would certainly reflect the average effect of the $N-1$ electrons. In this case, where the particles might be described as quasi-independent, we should have $\rho^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \rho^{(1)}(\boldsymbol{x}_1)\rho^{(1)}(\boldsymbol{x}_2)$."

However, as indicated by Bader and Kutzelnigg (R. F. Bader & Stephens, 1974; Kutzelnigg, 2003), a pair density can never be equal to a product of electron densities and this quasi-independence should instead be reflected by

$$\rho^{(2)}_{\text{ind}}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \frac{n-1}{n}\rho^{(1)}(\boldsymbol{r}_1)\rho^{(1)}(\boldsymbol{r}_2) , \qquad (105)$$

where the prefactor can at best be justifiably neglected for extended systems. This renormalization factor is related to the fact that the electrons are countable particles and is independent of the particle statistics (Kutzelnigg, 1973). As such, assuming that the cumulant is the difference between a correlated second-order reduced density and (quasi)-independent first-order reduced densities does not lead to consistent interpretations.

We can obtain the $k+l$-order reduced density cumulant by successive differentiations

$$\kappa_{k+l}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_k, \Omega_1, \ldots, \Omega_l) = \lim_{\substack{g \to 0 \\ \boldsymbol{t} \to \boldsymbol{0}}} \frac{\partial^{k+l}\mathcal{C}[g, \boldsymbol{t}]}{\partial g(\boldsymbol{r}_1) \ldots \partial g(\boldsymbol{r}_k) \partial t_1 \ldots \partial t_l} . \qquad (106)$$

electron density $\mathcal{C}[g, \boldsymbol{t}] = \ln\left(\mathcal{M}[g, \boldsymbol{t}]\right)$ domain population

$\frac{\partial}{\partial g(\boldsymbol{r}_1)}$ $\frac{\partial}{\partial t_a}$

exchange-correlation hole $\kappa^{(1)}(\boldsymbol{r}_1) = \rho^{(1)}(\boldsymbol{r}_1)$ $\kappa^{(1)}(\Omega_a) = \rho^{(1)}(\Omega_a)$ (de)localization index

$\frac{\partial}{\partial g(\boldsymbol{r}_2)}$ $\frac{\partial}{\partial t_a}$ $\frac{\partial}{\partial g(\boldsymbol{r}_1)}$ $\frac{\partial}{\partial t_b}$

$\kappa^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2)$ $\kappa^{(2)}(\boldsymbol{r}_1, \Omega_a)$ $\kappa^{(2)}(\Omega_a, \Omega_b)$

$\frac{\partial^{N-2}}{\partial t_l \ldots \partial t_k}$

DAFH $\kappa^{(N)}(\boldsymbol{r}_1, \Omega_a, \ldots, \Omega_k)$
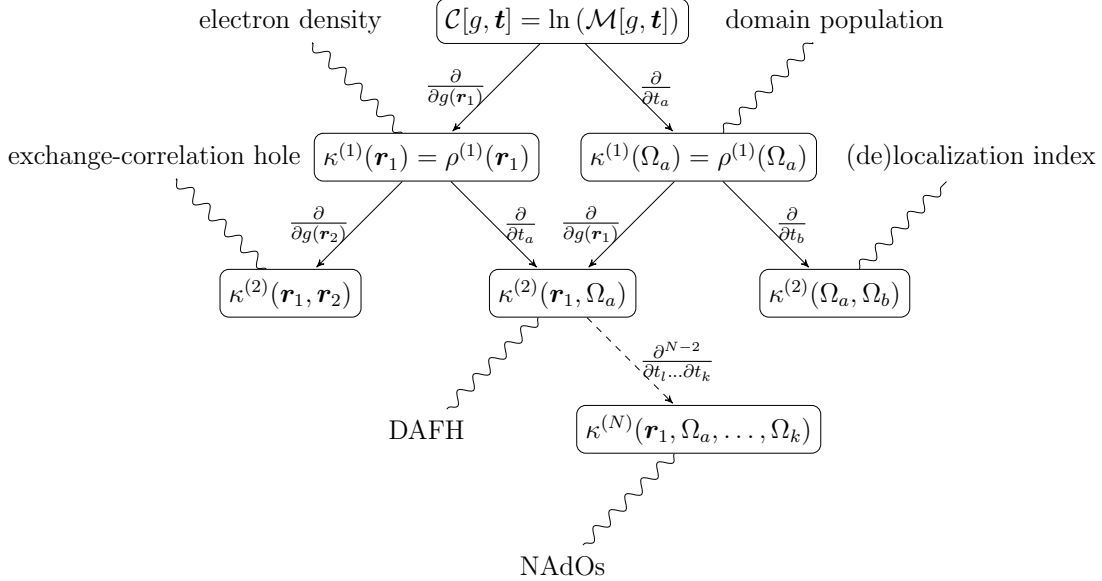
NAdOs

Figure 4: Overview of descriptors derived from the reduced density cumulant generating functional $\mathcal{C}[g, \boldsymbol{t}]$. Arrows indicate successive differentiations of that generating function.

In this way, the cumulant generating function leads to a hierarchy of bonding descriptors visualized in Figure 4.

Of particular interest, are the quantities obtained by differentiating only one time with respect to $g(\boldsymbol{r})$ (and e.g. k times with respect to $\boldsymbol{t}$) as they provide an effective one-electron description of a k-domain integrated cumulant density. As such, they offer a reflection of the degree of statistical correlation between the fluctuations of a single occupation-number and k domain occupation-numbers. In position space, these quantities have been coined the 'Natural Adaptive Orbitals' or NAdOs by Francisco et al. (Francisco, Martín Pendás, García-Revilla, & Álvarez Boto, 2013; Menéndez, Álvarez Boto, Francisco, & Martín Pendás, 2015), since they can be used to extend the applicability of concepts based on orbital models to correlated levels of theory. In this theory, the resulting one-electron functions are considered to provide at each point in spin-position space an indication of the interaction of (QTAIM) domains with that position. As noted by Pendás et al. (Menéndez et al., 2015), these one-electron functions are contained inside those domains, if no $(k+1)$-center interactions exist.

We note that most of the hierarchy detailed in Figure 4 is still left largely unexplored. In

view of the insight that has already been gleaned from those descriptors that already have been received attention from the community, we believe that many more important chemical insights lie can be revealed using these currently underutilized descriptors.

We also note that the above hierarchies are based on occupation number operators that lead to reduced densities. By generalizing this occupation number operator to transition elements such as

$$\hat{n}(\boldsymbol{x};\boldsymbol{x}') = \hat{\psi}^+(\boldsymbol{x})\hat{\psi}^-(\boldsymbol{x}') \ , \tag{107}$$

we can obtain reduced density matrices, such as the first order reduced density matrix $\rho^{(1)}(\boldsymbol{x};\boldsymbol{x}')$

$$\Gamma^{(1)}(\boldsymbol{x};\boldsymbol{x}') = \left\langle \hat{\psi}^+(\boldsymbol{x})\hat{\psi}^-(\boldsymbol{x}') \right\rangle \ . \tag{108}$$

In this case, the concomitant generating functions are built using Grassmann variables (Ziesche, 1992, 2000) and the reduced cumulant density matrices have additional off-diagonal contributions (Kutzelnigg & Mukherjee, 1999). For instance, the second order reduced cumulant density matrix $K^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2; \boldsymbol{x}'_1, \boldsymbol{x}'_2)$ is given by

$$K^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2; \boldsymbol{x}'_1, \boldsymbol{x}'_2) = \Gamma^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2; \boldsymbol{x}'_1, \boldsymbol{x}'_2) + \hat{A}\Gamma^{(1)}(\boldsymbol{x}_1; \boldsymbol{x}'_1)\Gamma^{(1)}(\boldsymbol{x}_2; \boldsymbol{x}'_2) \ , \tag{109}$$

where $\hat{A}$ is an antisymmetrizer (Ziesche, 2000). According to Bochicchio et al. (Bochicchio et al., 2003), $\kappa^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2)$ also describes effectively unpaired electrons (Lain, Torre, Bochicchio, & Ponec, 2001), which hinders its use in describing bonding phenomena based on paired electrons. For bonding analysis purposes, they propose to focus on the difference between the $K^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2) = K^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2; \boldsymbol{x}_1, \boldsymbol{x}_2)$ and $\kappa^{(2)}$, which reduces to an exchange term

$$K^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2) - \kappa^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \Gamma^{(1)}(\boldsymbol{x}_1; \boldsymbol{x}_2)\Gamma^{(1)}(\boldsymbol{x}_2; \boldsymbol{x}_1) \ . \tag{110}$$

Despite their inherent utility for quantum chemical theories (Mundim, Giambiagi, & de Giambiagi, 1994), most of the known quantum chemical descriptors are still based on a density based perspective, leaving much of the hierarchy of (cumulant) density matrices unexplored. In contrast, the resulting (cumulant) density matrix quantities have received much attention in modern method development, such as in the Contracted Schrödinger Equation method (Mazziotti, 2000) and for devising reconstruction functionals for density matrices (Herbert & Harriman, 2002; DePrince III & Mazziotti, 2007).

CONCLUSIONS

We have reviewed the statistical theory of occupancy by constructing (domain) occupation number operators and determining their (central) moments and cumulants through a suitable generating function. Many of the existing quantum chemical descriptors can be matched to a statistical quantity in the resulting hierarchy. This hierarchy connects seemingly unrelated concepts and provides a unified interpretation for all these descriptors. Furthermore, it points to quantum chemical concepts that not yet have been explored.

The resulting framework provides a common language between conceptual quantum chemists and theoretical quantum physicists. By elucidating this language in this review, we hope that this review will stimulate further research into these quantum chemical descriptors through such interdisciplinary collaborations.

APPENDIX: SECOND QUANTIZED MACHINERY

Creation and annihilation operators

The quantum chemistry of electrons is described more succinctly when allowing for varying particle number (Helgaker, Jørgensen, & Olsen, 2000; Surján, 1989). Relations between states of varying particle number can be described in Fock space $\mathcal{F}$, which is the direct sum of the $i = 0$- to $i = \infty$-particle Hilbert spaces

$$\mathcal{F} = \bigoplus_{i=0}^{\infty} \mathcal{H}_i \ . \tag{111}$$

We can define a creation operator in $\mathcal{F}$ by its action on an antisymmetric N-particle state (which is essentially equivalent to a single Slater determinant)

$$\hat{\phi}_I^+ \left| \phi_1 \ldots \phi_N \right\rangle := \left(\prod_{J=1}^{N} (1 - \delta_{IJ}) \right) \left| \phi_I \phi_1 \ldots \phi_N \right\rangle \ . \tag{112}$$

in which an electron is 'created' in orbital ϕ_I. The factors $(1 - \delta_{IJ})$ make sure that a particle cannot be created more than once. The associated annihilation operator can be defined as

$$\hat{\phi}_I^- \left| \phi_I \phi_1 \ldots \phi_N \right\rangle := \left| \phi_1 \ldots \phi_N \right\rangle \ , \tag{113}$$

and 'annihilates' the electron in orbital ϕ_I. Whenever there is no electron in orbital ϕ_I, the operator $\hat{\phi}_I^-$ annihilates the state altogether

$$\hat{\phi}_I^- |\phi_1 \ldots \phi_N\rangle := 0 \ . \tag{114}$$

In this review, we will employ Longuet-Higgins notation (Longuet-Higgins, 1966), which explicitly incorporates the identity of the underlying single-particle basis in the notation of the operators. In particular, we will employ the following notations

$$\hat{\chi}_I^+ |0\rangle = |\chi_I\rangle \quad \text{finite, non-orthogonal spin-orbital basis} \tag{115}$$

$$\hat{\phi}_I^+ |0\rangle = |\phi_I\rangle = |I\rangle \quad \text{finite, orthonormal spin-orbital basis} \tag{116}$$

$$\hat{\psi}^+(\boldsymbol{x}) |0\rangle = |\boldsymbol{x}\rangle \quad \text{spin-position orbital creation operator (field operator) ,} \tag{117}$$

where an orthonormal spin-orbital basis is characterized by

$$\langle \phi_I | \phi_J \rangle = \int \phi_I^*(\boldsymbol{x}) \phi_J(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \delta_{IJ} \ , \tag{118}$$

and a non-orthogonal spin-orbital basis by (Surján, 1989)

$$\langle \chi_I | \chi_J \rangle = \int \chi_I^*(\boldsymbol{x}) \chi_J(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \boldsymbol{\Sigma}_{IJ} \neq \delta_{IJ} \ , \tag{119}$$

with $\boldsymbol{\Sigma}$ the spin-orbital overlap matrix. For an orthonormal single-particle basis, the associated annihilation operator is also the Hermitian conjugate of the creation operator

$$\hat{\phi}_I^- = (\hat{\phi}_I^+)^\dagger \ , \tag{120}$$

and can hence be used to create particles in the dual vector. The creation and annihilation operators codify the Pauli antisymmetry principle into their anticommutation relations (where $[\hat{A}, \hat{B}]_+ = \hat{A}\hat{B} + \hat{B}\hat{A}$)

$$[\hat{\phi}_I^+, \hat{\phi}_J^+]_+ = 0 \tag{121}$$

$$[\hat{\phi}_I^-, \hat{\phi}_J^-]_+ = 0 \tag{122}$$

$$[\hat{\phi}_I^+, \hat{\phi}_J^-]_+ = \delta_{IJ} = [\hat{\phi}_J^-, \hat{\phi}_I^+]_+ \ . \tag{123}$$

For non-orthogonal spin-orbitals, these anticommutation relations are replaced by

$$[\hat{\chi}_I^+, \hat{\chi}_J^+]_+ = 0 \tag{124}$$

$$[\hat{\chi}_I^-, \hat{\chi}_J^-]_+ = 0 \tag{125}$$

$$[\hat{\chi}_I^+, \hat{\chi}_J^-]_+ = \Sigma_{IJ} \ . \tag{126}$$

We can restore the usual anticommutation relations by constructing a bio-orthogonal basis $\{\varphi_I\}$ (Surján, 1989) as

$$\hat{\varphi}_I^- = \sum_J \Sigma_{IJ}^{-1} \hat{\chi}_J^- \tag{127}$$

for which

$$\left[\hat{\chi}_I^+, \hat{\varphi}_I^-\right]_+ = \delta_{IJ} \ . \tag{128}$$

Although this condition allows us to use any result derived for the orthogonal case (Surján, 1989), it does not solve the problem that the annihilation operator $\hat{\chi}_I^-$ is no longer the Hermitian conjugate of the creation operator $\hat{\chi}_I^+$. As such, the construction of bra wave functions is only possible by means of the inverse transformation

$$\hat{\chi}_I^- = \sum_J \Sigma_{IJ} \hat{\varphi}_J^- \ , \tag{129}$$

which reintroduces the need to take the overlap matrix Σ explicitly into account when using non-orthogonal orbitals.

In non-relativistic theory, spin-orbitals $\{\phi_I\}$ are commonly of the following form (Helgaker et al., 2000)

$$\phi_{i\sigma}(\boldsymbol{r}, m_s) = \phi_i(\boldsymbol{r})\sigma(m_s) \ . \tag{130}$$

Hence, we may separate out spin from spatial parts of the creation and annihilation operators, giving

$$[\hat{\phi}_{i\sigma}^+, \hat{\phi}_{j\sigma'}^+]_+ = 0 \tag{131}$$

$$[\hat{\phi}_{i\sigma}^-, \hat{\phi}_{j\sigma'}^-]_+ = 0 \tag{132}$$

$$[\hat{\phi}_{i\sigma}^+, \hat{\phi}_{j\sigma'}^-]_+ = \delta_{ij}\delta_{\sigma\sigma'} \ , \tag{133}$$

where the M spin-orbitals $\{\phi_I\}$ can be expressed in terms of m spatial orbitals $\{\phi_i\}$.

Field operators

Given a certain basis, we can express this basis in an alternative basis with the aid of the single-particle completeness relations. If we start from the complete set of eigenvectors of the operator $\hat{\boldsymbol{x}}$

$$\hat{\boldsymbol{x}} \left|\boldsymbol{x}\right\rangle = \boldsymbol{x} \left|\boldsymbol{x}\right\rangle \ , \tag{134}$$

the corresponding basis transformations are given by

$$|\phi_I\rangle = \int |\boldsymbol{x}\rangle \langle \boldsymbol{x}|\phi_I\rangle \, \mathrm{d}\boldsymbol{x} = \int \phi_I(\boldsymbol{x}) |\boldsymbol{x}\rangle \, \mathrm{d}\boldsymbol{x} \tag{135}$$

$$|\boldsymbol{x}\rangle = \sum_I |\phi_I\rangle \langle \phi_I|\boldsymbol{x}\rangle = \sum_I \phi_I^*(\boldsymbol{x}) |\phi_I\rangle \ . \tag{136}$$

From this, we can derive the transformation for the associated creation operator

$$\hat{\psi}^+(\boldsymbol{x}) |0\rangle = |\boldsymbol{x}\rangle = \sum_I |\phi_I\rangle \langle \phi_I|\boldsymbol{x}\rangle = \sum_I \phi_I^*(\boldsymbol{x}) |\phi_I\rangle = \sum_I \phi_I^*(\boldsymbol{x})\hat{\phi}_I^+ |0\rangle \tag{137}$$

In a similar way, the inverse transformation can be derived from equation (135)

$$\hat{\phi}_I^+ = \int \phi_I(\boldsymbol{x})\hat{\psi}^+(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \ , \tag{138}$$

and the following transformation rules for annihilation operators can be derived

$$\hat{\psi}^-(\boldsymbol{x}) = \sum_I \phi_I(\boldsymbol{x})\hat{\phi}_I^- \tag{139}$$

$$\hat{\phi}_I^- = \int \phi_I^*(\boldsymbol{x})\hat{\psi}^-(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \ . \tag{140}$$

The operator $\hat{\psi}^+(\boldsymbol{x})$ is called a 'field operator' and acts as a hybrid between first and second quantization (Fetter & Walecka, 2003; Engel & Dreizler, 2013; Lancaster & Blundell, 2014). As such, it is very convenient for porting first quantized expression into second quantization.

Normal ordering

Given a sequence of second quantized operators $\hat{O}$, the normal ordered form $\hat{\mathcal{N}}\hat{O}$ with respect to the vacuum state $|0\rangle$ corresponds to moving all creation operators to the left and all annihilation operators to the right as if they all anticommuted (Saue, 2015). Hence,

$$\hat{\mathcal{N}}\hat{\phi}_I^- \hat{\phi}_J^- = \hat{\phi}_I^- \hat{\phi}_J^- \tag{141}$$

$$\hat{\mathcal{N}}\hat{\phi}_I^+ \hat{\phi}_J^+ = \hat{\phi}_I^+ \hat{\phi}_J^+ \tag{142}$$

$$\hat{\mathcal{N}}\hat{\phi}_I^+ \hat{\phi}_J^- = \hat{\phi}_I^+ \hat{\phi}_J^- \tag{143}$$

$$\hat{\mathcal{N}}\hat{\phi}_I^- \hat{\phi}_J^+ = -\hat{\phi}_I^+ \hat{\phi}_J^- \tag{144}$$

Note that the action of the normal ordered form of a collection of operators $\hat{O}$ that contains duplicate creation or annihilation operators on an arbitrary state $|\boldsymbol{K}\rangle$ is such that the state is annihilated

$$\hat{\mathcal{N}}\hat{\phi}_I^+ \hat{\phi}_I^- \hat{\phi}_I^+ \hat{\phi}_I^- |\boldsymbol{K}\rangle = -\hat{\phi}_I^+ \hat{\phi}_I^+ \hat{\phi}_I^- \hat{\phi}_I^- |\boldsymbol{K}\rangle = 0 \ . \tag{145}$$

A term related to the idea of normal ordering, is a 'contraction'. This contraction is defined as the difference between the operator string itself and its normal ordered form

$$\overbrace{\hat{\phi}_I^+ \hat{\phi}_J^-} = \hat{\phi}_I^+ \hat{\phi}_J^- - \hat{\mathcal{N}}\hat{\phi}_I^+ \hat{\phi}_J^- \ . \tag{146}$$

This difference can only give rise to zero or a complex number times the identity (Fetter & Walecka, 2003). As such, these contractions play a central role in Wick's theorem, which states that any operator string can be written as a linear combination of normal-ordered strings, were only fully contracted terms contribute to expectation values with respect to vacuum $|0\rangle$.

References

Acke, G., & Bultinck, P. (2018). The influence of correlation on (de) localization indices from a valence bond perspective. *Journal of molecular modeling*, *24*(10), 275.

Acke, G., De Baerdemacker, S., Claeys, P. W., Van Raemdonck, M., Poelmans, W., Van Neck, D., & Bultinck, P. (2016). Maximum probability domains for hubbard models. *Molecular Physics*, *114*, 1392-1405.

Agostini, F., Ciccotti, G., Savin, A., & Vuilleumier, R. (2015). Maximum probability domains for the analysis of the microscopic structure of liquids. *The Journal of chemical physics*, *142*(6), 064117.

Andrews, G. E. (1998). *The theory of partitions* (No. 2). Cambridge university press.

Angyan, J. G., Loos, M., & Mayer, I. (1994). Covalent bond orders and atomic valence indices in the topological theory of atoms in molecules. *The Journal of Physical Chemistry*, *98*(20), 5244–5248.

Aslangul, C., Constanciel, R., Daudel, R., Esnault, L., & Ludena, E. (1974). The loge theory as a starting point for variational calculations. i. general formalism. *International Journal of Quantum Chemistry*, *8*(4), 499–522.

Aslangul, C., Constanciel, R., Daudel, R., & Kottis, P. (1972). Aspects of the localizability of electrons in atoms and molecules: Loge theory and related methods. *Advances in quantum chemistry*, *6*, 93–141.

Bader, R., & Stephens, M. (1975). Spatial localization of the electronic pair and number distributions in molecules. *Journal of the American Chemical Society*, *97*(26), 7391–7399.

Bader, R. F. (1990). *Atoms in molecules: a quantum theory*. Oxford University Press, Oxford.

Bader, R. F., & Stephens, M. E. (1974). Fluctuation and correlation of electrons in molecular systems. *Chemical Physics Letters*, *26*(3), 445–449.

Baranov, A. I., & Kohout, M. (2011). Electron localization and delocalization indices for solids. *Journal of computational chemistry*, *32*(10), 2064–2076.

Bochicchio, R., Lain, L., & Torre, A. (2003). On the definition of bond orders at correlated level. *Chemical physics letters*, *374*(5-6), 567–571.

Bochicchio, R., Ponec, R., Lain, L., & Torre, A. (2000). Pair population analysis within aim theory. *The Journal of Physical Chemistry A*, *104*(40), 9130–9135.

Bochicchio, R., Ponec, R., Torre, A., & Lain, L. (2001). Multicenter bonding within the aim theory. *Theoretical Chemistry Accounts*, *105*(4-5), 292–298.

Breuer, H.-P., Petruccione, F., et al. (2002). *The theory of open quantum systems*. Oxford University Press on Demand.

Bultinck, P., Ponec, R., & Van Damme, S. (2005). Multicenter bond indices as a new measure of aromaticity in polycyclic aromatic hydrocarbons. *Journal of physical organic chemistry*, *18*(8), 706–718.

Bultinck, P., Van Alsenoy, C., Ayers, P. W., & Carbó-Dorca, R. (2007). Critical analysis and extension of the hirshfeld atoms in molecules. *The Journal of chemical physics*, *126*(14), 144111.

Bultinck, P., Van Neck, D., Acke, G., & Ayers, P. W. (2012). Influence of electron correlation

and degeneracy on the fukui matrix and extension of frontier molecular orbital theory to correlated quantum chemical methods. *Physical Chemistry Chemical Physics*, *14*(7), 2408–2416.

Cancès, E., Keriven, R., Lodier, F., & Savin, A. (2004). How electrons guard the space: shape optimization with probability distribution criteria. *Theoretical Chemistry Accounts*, *111*(2-6), 373–380.

Causa, M., DAmore, M., Garzillo, C., Gentile, F., & Savin, A. (2013). The bond analysis techniques (elf and maximum probability domains) application to a family of models relevant to bio-inorganic chemistry. In *Applications of density functional theory to biological and bioinorganic chemistry* (pp. 119–141). Springer.

Causà, M., D'Amore, M., Gentile, F., Menéndez, M., & Calatayud, M. (2015). Electron Localization Function and Maximum Probability Domains analysis of semi-ionic oxides crystals, surfaces and surface defects. *Computational and Theoretical Chemistry*, *1053*, 315–321.

Causa, M., & Savin, A. (2011). Maximum probability domains in crystals: the rock-salt structure. *The Journal of Physical Chemistry A*, *115*(45), 13139–13148.

Causà, M., & Savin, A. (2011). Maximum probability domains in the solid-state structures of the elements: the diamond structure. *Zeitschrift für anorganische und allgemeine Chemie*, *637*(7-8), 882–884.

Claverie, P., & Diner, S. (1976). Statistical and stochastic aspects of the delocalization problem in quantum mechanics. In *Localization and delocalization in quantum chemistry* (pp. 395–448). Springer.

Cooper, D. L., & Ponec, R. (2008). A one-electron approximation to domain-averaged fermi hole analysis. *Physical Chemistry Chemical Physics*, *10*(9), 1319–1329.

Daley, D. J., & Vere-Jones, D. (2002). *An introduction to the theory of point processes: volume i: Elementary theory and methods (second edition)* (J. Gani, C. C. Heyde, P. Jagers, & T. G. Kurtz, Eds.). Springer.

Daley, D. J., & Vere-Jones, D. (2008). *An introduction to the theory of point processes: volume ii: general theory and structure (second edition)* (J. Gani, C. C. Heyde, P. Jagers, & T. G. Kurtz, Eds.). Springer.

Daudel, R., Bader, R., Stephens, M., & Borrett, D. (1974). The electron pair in chemistry. *Canadian Journal of Chemistry*, *52*(8), 1310–1320.

Daudel, R., Brion, H., & Odiot, S. (1955). Localizability of electrons in atoms and molecules - application to the study of the notion of shell and of the nature of chemical bonds. *The Journal of Chemical Physics*, *23*(11), 2080–2083.

DePrince III, A. E., & Mazziotti, D. A. (2007). Cumulant reconstruction of the three-electron reduced density matrix in the anti-hermitian contracted schrödinger equation. *The Journal of chemical physics*, *127*(10), 104104.

De Silva, P., & Corminboeuf, C. (2014). Simultaneous visualization of covalent and non-covalent interactions using regions of density overlap. *Journal of chemical theory and computation*, *10*(9), 3745–3756.

Diner, S., & Claverie, P. (1975). Static and dynamical aspects of delocalization. electrons and orbitals. In O. Chalvet, R. Daudel, S. Diner, & J. P. Malrieu (Eds.), *Localization and delocalization in quantum chemistry* (Vol. 2, p. 461-464). D. Reidel Publishing Company.

Engel, E., & Dreizler, R. M. (2011). *Density functional theory: an advanced course.* Springer.

Engel, E., & Dreizler, R. M. (2013). *Density functional theory.* Springer.

Feixas, F., Matito, E., Poater, J., & Solà, M. (2015). Quantifying aromaticity with electron delocalisation measures. *Chemical Society Reviews*, *44*(18), 6434–6451.

Fetter, A. L., & Walecka, J. D. (2003). *Quantum theory of many-particle systems.* Courier Corporation.

Fradera, X., Austen, M. A., & Bader, R. F. (1999). The lewis model and beyond. *The Journal of Physical Chemistry A*, *103*(2), 304–314.

Fradera, X., Poater, J., Simon, S., Duran, M., & Solà, M. (2002). Electron-pairing analysis from localization and delocalization indices in the framework of the atoms-in-molecules theory. *Theoretical Chemistry Accounts*, *108*(4), 214–224.

Francisco, E., & Martín Pendás, A. (2014). Electron number distribution functions from molecular wavefunctions. version 2. *Computer Physics Communications*.

Francisco, E., Martín Pendás, A., García-Revilla, M., & Álvarez Boto, R. (2013). A hierarchy of chemical bonding indices in real space from reduced density matrices and cumulants.

Daudel, R., Bader, R., Stephens, M., & Borrett, D. (1974). The electron pair in chemistry. *Canadian Journal of Chemistry*, *52*(8), 1310–1320.

Daudel, R., Brion, H., & Odiot, S. (1955). Localizability of electrons in atoms and molecules - application to the study of the notion of shell and of the nature of chemical bonds. *The Journal of Chemical Physics*, *23*(11), 2080–2083.

DePrince III, A. E., & Mazziotti, D. A. (2007). Cumulant reconstruction of the three-electron reduced density matrix in the anti-hermitian contracted schrödinger equation. *The Journal of chemical physics*, *127*(10), 104104.

De Silva, P., & Corminboeuf, C. (2014). Simultaneous visualization of covalent and non-covalent interactions using regions of density overlap. *Journal of chemical theory and computation*, *10*(9), 3745–3756.

Diner, S., & Claverie, P. (1975). Static and dynamical aspects of delocalization. electrons and orbitals. In O. Chalvet, R. Daudel, S. Diner, & J. P. Malrieu (Eds.), *Localization and delocalization in quantum chemistry* (Vol. 2, p. 461-464). D. Reidel Publishing Company.

Engel, E., & Dreizler, R. M. (2011). *Density functional theory: an advanced course.* Springer.

Engel, E., & Dreizler, R. M. (2013). *Density functional theory.* Springer.

Feixas, F., Matito, E., Poater, J., & Solà, M. (2015). Quantifying aromaticity with electron delocalisation measures. *Chemical Society Reviews*, *44*(18), 6434–6451.

Fetter, A. L., & Walecka, J. D. (2003). *Quantum theory of many-particle systems.* Courier Corporation.

Fradera, X., Austen, M. A., & Bader, R. F. (1999). The lewis model and beyond. *The Journal of Physical Chemistry A*, *103*(2), 304–314.

Fradera, X., Poater, J., Simon, S., Duran, M., & Solà, M. (2002). Electron-pairing analysis from localization and delocalization indices in the framework of the atoms-in-molecules theory. *Theoretical Chemistry Accounts*, *108*(4), 214–224.

Francisco, E., & Martín Pendás, A. (2014). Electron number distribution functions from molecular wavefunctions. version 2. *Computer Physics Communications*.

Francisco, E., Martín Pendás, A., García-Revilla, M., & Álvarez Boto, R. (2013). A hierarchy of chemical bonding indices in real space from reduced density matrices and cumulants.

Computational and Theoretical Chemistry, *1003*, 71–78.

Francisco, E., Pendás, A. M., & Blanco, M. (2007). Electron number probability distributions for correlated wave functions. *The Journal of chemical physics*, *126*(9), 094102.

Francisco, E., Pendás, A. M., & Blanco, M. (2008). Edf: Computing electron number probability distribution functions in real space from molecular wave functions. *Computer Physics Communications*, *178*(8), 621–634.

Francisco, E., Pendás, A. M., & Blanco, M. (2009). A connection between domain-averaged fermi hole orbitals and electron number distribution functions in real space. *The Journal of chemical physics*, *131*(12), 124125.

Francisco, E., Pendás, A. M., & Blanco, M. A. (2011). Generalized electron number distribution functions: real space versus orbital space descriptions. *Theoretical Chemistry Accounts*, *128*(4-6), 433–444.

Francisco, E., Pendás, A. M., & Costales, A. (2014). On the interpretation of domain averaged fermi hole analyses of correlated wavefunctions. *Physical Chemistry Chemical Physics*, *16*(10), 4586–4597.

Fulde, P. (2012). *Correlated electrons in quantum matter*. World Scientific.

Gallegos, A., Carbó-Dorca, R., Lodier, F., Cancès, E., & Savin, A. (2005). Maximal probability domains in linear molecules. *Journal of computational chemistry*, *26*(5), 455–460.

Geerlings, P., De Proft, F., & Langenaeker, W. (2003). Conceptual density functional theory. *Chemical reviews*, *103*(5), 1793–1874.

Giambiagi, M., de Giambiagi, M. S., & Mundim, K. C. (1990). Definition of a multicenter bond index. *Structural Chemistry*, *1*(5), 423–427.

Gottlieb, A. D. (2005). Introduction to determinantal point processes from a quantum probability viewpoint. *arXiv preprint math/0511334*.

Hanauer, M., & Köhn, A. (2012). Meaning and magnitude of the reduced density matrix cumulants. *Chemical Physics*, *401*, 50–61.

Helgaker, T., Jørgensen, P., & Olsen, J. (2000). *Molecular electronic-structure theory*. John Wiley & Sons.

Herbert, J. M., & Harriman, J. E. (2002). Contraction relations for grassmann products of reduced density matrices and implications for density matrix reconstruction. *Physical*

Review A, *65*(2), 022511.

Hirshfeld, F. L. (1977). Bonded-atom fragments for describing molecular charge densities. *Theoretica chimica acta*, *44*(2), 129–138.

Johnson, E. R., Keinan, S., Mori-Sánchez, P., Contreras-García, J., Cohen, A. J., & Yang, W. (2010). Revealing noncovalent interactions. *Journal of the American Chemical Society*, *132*(18), 6498–6506.

Knizia, G., & Chan, G. K.-L. (2013). Density matrix embedding: A strong-coupling quantum embedding theory. *Journal of chemical theory and computation*, *9*(3), 1428–1432.

Kohout, M. (2004). A measure of electron localizability. *International journal of quantum chemistry*, *97*(1), 651–658.

Kong, L., & Valeev, E. F. (2011). A novel interpretation of reduced density matrix and cumulant for electronic structure theories. *The Journal of chemical physics*, *134*(21), 214109.

Kubo, R. (1962). Generalized cumulant expansion method. *Journal of the Physical Society of Japan*, *17*(7), 1100–1120.

Kutzelnigg, W. (1973). Electron correlation and electron pair theories. In *New concepts i* (pp. 31–73). Springer.

Kutzelnigg, W. (2003). Theory of electron correlation. In J. Rychlewski (Ed.), *Explicitly correlated wave functions in chemistry and physics: theory and applications*. Springer.

Kutzelnigg, W., & Mukherjee, D. (1999). Cumulant expansion of the reduced density matrices. *Journal of Chemical Physics*, *110*(6), 2800–2809.

Lain, L., Torre, A., Bochicchio, R., & Ponec, R. (2001). On the density matrix of effectively unpaired electrons. *Chemical physics letters*, *346*(3), 283–287.

Lancaster, T., & Blundell, S. J. (2014). *Quantum field theory for the gifted amateur*. OUP Oxford.

Longuet-Higgins, H. (1966). p. 105 in "quantum theory of atoms. *Molecules and the Solid State"(ed. P.-O. Löwdin) Academic Press, New York*.

Lopes Jr, O. M., Braïda, B., Causa, M., & Savin, A. (2012). Understanding maximum probability domains with simple models. In *Advances in the theory of quantum systems in chemistry and physics* (pp. 173–184). Springer.

Löwdin, P.-O. (1955a). Quantum theory of many-particle systems. iii. extension of the hartree-fock scheme to include degenerate systems and correlation effects. *Physical Review*, *97*(6), 1509–1520.

Löwdin, P.-O. (1955b). Quantum theory of many-particle systems. ii. study of the ordinary hartree-fock approximation. *Physical Review*, *97*(6), 1490.

Löwdin, P.-O. (1955c). Quantum theory of many-particle systems. i. physical interpretations by means of density matrices, natural spin-orbitals, and convergence problems in the method of configurational interaction. *Physical Review*, *97*(6), 1474.

Macchi, O. (1975). The coincidence approach to stochastic point processes. *Advances in Applied Probability*, 83–122.

Matta, C. F., & Boyd, R. J. (2007). *The quantum theory of atoms in molecules: from solid state to dna and drug design*. John Wiley & Sons.

Matta, C. F., Sumar, I., Cook, R., & Ayers, P. W. (2016). Localization-delocalization matrices and electron density-weighted adjacency/connectivity matrices: A bridge between the quantum theory of atoms in molecules and chemical graph theory. In *Applications of topological methods in molecular chemistry* (pp. 53–88). Springer.

Mayer, I. (1983). Charge, bond order and valence in the ab initio scf theory. *Chemical Physics Letters*, *97*(3), 270–274.

Mayer, I. (1985). Charge, bond order and valence in the ab initio scf theory (addendum). *Chem. Phys. Lett*, *117*(4), 396.

Mayer, I. (2007). Bond order and valence indices: a personal account. *Journal of computational chemistry*, *28*(1), 204–221.

Mayer, I. (2016). *Bond orders and energy components: extracting chemical information from molecular wave functions*. CRC Press.

Mayer, I., & Hamza, A. (2005). Atomic decomposition of identity: General formalism for population analysis and energy decomposition. *International journal of quantum chemistry*, *103*(6), 798–807.

Mazziotti, D. A. (2000). Cumulants and the contracted schrödinger equation. In *Many-electron densities and reduced density matrices* (pp. 139–163). Springer.

McWeeny, R. (1960). Some recent advances in density matrix theory. *Reviews of Modern*

Physics, *32*(2), 335.

Menéndez, M., Álvarez Boto, R., Francisco, E., & Martín Pendás, Á. (2015). One-electron images in real space: Natural adaptive orbitals. *Journal of computational chemistry*, *36*(11), 833–843.

Menéndez, M., & Pendás, A. M. (2014). On the stability of some analytically solvable maximum probability domains. *Theoretical Chemistry Accounts*, *133*(9), 1–8.

Menéndez, M., Pendás, A. M., Braïda, B., & Savin, A. (2014). A view of covalent and ionic bonding from maximum probability domains. *Computational and Theoretical Chemistry*.

Mundim, K. C., Giambiagi, M., & de Giambiagi, M. S. (1994). Multicenter bond index: Grassmann algebra and n-order density functional. *The Journal of Physical Chemistry*, *98*(24), 6118–6119.

Parr, R. G., & Yang, W. (1989). *Density-functional theory of atoms and molecules*. Oxford University Press.

Pendás, A. M., & Francisco, E. (2018a). Decoding real space bonding descriptors in valence bond language. *Physical Chemistry Chemical Physics*, *20*(18), 12368–12372.

Pendás, A. M., & Francisco, E. (2018b). Quantum chemical topology as a theory of open quantum systems. *Journal of chemical theory and computation*, *15*(2), 1079–1088.

Pendás, A. M., Francisco, E., & Blanco, M. (2007a). An electron number distribution view of chemical bonds in real space. *Physical Chemistry Chemical Physics*, *9*(9), 1087–1092.

Pendás, A. M., Francisco, E., & Blanco, M. (2007b). Pauling resonant structures in real space through electron number probability distributions. *The Journal of Physical Chemistry A*, *111*(6), 1084–1090.

Pendás, A. M., Francisco, E., & Blanco, M. (2007c). Spin resolved electron number distribution functions: How spins couple in real space. *The Journal of chemical physics*, *127*(14), 144103.

Pendás, Á. M., Francisco, E., & Blanco, M. A. (2007). Charge transfer, chemical potentials, and the nature of functional groups: answers from quantum chemical topology. *Faraday discussions*, *135*, 423–438.

Poater, J., Duran, M., Solá, M., & Silvi, B. (2005). Theoretical evaluation of electron

delocalization in aromatic molecules by means of atoms in molecules (aim) and electron localization function (elf) topological approaches. *Chemical reviews*, *105*(10), 3911–3947.

Poater, J., Sola, M., Duran, M., & Fradera, X. (2002). The calculation of electron localization and delocalization indices at the hartree–fock, density functional and post-hartree–fock levels of theory. *Theoretical Chemistry Accounts*, *107*(6), 362–371.

Ponec, R. (1997). Electron pairing and chemical bonds. chemical structure, valences and structural similarities from the analysis of the fermi holes. *Journal of Mathematical Chemistry*, *21*(3), 323–333.

Ponec, R. (1998). Electron pairing and chemical bonds. molecular structure from the analysis of pair densities and related quantities. *Journal of mathematical chemistry*, *23*(1-2), 85–103.

Ponec, R., Bultinck, P., & Saliner, A. G. (2005). Multicenter bond indices as a new means for the quantitative characterization of homoaromaticity. *The Journal of Physical Chemistry A*, *109*(29), 6606–6609.

Ponec, R., Cooper, D. L., & Savin, A. (2008). Analytic models of domain-averaged fermi holes: A new tool for the study of the nature of chemical bonds. *Chemistry–A European Journal*, *14*(11), 3338–3345.

Ponec, R., & Feixas, F. (2009). Domain averaged fermi hole analysis for open-shell systems. *The Journal of Physical Chemistry A*, *113*(19), 5773–5779.

Popelier, P. L. (2000). *Atoms in molecules: An introcuction*. Prentice Hall.

Saue, T. (2015). Second quantization. In *European summerschool in quantum chemistry*.

Savin, A. (2001). *Probability distributions and valence shells in atoms*. Reviews in modern quantum chemistry by Sen.

Savin, A. (2004). Phase transition in iodine: a chemical picture. *Journal of Physics and Chemistry of Solids*, *65*(12), 2025–2029.

Savin, A. (2005). On the significance of elf basins. *JOURNAL OF CHEMICAL SCIENCES-BANGALORE-*, *117*(5), 473.

Savin, A., Nesper, R., Wengert, S., & Fässler, T. F. (1997). Elf: The electron localization function. *Angewandte Chemie International Edition in English*, *36*(17), 1808–1832.

Scemama, A., Caffarel, M., & Savin, A. (2007). Maximum Probability Domains from Quantum Monte Carlo calculations. *Journal of computational chemistry*, *28*(1), 442–454.

Silvi, B. (2004). How topological partitions of the electron distributions reveal delocalization. *Physical Chemistry Chemical Physics*, *6*(2), 256–260.

Silvi, B., & Savin, A. (1994). Classification of chemical bonds based on topological analysis of electron localization functions. *Nature*, *371*(6499), 683–686.

Slater, J. C. (1951). A simplification of the hartree-fock method. *Physical Review*, *81*(3), 385.

Soshnikov, A. (2000). Determinantal random point fields. *Russian Mathematical Surveys*, *55*(5), 923–975.

Surján, P. R. (1989). *Second quantized approach to quantum chemistry: an elementary introduction.* Springer-Verlag.

Truskett, T. M., Torquato, S., & Debenedetti, P. G. (1998). Density fluctuations in many-body systems. *Physical Review E*, *58*(6), 7369.

Turek, J., Braïda, B., & De Proft, F. (2017). Bonding in heavier group 14 zero-valent complexes—a combined maximum probability domain and valence bond theory approach. *Chemistry–A European Journal*, *23*(58), 14604–14613.

Vezzetti, D. J. (1975). A new derivation of some fluctuation theorems in statistical mechanics. *Journal of Mathematical Physics*, *16*(1), 31-33.

Yamasaki, T., & Goddard, W. A. (1998). Correlation analysis of chemical bonds. *The Journal of Physical Chemistry A*, *102*(17), 2919–2933.

Yamasaki, T., Mainz, D. T., & Goddard, W. A. (2000). Correlation analysis of chemical bonds (cacb) ii: quantum mechanical operators for classical chemical concepts. *The Journal of Physical Chemistry A*, *104*(11), 2221–2229.

Ziesche, P. (1992). Definition of exchange based on cumulant expansion: Correlation induced narrowing of the exchange hole. *Solid state communications*, *82*(8), 597–602.

Ziesche, P. (2000). Cumulant expansions of reduced densities, reduced density matrices and green's functions. In C. Jerzy (Ed.), *Many-electron densities and reduced density matrices.* Springer.

ACKNOWLEDGEMENTS

FURTHER READING

The works that have been most influential in writing this review are:

- Daudel, R.; Chalvet, O.; Diner, S.; Malrieu, J. P. (1975) Localization and delocalization in quantum chemistry, Volumes 1 and 2, D. Reidel Publishing Company.

- Ziesche, P. (2000). Cumulant expansions of reduced densities, reduced density matrices and green's functions. In C. Jerzy (Ed.), Many-electron densities and reduced density matrices. Springer.

- Kutzelnigg, W. (2003). Theory of electron correlation. In. J. Rychlewski (Ed.), Explicitly correlated wave functions in chemistry and physics: theory and applications. Springer.

- Herbert, J. M. (2003). Reconstructive approaches to one- and two-electron density matrix theory. PhD Thesis. University of Wisconsin-Madison.

- Mayer, I. (2016). Bond orders and energy components: Extracting chemical information from molecular wave functions. CRC Press.