Journal of Molecular Modeling (2024) 30:289
https://doi.org/10.1007/s00894-024-06070-4

RESEARCH

KLD: a program to elucidate the localization of the Fermi and Coulomb holes in molecular systems

Valeria Bedoya[1] · Vladimir Rodríguez[1,2] · Luis Rincón[1] · Cesar Zambrano[1] · Luis Seijas[3] · F. Javier Torres[1]

Received: 16 May 2024 / Accepted: 9 July 2024 / Published online: 29 July 2024

Abstract

Context The electron localization is a concept that allows scientists to better understand the physical and chemical properties of electronic systems. It is associated with the propensity of electron pairs with opposite spins to accumulate as well as with their response to external perturbations. This paper contains a detailed description of the design and implementation of the program KLD, which was primarily developed in our research group to elucidate electron localization in molecular systems by evaluating the information content of electron-pair density functions. KLD employs two information-based functions as a real space measure of the Fermi and Coulomb holes for same-spin electrons and shows a better resolution as compared to other methods (i.e., ELF). Information about the acceleration of the code is also included in the present work, being noticeable the reduction of wall-time calculation and the error calculation between versions.

Methods KLD was designed to be easy to use, extend, and maintain; thus, many principles of modern software development, extensive testing, and package management were adopted. The latest version of the KLD program was created utilizing the Compute Unified Device Architecture (CUDA) version, which allows it to use the computational capacity of NVIDIA Graphics Processing Units (GPUs) for processing purposes. The electron-pair conditional density was calculated from the canonical molecular orbitals obtained at the HF/6-31G(2df,p) level, or alternatively the natural orbitals in the case of explicit correlated wavefunctions computed at the MP2/6-31G(2df,p)//HF/6-31G(2df,p) level.

Keywords Conditional electron-pair density · Information theory · Electron localization · Fermi hole · Coulomb hole

Introduction

It is widely recognized that a precise description of the physicochemical properties of molecular systems requires a clear understanding of the interaction existing between the electrons embedded within a nuclear field [1]. From the early principles of quantum mechanics, it is known that two main mechanisms govern the electron–electron interaction in atoms, molecules, and solids. The first one arises as a consequence of the classical electrostatic repulsion between electrons [2, 3], whereas the second one has a quantum origin, being associated to the antisymmetric character of Fermion wavefunctions and later formalized in the Pauli exclusion principle [4]. In order to provide context to the present work, it is important to annotate that both interaction mechanisms delineate regions around electrons where the probability to find another electron is negligible. These zones of diminished probability are customarily referred to as Coulomb hole and Fermi hole, and their accurate description is fundamental to unveil the intricate quantum behavior of electrons in molecular frames [5, 6].

Taking into consideration that the electron exclusion of the same spin electrons due to the Pauli principle is significantly more dominant than the electrostatic repulsion due to the correlation effects [7]; the Fermi hole markedly hinders the Coulomb hole. As a consequence, the former plays

Dedicated to Prof. Alejandro Toro-Labbé dear friend and collaborator.

✉ F. Javier Torres
jtorres@usfq.edu.ec

[1] Departamento de Ingeniería Química, Grupo de Química Computacional y Teórica (QCT-USFQ), Universidad San Francisco de Quito (USFQ), Diego de Robles S/N y Vía Interoceánica, Quito 170901, Ecuador

[2] Departamento de Matemática, Universidad San Francisco de Quito (USFQ), Diego de Robles S/N y Vía Interoceánica, Quito 170901, Ecuador

[3] Escuela de Ingeniería, Ciencia y Tecnología, Universidad del Rosario, 11171 Bogotá, Colombia

a pivotal role in defining the nature of chemical bonding, influencing the molecular geometry, and shaping the electronic configuration of both atomic and molecular orbitals [8]. Furthermore, it is well-established that the distribution of electrons within a molecule can be elucidated by examining the Fermi hole [9, 10]. In these regards, the electron localization is considered a crucial concept in quantum chemistry because it provides valuable insights on a number of important properties, including reactivity, electron-donating/withdrawing character, polarizability, and molecular spectroscopy [11–14]. Unfortunately, from the quantum mechanical point of view, the electron localization is a non-observable quantity, and for this reason, its univocal definition is not attainable. Consequently, different interpretations and methodologies for electron localization have been proposed. Bader's work on the electron accumulation/depletion based on the analysis of the Laplacian of the electron density [15], the electron localizability indicator (ELI) of Kohout [16], the domain-averaged Fermi holes (DAFH) proposed by Ponec et. al. [17], the local measurement of the Fermi hole covariance (FHC) introduced by Ayers [18], the maximum probability domains (MPD) defined by Savin [19], and the electron delocalization range function (EDR) of Frisch and co-workers [20] are some examples of electron localization methods. Among them, the electron localization function (ELF), derived from the leading term of the spherical average conditional pair density by Becke and Edgecombe [21], emerges as one of the most employed methods, being its wide acceptance associated not precisely to theoretical grounds but to its intuitive graphical representation, alignment with classical chemical theories, numerical stability regardless the level of theory, and accessibility in most of the modern quantum chemistry programs. The popularity of ELF has motivated further elaborations and interpretations. For instance, Savin interpreted this function using the kinetic energy density [22, 23], and later, Silvi and Savin proposed a chemical bond classification through ELF basin analysis, resulting in the merge of classical chemical concepts with quantum mechanics [24, 25]. Subsequent developments concerning ELF include a time-dependent version by Burnus, Marques, and Gross [26] and a variant for post Hartree–Fock methods presented by Matito, Silvi, Duran, and Solà [27].

Beyond the previous non-exhaustive list, methodologies that employ information theory (IT) concepts to obtain insights into electron localization have recently gained attention. This interest arises from the fact that IT provides adequate tools to extract and study the information content of probability distribution functions [28]. In the realm of quantum chemistry, the one-electron density, $\rho(\vec{r})$, computed from the wavefunction of a given N-electron system, $\psi\left(\vec{r_1}, \sigma_1, \vec{r_2}, \sigma_2, \vec{r_3}, \sigma_3, \ldots, \vec{r_N}, \sigma_N\right)$, represents an ideal source of information [29–31] for IT-based approaches. Some examples of the previous are the quantum information theory developed by Nalewajski [32, 33], the quantum mutual information method proposed by Henderson and Vedral [34], and the analysis of the negative entropy and information presented by Cerf and Adami [35]. In line with these recent advances, we have also contributed to the field by presenting in 2016, a method that allows the elucidation of the electron localization in molecules [36] as inferred from the information content of the conditional electron-pair probability [37]. A detailed description of our method and the implemented algorithm is provided in the next section of the present work. Here, it is only foretold that the Kullback-Liebler divergence defined for two probability functions, $P(x)$ and $Q(x)$, [38–40] as follows:)

$$\mathrm{KLD}[P(x)||Q(x)] = \int P(x)\log_2\left(\frac{P(x)}{Q(x)}\right)\mathrm{dx} \tag{1}$$

is the IT-based concept at the center of our method. In fact, its relevance is such that the computational program containing our algorithm is named **KLD** after it. Before continuing with the present section, it is worth mentioning that Eq. (1) represents a direct measure of the information missed when the exact or true density function $P(x)$ is approximated with the density function $Q(x)$, being this a property of great importance in our approach as discussed further in this document. It is also worth to mention that the Kullback–Leibler divergence has been also used as the central quantity in other studies [41]; however, these works do not use pair densities as the information source.

As a last remark, it should be noted that, despite the dominant role of the Fermi hole in the electronic structure of molecular systems, the necessity of knowing exact information on the Coulomb hole should not be disregarded. Indeed, such is its relevance that the genesis of density functional theory (DFT) itself can be attributed to the necessity of incorporating the information of the electron correlation into quantum mechanical models [42–44]. In the context of DFT, the Coulomb hole concept is employed to describe the deviation of $\rho(\vec{r})$ due to the electrostatic repulsion between the electrons (i.e., correlated particles) beyond the mean-field approach provided by the Hartree–Fock (HF) method [45, 46]. Thus, the inclusion of the Coulomb hole information in the correlation part of the different exchange–correlation functional forms [47–49] allow DFT methods to provide better estimates of the energetic and structural properties, chemical reactivity indexes, and electron-excitation energies [50–52]. The latter results have ensued in the wide acceptance and preference of DFT methodologies over the more computational-demanding ab initio approaches (i.e., Møller-Plesset, coupled cluster, configuration interaction). One important scenario where the Coulomb hole has a fundamental role is the study of non-covalent interactions [53]. Previous theoretical reports have shown that this particular type of interactions are the result of the interplay between

charge–transfer and dispersive components [54–57]. While the charge–transfer mechanism can be reliably described by means the exact-exchange included in the so-called hybrid DFT functionals (e.g., B3LYP), the dispersive forces are solely accessible through a precise description of the electron correlation. In these regards, a remarkable feature of the **KLD** program relies in the fact that it allows the separation of the Coulomb hole from the whole exchange and correlation landscape, providing the conditions for an independent analysis of the dispersion in molecular systems [58].

The aim of this work is to present detailed information of the program **KLD**, making emphasis in its theoretical framework and logic behind the algorithm. The latter is covered in "Theoretical background", whereas "Results and discussion" contains a thorough discussion on the results obtained with the latest version of our code, which includes a recent implementation proposed to increase the computational efficiency in calculating the Fermi hole as well as the Coulomb hole of molecular systems. Some conclusions and future perspectives are shared in "Conclusions and future perspectives."

Theoretical background

The same-spin electron-pair conditional probability density

As previously commented, the fundamental quantity in our method is the electron-pair density. Therefore, a brief discussion on the second-order reduced density matrix (2-RDM) and other related electron-pair density functions is presented in this section.

The 2-RDM, $\Gamma\left(\overrightarrow{\chi}_1,\overrightarrow{\chi}_2;\overrightarrow{\chi}'_1,\overrightarrow{\chi}'_2\right)$, is computed by integrating the wavefunction, $\Psi\left(\overrightarrow{\chi}_1,\overrightarrow{\chi}_2,\overrightarrow{\chi}_3,\ldots,\overrightarrow{\chi}_N\right)$, of a N-electron system over the spatial and spin states $\overrightarrow{\chi}_i=\left(\vec{r}_i,\vec{\sigma}_i\right)$ of all the electrons but the first two:

$$\Gamma\left(\overrightarrow{\chi}_1,\overrightarrow{\chi}_2;\overrightarrow{\chi}'_1,\overrightarrow{\chi}'_2\right)=\frac{N(N-1)}{2}\int d\overrightarrow{\chi}_3\cdots\int d\overrightarrow{\chi}_N\Psi^*\left(\overrightarrow{\chi}_1,\overrightarrow{\chi}_2\ldots,\overrightarrow{\chi}_N\right)\Psi\left(\overrightarrow{\chi}'_1,\overrightarrow{\chi}'_2\ldots,\overrightarrow{\chi}'_N\right) \tag{2}$$

where the factor $\frac{N(N-1)}{2}$ is the normalization with respect to the number of pairs of electrons.

By imposing the conditions $\overrightarrow{\chi}_1=\overrightarrow{\chi}'_1$ and $\overrightarrow{\chi}_2=\overrightarrow{\chi}'_2$ in Eq. (2), the so-called electron-pair density is defined $\Gamma\left(\overrightarrow{\chi}_1,\overrightarrow{\chi}_2;\overrightarrow{\chi}_1,\overrightarrow{\chi}_2\right):=\Gamma\left(\overrightarrow{\chi}_1,\overrightarrow{\chi}_2\right)$. The latter can be interpreted as the probability density of finding a first electron with spin state σ_1 at position $\vec{r}_1$, and, at the same time, a second electron with spin σ_2 at position $\vec{r}_2$. Note that, in the general form of $\Gamma\left(\overrightarrow{\chi}_1,\overrightarrow{\chi}_2\right)$, the spin states σ_1 and σ_2 can be either parallel or antiparallel from which the following arises:

$$\Gamma\left(\overrightarrow{\chi}_1,\overrightarrow{\chi}_2\right)=\Gamma^{\alpha,\alpha}\left(\vec{r}_1,\vec{r}_2\right)\alpha_1\alpha_2+\Gamma^{\alpha,\beta}\left(\vec{r}_1,\vec{r}_2\right)\alpha_1\beta_2\\ \Gamma^{\beta,\alpha}\left(\vec{r}_1,\vec{r}_2\right)\beta_1\alpha_2+\Gamma^{\beta,\beta}\left(\vec{r}_1,\vec{r}_2\right)\beta_1\beta_2 \tag{3}$$

Of the terms at the right-hand side of Eq. (3), the first and the fourth are denoted as same-spin electron-pair densities, which are quantities that contain all the information regarding the Pauli exclusion principle (i.e., Fermi hole). For restricted wavefunctions (i.e., closed-shell systems), the $\Gamma^{\alpha,\alpha}\left(\vec{r}_1,\vec{r}_2\right)=\Gamma^{\beta,\beta}\left(\vec{r}_1,\vec{r}_2\right)$ equality holds. Moreover, since the latter ones are obtained after integrating the spin coordinates out, either $\Gamma^{\alpha,\alpha}\left(\vec{r}_1,\vec{r}_2\right)$ or $\Gamma^{\beta,\beta}\left(\vec{r}_1,\vec{r}_2\right)$ can be simply represented as $\Gamma^{\sigma,\sigma}\left(\vec{r}_1,\vec{r}_2\right)$ without losing generality.

Some important properties of $\Gamma^{\sigma,\sigma}\left(\vec{r}_1,\vec{r}_2\right)$ are well-established:

- The same-spin electron-pair density vanishes at electron coalescence, $\vec{r}_1=\vec{r}_2$,

$$\Gamma^{\sigma,\sigma}\left(\vec{r}_1,\vec{r}_1\right)=0 \tag{4}$$

- $\Gamma^{\sigma,\sigma}\left(\vec{r}_1,\vec{r}_2\right)$ results normalized to the number of σ electron-pairs after integration over the electron positions,

$$\int d\vec{r}_1\int d\vec{r}_2\Gamma^{\sigma,\sigma}\left(\vec{r}_1,\vec{r}_2\right)=\frac{N^\sigma(N^\sigma-1)}{2} \tag{5}$$

- The σ-spin one-electron density is recovered after integration over one of the electron positions,

$$\int d\vec{r}_2\Gamma^{\sigma,\sigma}\left(\vec{r}_1,\vec{r}_2\right)=\frac{(N^\sigma-1)}{2}\rho^\sigma\left(\vec{r}_1\right) \tag{6}$$

It is important to comment that all these properties are in essence N-representability constraints. Moreover, there are some other interesting properties of the pair density, related to its N-representability. An appropriate compendium on this specific subject can be found elsewhere [59–62].

For the scope of the present work, it is useful to re-write the same-spin pair density as a product of the one-electron density and the conditional electron-pair density as follows:

$$\Gamma^{\sigma,\sigma}\left(\vec{r}_1,\vec{r}_2\right)=\frac{1}{2}\rho^\sigma\left(\vec{r}_1\right)\gamma^{\sigma,\sigma}_{\text{cond}}\left(\vec{r}_2|\vec{r}_1\right) \tag{7}$$

In this equation, $\gamma^{\sigma,\sigma}_{\text{cond}}\left(\vec{r}_2|\vec{r}_1\right)$ represents the probability of finding an electron with spin σ at the position $\vec{r}_2$, when the position $\vec{r}_1$ of another electron with the same σ spin is known with certainty. Thus, this quantity is closely related to the deviation of the electronic behavior from the mean-field model [63] as represented in the following expression:

$$\gamma^{\sigma,\sigma}_{\text{cond}}\left(\vec{r}_2|\vec{r}_1\right)=\rho^\sigma\left(\vec{r}_2\right)+\eta^{\sigma,\sigma}_{\text{xc}}\left(\vec{r}_2|\vec{r}_1\right) \tag{8}$$

In the previous expression, $\eta^{\sigma,\sigma}_{\text{xc}}\left(\vec{r}_2|\vec{r}_1\right)$ is denoted as the same-spin hole density, illustrating the fact that this quantity is

always negative and defines zones of diminishing conditional density [9, 64, 65]. It must be also noted that the xc subindex included in the notation indicates that this probability distribution is directly associated to the electron exchange and same-spin correlation effects in the frame of DFT.

By inspection of the above expressions, it results self-explanatory that $\gamma^{\sigma,\sigma}_{\text{cond}}\left(\vec{r}_2|\vec{r}_1\right)$ differs from $\rho^{\sigma}\left(\vec{r}_2\right)$ only by the hole density. Thus, the application of the Kullback–Leibler divergence shown in Eq. (1):

$$\text{KLD}_{\text{xc}}\left(\vec{r}_1\right) = \int d\vec{r}_2 \frac{\gamma^{\sigma,\sigma}_{\text{cond}}\left(\vec{r}_2|\vec{r}_1\right)}{N^{\sigma}-1}\log_2\left(\frac{\gamma^{\sigma,\sigma}_{\text{cond}}\left(\vec{r}_2|\vec{r}_1\right)/(N^{\sigma}-1)}{\rho^{\sigma}\left(\vec{r}_2\right)/N^{\sigma}}\right) \tag{9}$$

allows the quantity $\text{KLD}_{\text{xc}}\left(\vec{r}_1\right)$ to be defined as the missing information when $\gamma^{\sigma,\sigma}_{\text{cond}}\left(\vec{r}_2|\vec{r}_1\right)$ is approximated with $\rho^{\sigma}\left(\vec{r}_2\right)$, which, as observed in Eq. (8), is a direct measure of the electron exchange and correlation term, $\eta^{\sigma,\sigma}_{\text{xc}}\left(\vec{r}_2|\vec{r}_1\right)$. Note that the $N^{\sigma}-1$ and N^{σ} denominators are simply normalization factors for $\gamma^{\sigma,\sigma}_{\text{cond}}\left(\vec{r}_2|\vec{r}_1\right)$ and $\rho^{\sigma}\left(\vec{r}_2\right)$, respectively. $\text{KLD}_{\text{xc}}\left(\vec{r}_1\right)$ can be readily used to elucidate the electron localization in the direct molecular space by recognizing that large values of the function evaluated at $\vec{r}_1$ are associated with regions of the molecule with a significant count of electrons and *vice versa*. Finally, by considering the approximated scaling of $\text{KLD}_{\text{xc}}\left(\vec{r}_1\right)$ with respect to the number of electrons with σ spin, N^{σ}, the following quantity:

$$\Omega^{\sigma}_{\text{xc}}\left(\vec{r}_1\right) = (N^{\sigma}-1)\text{KLD}_{\text{xc}}\left(\vec{r}_1\right)f_{\text{cut}}\left(\vec{r}_1\right) \tag{10}$$

can be proposed as an electron localization descriptor. In Eq. (10), $f_{\text{cut}}\left(\vec{r}_1\right) = \frac{1}{2}\left(1.0 + \text{ERF}\left[\frac{1}{2}\log_{10}\left(\frac{\rho(\vec{r}_1)}{\rho_{\text{cut}}}\right)\right]\right)$ is a cut-off function that monotonically decreases to zero when the position $\vec{r}_1$ is characterized by a negligible value of the one-electron density (i.e., $\rho\left(\vec{r}_1\right) = 1.0\times10^{-3}$).

As anticipated previously in this work, an important advantage of our method is that it allows the separation of the Coulomb and Fermi holes [58], so that the analysis of each term can be carried out independently. For this purpose, it has to be considered that the amount of information contained in $\gamma^{\sigma,\sigma}_{\text{cond}}\left(\vec{r}_1|\vec{r}_2\right)$ is directly related to the accuracy of the wavefunction employed to compute $\Gamma^{\sigma,\sigma}\left(\vec{r}_1,\vec{r}_2\right)$. In these regards, it is well-known that the electron-pair density can be obtained from a single-determinant wavefunction as follows [65, 66]:

$$\Gamma^{\sigma,\sigma}\left(\vec{r}_1,\vec{r}_2\right) = \frac{1}{2}\left(\rho^{\sigma}\left(\vec{r}_1\right)\rho^{\sigma}\left(\vec{r}_2\right) - \left|\gamma^{\sigma,\sigma}\left(\vec{r}_1,\vec{r}_2\right)\right|^2\right) \tag{11}$$

where $\gamma^{\sigma,\sigma}\left(\vec{r}_1,\vec{r}_2\right)$ is the first-order reduced density matrix (1-RDM), and its explicit definition from a single-determinant wavefunction can be found elsewhere [67]. From the latter context, it is clear that applying a Hartree–Fock wavefunction in Eqs. (7) and (11) yields, by definition, a conditional electron-pair density containing only the exchange effects of a molecular system, $\gamma^{\sigma,\sigma}_{\text{cond},x}\left(\vec{r}_2|\vec{r}_1\right)$, which subsequently produces an exchange only hole density, $\eta^{\sigma,\sigma}_{x}\left(\vec{r}_2|\vec{r}_1\right)$, by means of Eq. (8). Conversely, the calculation of a hole density containing appropriate information on the electron correlation effects requires the use of a multi-configurational wavefunctions, such as those resulting from Møller-Plesset (MPn), coupled cluster (CC), or configuration interaction (CI) levels of theory. Because $\Gamma^{\sigma,\sigma}\left(\vec{r}_1,\vec{r}_2\right)$ cannot be exactly written in terms of the reduced first-order density in post-HF methods (e.g., MP2, CC, CI), our method relies on an approximated ansatz for $\Gamma^{\sigma,\sigma}\left(\vec{r}_1,\vec{r}_2\right)$ that depends on the natural orbital expansion and their associated occupation numbers which are accessible in most electronic structure programs. Derived from the latter, $\Gamma^{\sigma,\sigma}\left(\vec{r}_1,\vec{r}_2\right)$ depends on three contributions: (i) an uncorrelated Coulomb part, (ii) a Hartree–Fock like exchange part, and (iii) a correlation correction that restore the sum rules of the pair density:

$$\Gamma^{\sigma,\sigma}\left(\vec{r}_1,\vec{r}_2\right) = \frac{1}{2}\left(\rho^{\sigma}(\vec{r}_1)\rho^{\sigma}(\vec{r}_2) - \left|\gamma^{\sigma\sigma}\left(\vec{r}_1,\vec{r}_2\right)\right|^2 + \Gamma_C\left(\vec{r}_1,\vec{r}_2\right)\right) \tag{12}$$

This approximation for $\Gamma^{\sigma,\sigma}\left(\vec{r}_1,\vec{r}_2\right)$ is thoroughly discussed in Ref. [58] ("Section 2.3: Approximated pair-density expression from the natural orbitals"). For further details on the latter, the interested reader is also referred to Ref. [68]; here, it is only commented that upon obtaining $\eta^{\sigma,\sigma}_{\text{xc}}\left(\vec{r}_2|\vec{r}_1\right)$, which contains the information of the electron correlation provided by a given multi-configurational wavefunction, the $\text{KLD}_{c}\left(\vec{r}_1\right)$ function can be defined as follows:

$$\text{KLD}_{c}\left(\vec{r}_1\right) = -\int d\vec{r}_2\eta^{\sigma,\sigma}_{\text{xc}}\left(\vec{r}_2|\vec{r}_1\right)\log_2\left(\frac{\eta^{\sigma,\sigma}_{\text{xc}}\left(\vec{r}_2|\vec{r}_1\right)}{\eta^{\sigma,\sigma}_{x}\left(\vec{r}_2|\vec{r}_1\right)}\right) \tag{13}$$

It must be noted that, in contrast to Eq. (9), the latter expression considers hole densities instead of conditional electron-pair ones. Therefore, the errors derived for the intrinsic differences between the one-electron density computed from different ab initio levels (i.e., HF and MP2) are mitigated. Moreover, the minus sign at front of the integral is included to cancel the negative sign in the kernel of the expression. In analogy with the $\text{KLD}_{\text{xc}}\left(\vec{r}_1\right)$ case, the $\text{KLD}_{c}\left(\vec{r}_1\right)$ function can be interpreted as the missing information when the exchange–correlation hole density, $\eta^{\sigma,\sigma}_{\text{xc}}\left(\vec{r}_2|\vec{r}_1\right)$, is approximated with the Hartree–Fock exchange-only $\eta^{\sigma,\sigma}_{x}\left(\vec{r}_2|\vec{r}_1\right)$ hole density, being therefore, a direct measure of the Coulomb hole. Taking this into consideration, a descriptor for the electron-correlation, $\Omega^{\sigma}_{c}\left(\vec{r}_1\right)$, in molecular systems can be defined in analogy with Eq. (10):

$$\Omega^{\sigma}_{c}\left(\vec{r}_1\right) = (N^{\sigma}-1)\text{KLD}_{c}\left(\vec{r}_1\right)f_{\text{cut}}\left(\vec{r}_1\right) \tag{14}$$

The algorithm within the KLD code

Whether the analysis of the Fermi hole or the Coulomb hole is to be conducted, the step with the highest computational

cost in our method is the calculation of $\mathrm{KLD}_{xc}(\vec{r}_1)$ or $\mathrm{KLD}_c(\vec{r}_1)$, respectively. Therefore, the algorithm within the **KLD** code was implemented to evaluate, as efficiently as possible, the integrals shown in Eqs. (9) and (13) by employing our adapted version of the Becke's approach [69].

Figure 1 contains a schematic representation of the different routines in our program. As observed in the figure, **KLD** has two sources of input data. The first one consists of parameters provided by the user to define the position of a reference electron (i.e., labeled as $\vec{r}_1$ in the equations above) in terms of an initial position $\vec{r}_{0,0,0} = (x_0, y_0, z_0)$, an incremental distance step s, and the total number of points to be considered in each Cartesian direction N_x, N_y, N_z. On the other hand, **KLD** reads a standard WFN file containing both the system specification defined by the total number of atoms NAT, the atomic number of each atom Z_A, and the atomic coordinates $\vec{R}_{Z_A}$, as well as the wavefunction expanded in terms of the occupied molecular orbitals NMO, which in turn are built as a linear combination of a given set of primitive basis functions NBS.

For a given position $\vec{r}_1$ (i.e., $\vec{r}_{0,0,0}$ at the first iteration), **KLD** uses the geometric data to decompose the molecular space into atomic basins by means of the Voronoi tessellation method [70]. For each basin, both the Gauss–Chebyshev and the Lebedev quadratures are employed to define a set of radial and angular points [71, 72], consisting of one weight and three spatial coordinates $\{w_i, x_i, y_i, z_i\} = \{w_i, \vec{r}_i\}$, which are stored in memory to be used in the subsequent stages of the program. It is important to comment that **KLD** considers different number of radial points for atoms belonging to different rows of the periodic table. Specifically, 20, 25, 30, 35, 40, and 45 radial points are employed for atoms in the first, second, third, fourth, fifth, and sixth rows, respectively. In contrast, the number of angular points is kept constant for each radial distance and are defined on the basis of the standard rules within the Lebedev scheme [73, 74]. For the sake of clarity, the position of the reference electron is distinguished from the position of the quadrature points by adopting two font types in Fig. 1. Italic $\vec{r}_1$ and normal $\vec{r}_i$ typefaces are employed to refer to the former and latter, respectively. Both the position of the reference electron, $\vec{r}_1$, and the set of quadrature points, $\{\vec{r}_i\}$, are employed to compute all the density functions (i.e., $\rho^\sigma(\vec{r}_i)$, $\Gamma^{\sigma,\sigma}(\vec{r}_i|\vec{r}_1)$, $\gamma^{\sigma,\sigma}_{\mathrm{cond}}(\vec{r}_i|\vec{r}_1)$, $\eta^{\sigma,\sigma}_{\mathrm{xc}}(\vec{r}_i|\vec{r}_1)$, and $\eta^{\sigma,\sigma}_{x}(\vec{r}_i|\vec{r}_1)$) required to evaluate the kernels defining the $\mathrm{KLD}_{xc}(\vec{r}_1)$ and $\mathrm{KLD}_c(\vec{r}_1)$ functions. In the final step of the algorithm, the integrals shown in Eqs. (9) and (13) are numerically evaluated to subsequently compute the Fermi hole descriptor, $\Omega^\sigma_{xc}(\vec{r}_1)$, or the Coulomb hole descriptor, $\Omega^\sigma_c(\vec{r}_1)$. The value of either function is printed in the output file of the program together with the position of the reference electron $\vec{r}_1$, then this position is updated $\vec{r}_1 = \vec{r}_{IJK} = (x_0 + I \times s, y_0 + J \times s, z_0 + K \times s)$ to repeat the numerical integration and the calculation of the descriptors until the conditions $I = N_x$, $J = N_y$, and $K = N_z$ are fulfilled. Once the calculation of either descriptor is completed for all the $N_x \times N_y \times N_z$ total points, the results of our program can be plotted by building a standard cube file with the results.

It is precise to comment that the reliability of our method as electron localization descriptor has been discussed in previous works [36] on the basis of a thorough comparison with ELF results. Here, it is only reminded that, while the average value of ELF is observed to be rather constant for the different basins of a molecular system, the average value of our descriptor varies depending on the specific characteristics of each basin (i.e., chemical environment). Therefore, it can be stated that our method provides an appreciable amount of additional information on the electron localization of molecules when contrasted with ELF. Consider as an example the H_2O molecule where the average values of $\Omega^\sigma_{xc}(\vec{r}_1)$ are 0.781, 0.418, and 0.382 e for the O core, O–H bond, and O lone pairs, respectively. As commented, this clearly shows the value variability with respect to the localization degree of each basin. However, as shown in Fig. S1 (Supplementary Information Material), the ELF results and those obtained with our method result qualitatively similar. An alike behavior was observed for other molecules (see for instance Fig. 2) as properly presented in Refs. [36, 68]. As a final important comment on these regards, it has been also observed that our electron localization method shows stability with respect to the basis set size, producing negligible variability when the basis set is consistently increased. The latter can be noted in Table S1 (Supplementary Information Material) where the $\Omega^\sigma_{xc}(\vec{r}_1)$ results computed at different levels of theory (i.e., basis set size) for the molecules depicted in Fig. 2 are contrasted with a reference calculation at the level of theory employed in the present study (see below).

Although the promising results of our methodology, the extended applicability of the program **KLD** is undoubtedly limited by the excessive computational cost involved in the calculation of the $\Gamma^{\sigma,\sigma}(\vec{r}_i|\vec{r}_1)$ and $\gamma^{\sigma,\sigma}_{\mathrm{cond}}(\vec{r}_i|\vec{r}_1)$ (or $\eta^{\sigma,\sigma}_{\mathrm{xc}}(\vec{r}_i|\vec{r}_1)$, and $\eta^{\sigma,\sigma}_{x}(\vec{r}_i|\vec{r}_1)$) functions required for the numerical integration of (9) and (13). In these regards, it must be noted that, at each iteration of the algorithm, the position of the reference electron is updated (i.e., $\vec{r}_1 = \vec{r}_{IJK} = (x_0 + I \times s, y_0 + J \times s, z_0 + K \times s)$), and the aforementioned density functions must be re-computed. Considering that the quality of the results relies on the density of points considered in the molecular space (i.e., the grid of Cartesian points), appropriate values of the incremental distance step, s, (refer to Fig. 1) lie within the 0.15–0.25 a.u. range, producing hundreds of thousands of points to be evaluated as positions for the reference electron. Furthermore, it has been determined that the wall-clock time associated with these operations directly depends on both the total number of molecular orbitals, NMO, as well as the primitive basis

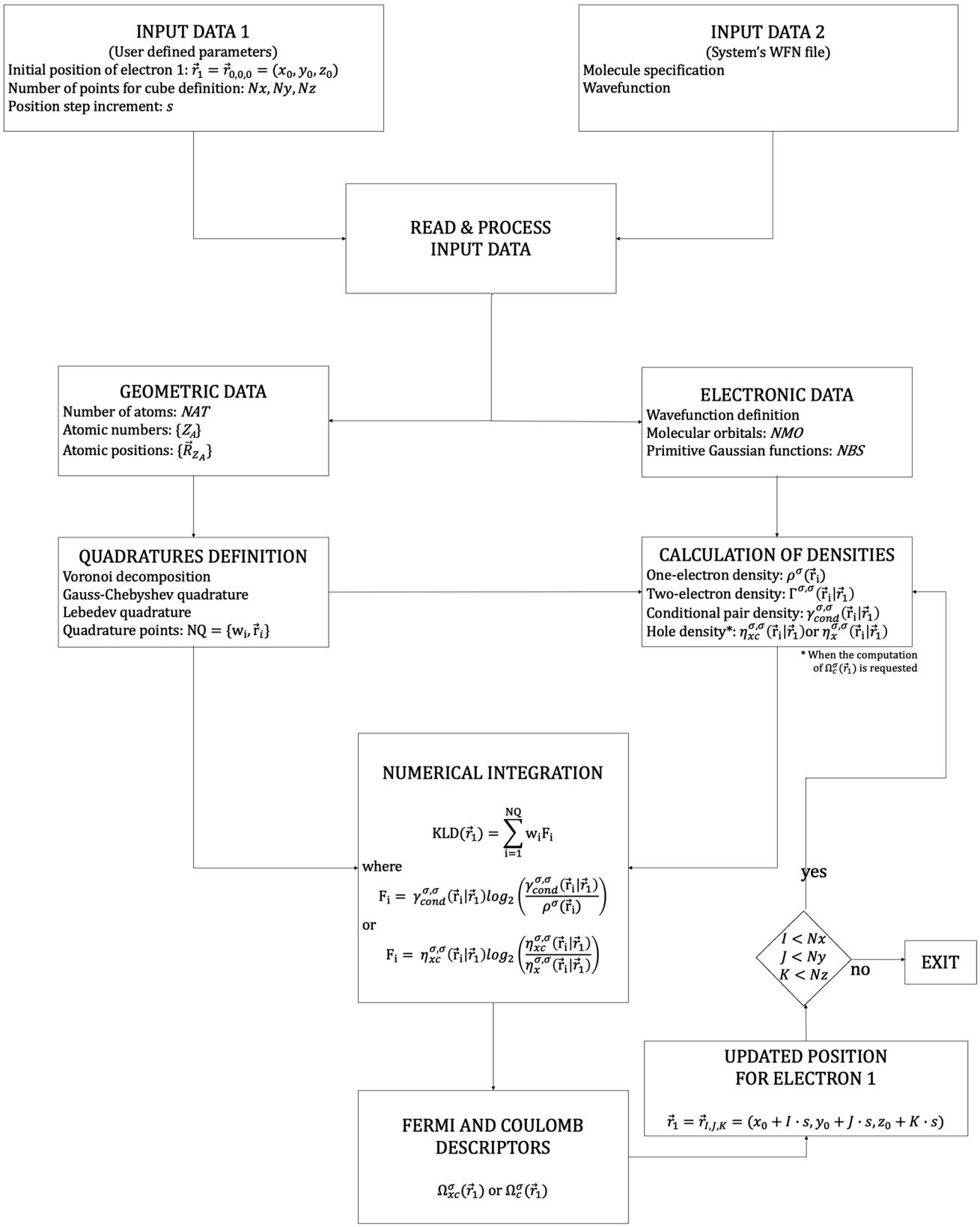

Fig. 1 Schematic diagram of the basic algorithm within **KLD** program

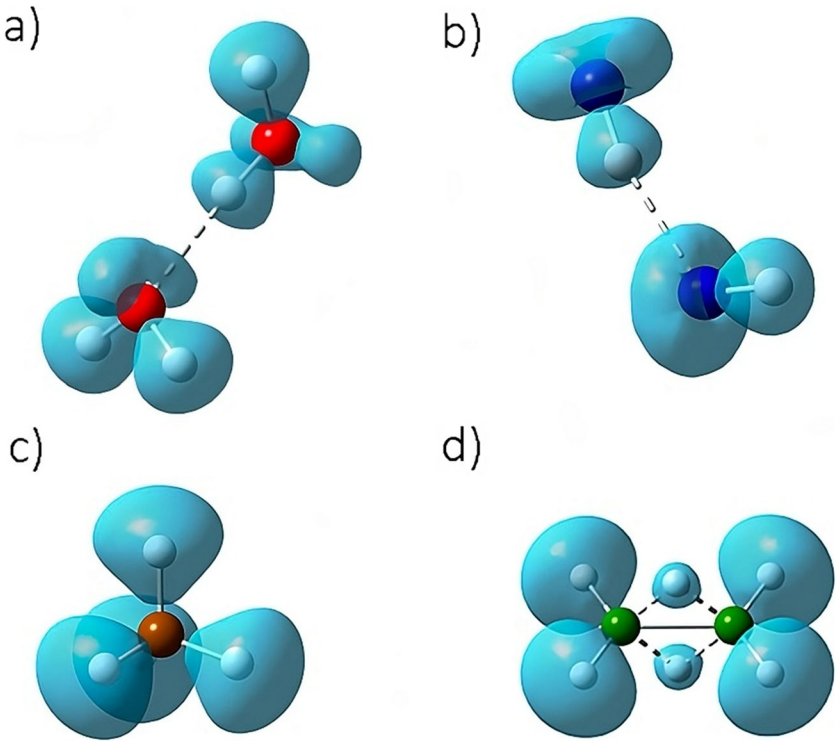

Fig. 2 Graphical representation of $\Omega^{\sigma}_{\chi_c}(\vec{r}_1)$ for **a** $[H_2O]_2$, **b** $[HF]_2$, **c** CH_4, and **d** B_2H_6. The surfaces are obtained with 0.59, 0.46, 0.55, and 0.75 bits times electron (bte) isovalues for a, b, c, and d, respectively. Hydrogen, oxygen, fluor, and beryllium atoms are represented with white, red, blue, brown, and green spheres, respectively

functions, NBS, employed to expand each orbital. With the purpose to improve the performance of our code, a strategy to accelerate the calculation of the $\Omega^{\sigma}_{\mathrm{xc}}(\vec{r}_1)$ and $\Omega^{\sigma}_{c}(\vec{r}_1)$ descriptors has been implemented in the latest version of our computational code. Before describing the details of the current version, it must be mentioned that the program **KLD** fully exploits the capabilities of modern hardware (see below). Thus, the definition of an adequate improvement strategy requires the analysis of Eqs. (10) and (14). As indicated in the previous subsection, the effect of the cut-off function $f_{\mathrm{cut}}(\vec{r}_1)$ is to gradually reduce the value of either descriptor until reaching a value of zero for small enough values of the one-electron density $\rho(\vec{r}_1)$. As observed in Fig. 1, the latter takes place at the final stage of each iteration of the program, which implies that all the computationally expensive functions and integrals have been already calculated regardless the significance of the $\Omega^{\sigma}_{\mathrm{xc}}(\vec{r}_1)$ and $\Omega^{\sigma}_{c}(\vec{r}_1)$ results. In order to avoid the evaluation of negligible values of the descriptors, the latest version of the **KLD** program computes $\rho(\vec{r}_1)$ at each of the $N_x \times N_y \times N_z$ grid points as one of the first steps of its algorithm. Then, a logical conditional structure is employed to decide whether $\mathrm{KLD}_{\mathrm{xc}}(\vec{r}_1)$ or $\mathrm{KLD}_{c}(\vec{r}_1)$ integrals are to be computed at the $\vec{r}_1 = \vec{r}_{\mathrm{IJK}}$ position of the reference electron. The criterium to do so is defined by the user through a threshold value for the one-electron density $\rho_{\mathrm{thresh}}(\vec{r}_1) = 1 \times 10^{-N}$ (see Appendix: Input file structure). Therefore, the $\mathrm{KLD}_{\mathrm{xc}}(\vec{r}_1)$ or $\mathrm{KLD}_{c}(\vec{r}_1)$ integrals are computed if $\rho(\vec{r}_1) > \rho_{\mathrm{thresh}}(\vec{r}_1)$ at $\vec{r}_1 = \vec{r}_{\mathrm{IJK}}$; otherwise, the position of the reference electron is updated and the program goes to the next iteration. Naturally, the application of the latter strategy has a direct impact in the accuracy of the results. Too small values of N in the $\rho_{\mathrm{thresh}}(\vec{r}_1)$ definition produces poor $\Omega^{\sigma}_{\mathrm{xc}}(\vec{r}_1)$ or $\Omega^{\sigma}_{c}(\vec{r}_1)$ results; conversely, too large values of N constrict the effect of implemented acceleration strategy. To properly assess the influence of $\rho_{\mathrm{thresh}}(\vec{r}_1)$ on the $\Omega^{\sigma}_{\mathrm{xc}}(\vec{r}_1)$ results, a set of 30 molecular systems of different sizes and characteristics are investigated in the present work. This analysis will be presented and discussed in the next section. Here, it is only indicated that the same grid of points was employed for all the studied molecules by defining $\vec{r}_{0,0,0} = (-10, -9, -9)$ *a.u.*, $s = 0.2$ a.u., $N_x = 101$ and $N_y = N_z = 91$. In this manner, the wall-clock time is constrained to exclusively depend on the NMO and NBS of each molecule. Likewise, the effect of $\rho_{\mathrm{thresh}}(\vec{r}_1)$ on the $\Omega^{\sigma}_{c}(\vec{r}_1)$ results was evaluated by employing a subset containing the smaller systems of the original set of 30 molecules and a more modest grid of $71 \times 61 \times 61$ points. The latter is conducted due to larger number of molecular orbitals NMO when explicitly correlated wavefunctions are employed for the descriptor calculation.

Computational details

The first version of the **KLD** program was written using the high-performance Fortran (HPF) programming language, and the code was parallelized employing shared and distributed memory strategies as defined by standard OpenMPI

and MPICH directives [75], respectively. Although this first version allowed the fine-tuning of our method as reported in Ref. [36], its excessive computational time motivated the migration of the code to a Compute Unified Device Architecture (CUDA) version capable to utilize the computational power of NVIDIA Graphics Processing Units (GPUs) for processing purposes [76, 77]. Specifically, the base GPU version of **KLD** (Fig. 1) was developed using CUDA version 12.1 and GPUs with capability SM_80 or higher. All the **KLD** calculations presented in this work were carried out in a GPU NVIDIA GeForce RTX 4090 with 24 GB based on CPU Intel Core i9-14900 architecture.

Concerning the studied molecules, all the systems under investigation (see "Results and discussion") were fully optimized at the HF/6-31G(2df,p) level of theory as implemented in the GAUSSIAN16 Rev. C suit of programs [77]. Upon obtaining the equilibrium geometries, single-point calculations were carried out in order to produce the Hartree–Fock WFN files required by **KLD** to compute the Fermi hole descriptor $\Omega_{xc}^{\sigma}(\vec{r}_1)$. In a subsequent stage, supplementary single-point calculations were performed at the MP2/6-31G(2df,p)//HF/6-31G(2df,p) level to obtain Møller-Plesset WFN files, which are used in conjunction to their Hartree–Fock counterparts to compute the Coulomb hole descriptor $\Omega_{c}^{\sigma}(\vec{r}_1)$.

Results and discussion

In this section, the results for wall-clock time of **KLD** are reported and discussed, making emphasis in the time scaling of the code and the speedup of the latest version of our program as compared with a stable previous version. Before commenting on the results, it is important to mention that the time scaling is a computational parameter that indicates how much time a program takes to execute a specific algorithm if the size of the problem is increased by some factor. In other words, the time scaling allows the increase of the number of operations and the computational time to be estimated when the problem is doubled for example. On the other hand, the wall-clock time is simply defined as the elapsed time from the start to the end of a given **KLD** calculation. Along this work, the following wall-clock time function is employed:

$$\tau = A\mathcal{N}^{\alpha} \tag{15}$$

where τ is the wall-clock time (in seconds), $\mathcal{N}$ is the size of the computational problem (note that in KLD, $\mathcal{N}$ is the product of the number of molecular orbitals *NMO* times the number of primitive basis functions NBS), α is the scaling factor associated to the complexity of the algorithm, and A is a pre-factor constant, which is a measure of the intrinsic computational cost of the algorithm.

To test the validity of Eq. (15), τ and $\mathcal{N}$ data obtained from a set of 30 simple molecular systems (see "Computational details") are plotted on a logarithmic scale in Fig. 3 and reported in Table 1, where some interesting features are revealed. As observed in the figure, the variable $\mathcal{N}$ defining the problem size spans the 255–14,688 range, covering a region of about ~ 3 orders of magnitude. Moreover, it is worth of mentioning that the size problem is quite independent of the dimension or complexity of the molecule. For instance, the value $\mathcal{N}$ for the very simple F_2 molecule is comparable to the one obtained for the bulkier CH_2O system. As observed in Fig. 3, the log (τ) values obtained by means of the reference version of the **KLD** program lie above the results computed with the latest version of the code. This is expected since, in the former case, the $\Omega_{xc}^{\sigma}(\vec{r}_1)$ function is evaluated in all the points of the Cartesian grid, whereas in the latter version, only the points with one-electron density greater than ρ_{thresh} are considered as previously exposed in "The algorithm within the KLD code Section." It is also observed in the Fig. 3 that, for a given molecule, the larger the value of ρ_{thresh} the smaller the elpased time τ, which means that the acceleration of the code directly depends on the ρ_{thresh} setting. Concerning the correlation of τ and $\mathcal{N}$, a strong linearity ($R^2 \geq 0.97$) is determined between the two quantities. The results for the linear regression, $\log(\tau) = \alpha\log(\mathcal{N}) + \log(A)$, are summarized in Table 2, where it is observed that the value of the scaling factor α is roughly ~ 1.5 for all the cases. Of them, the best scaling factor (i.e., $\alpha_{min} = 1.42$), corresponds to the calculations with the reference version as well as the results obtained with the latest version by setting $\rho_{thresh} = 10^{-3}$. Surprisingly, these are also the slowest and the fastest codes. The fact that the scaling factor is similar in all cases means that the code speedup due to the use of ρ_{thresh} is mostly related to the pre-factor constant in Eq. (15). Considering the latter, the speedup between two different cases can be evaluated by adopting the following expression:

$$\text{Speedup} = \frac{\tau_2}{\tau_1} = \frac{A_2 N^{\alpha_2}}{A_1 N^{\alpha_1}} \approx \frac{A_2}{A_1}. \tag{16}$$

Along this work, τ_2 and τ_1 are the wall-clock time obtained with the reference version and the latest version of the code, respectively. In this manner, the resulting speedup values are always larger than 1; furthermore, the greater the speedup the better the acceleration in the code. The average speedup obtained for the new version of the code with different values of ρ_{thresh} lie between 8 and 10 (i.e., an order of magnitude), which certainly represents a significant decrease of the computational times.

Since the use of ρ_{thresh} undoubtedly introduces errors in the **KLD** calculation, it is convenient to assess its potential effect on the final results. For this purpose, the mean

Table 1 Mean absolute errors (MAE) calculated between the Fermi hole results obtained with the reference version of the program and the latest version considering different threshold densities (ρ_{thresh})

Molecule	NBS	MAE	MAE			
ρ_{thresh}			10^{-3}	10^{-4}	10^{-5}	10^{-6}
HF	5	51	0.000908	6.32E-05	6.03E-07	5.90E-09
H2O	5	58	0.001454	9.75E-05	8.54E-07	7.54E-09
NH3	5	65	0.002222	0.000146	1.21E-06	1.20E-08
CH4	5	72	0.003794	0.000253	2.03E-06	2.65E-08
CO	7	88	0.002771	0.000170	1.28E-06	1.01E-08
N2	7	88	0.002314	0.000131	9.63E-07	9.02E-09
CNH	7	95	0.002859	0.000178	1.36E-06	1.81E-08
C2H2	7	102	0.003575	0.000234	1.87E-06	2.35E-08
H2S	9	82	0.003812	0.000278	2.22E-06	1.40E-08
F2	9	88	0.001235	6.97E-05	4.76E-07	7.90E-09
CH2O	8	102	0.003399	2.17E-04	1.66E-06	2.36E-08
C2H4	8	116	0.004769	0.000311	2.43E-06	4.94E-08
(HF)2	10	102	0.001858	0.000114	9.58E-07	2.34E-08
B2H6	8	130	0.008330	0.000542	4.01E-06	9.38E-08
CH4O	9	116	0.004265	0.000277	2.18E-06	3.02E-08
(H2O)2	10	116	0.002989	0.000189	1.59E-06	3.17E-08
C2H6	9	130	0.005896	0.000380	2.97E-06	6.57E-08
C3H4	11	160	0.005576	0.000352	2.72E-06	5.64E-08
C2OH4	12	160	0.005276	0.000340	2.60E-06	6.10E-08
C3H6	12	174	0.006412	0.000411	3.18E-06	6.69E-08
C2OH6	13	174	0.006036	0.000378	2.87E-06	6.76E-08
C2N2	13	176	0.003811	0.000214	1.53E-06	3.77E-08
C3H8	13	188	0.007651	0.000479	3.61E-06	9.04E-08
C2HON	14	183	0.004179	0.000242	1.76E-06	4.62E-08
C3H2O	14	190	0.004874	0.000295	2.24E-06	6.44E-08
C4OH8	20	276	0.007985	0.000481	3.62E-06	1.25E-07
C6H6	21	306	0.007810	0.000476	3.70E-06	9.36E-08
C4O2H6	23	306	0.007699	0.000455	3.39E-06	1.09E-07
C7NOH7	32	445	0.009109	0.000544	4.13E-06	1.68E-07
C8NH9	32	459	0.010569	0.000638	4.86E-06	1.74E-07
Average			0.004781	2.99E-04	2.30E-06	5.38E-08

Table 2 Linear fit parameters of the plots presented in Fig. 3

Plot	Intercept	Slope	*R*-square (COD)
Reference version	-0.68061 ± 0.04	1.41987 ± 0.012	0.99771
$\rho_{thresh} = 10^{-3}$	-1.56420 ± 0.15	1.41775 ± 0.040	0.96957
$\rho_{thresh} = 10^{-4}$	-1.70439 ± 0.14	1.50077 ± 0.044	0.97600
$\rho_{thresh} = 10^{-5}$	-1.78186 ± 0.13	1.55846 ± 0.041	0.98017
$\rho_{thresh} = 10^{-6}$	-1.80207 ± 0.12	1.59359 ± 0.040	0.98206

absolute error (MAE) calculated between the results of reference version of the program and those obtained with different ρ_{thresh} values in the new version are summarized in Table 1 At first sight, it is observed that, as $\mathcal{N}$ gets larger, the MAE value increases linearly. The MAE values obtained with $\rho_{thresh} = 10^{-6}$ are the smallest ones for all the considered molecular systems, being in the order of 1×10^{-8} followed by the MAE results obtained with $\rho_{thresh} = 10^{-5}$ and $\rho_{thresh} = 10^{-4}$ which are 2 and 4 orders of magnitude larger, respectively (see last column of Table 1). The largest values of MAE are those obtained with $\rho_{thresh} = 10^{-3}$, which is an expected result because, for this particular value of the one-electron density threshold, the number of grid points employed for the calculation of $\Omega_{xc}^{\sigma}(\vec{r}_1)$ is greatly restricted. At this point, it is important to indicate that the important features of $\Omega_{xc}^{\sigma}(\vec{r}_1)$ are about 1×10^{-2} bte; thus, MAE values lower than 1×10^{-3} can be considered insignificant. Following the latter, it can be suggested that the latest version of the code with $\rho_{thresh} = 10^{-4}$ represents the best compromise between reliability and speedup.

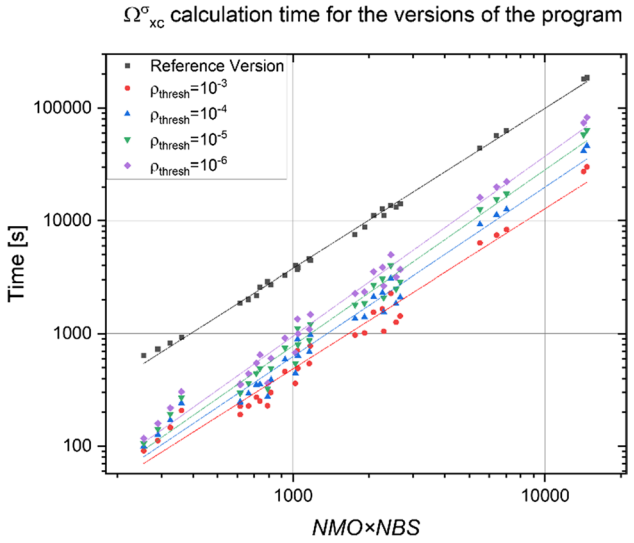

Fig. 3 Wall-clock time vs. number of molecular orbitals times base functions ($\mathcal{N}$) for the calculation of the Fermi hole using different versions of the program: reference version (black), and latest version with different values of ρ_{thresh}: 10^{-3} (red), 10^{-4} (blue), 10^{-5} (green), and 10^{-6} (purple)

Concerning the calculation of the Coulomb hole, an analysis similar to the previous one was conducted for the calculation of the $\Omega_c^{\sigma}(\vec{r}_1)$ function. Figure 4 reports the elapsed time τ as function of the problem size $\mathcal{N}$ obtained with the reference version of the code as well as the new version with $\rho_{\text{thresh}} = 10^{-4}$ and $\rho_{\text{thresh}} = 10^{-5}$. At first sight, the figure shows that the obtained values of $\log(\tau)$ are larger than those reported for the calculation of the $\Omega_{\text{xc}}^{\sigma}(\vec{r}_1)$ function (Fig. 3).

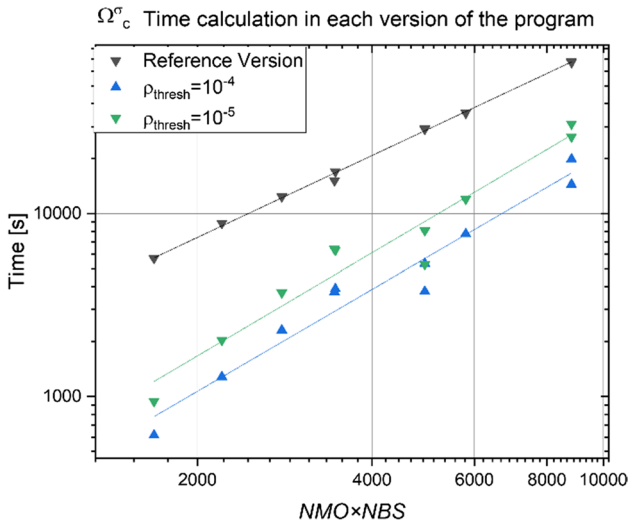

Fig. 4 Wall-clock time vs. number of molecular orbitals times base functions ($\mathcal{N}$) for the calculation of the Coulomb hole using different versions of the program: reference version (black), and latest version with different values of ρ_{thresh} :10^{-4}(blue) and 10^{-5}(green)

The latter is due to increase of the NMO to be considered when wavefunctions derived from correlated methods (i.e., MP2) are used in the **KLD** program. By comparing the NMO values of the HF and C_2H_4 molecules, for instance, data reported in Tables 1 and 3 indicates that the number of molecular orbitals that are employed for the calculation of $\Omega_c^{\sigma}(\vec{r}_1)$ are typically sevenfold those employed for the calculation of $\Omega_{\text{xc}}^{\sigma}(\vec{r}_1)$. As also observed in Fig. 4, the data obtained with the reference version of the code strongly correlate with a $R^2 = 0.99$ (see Table 4); in contrast, the values computed with the new version are more disperse, being characterized by R^2 values of about ~ 0.95. Moreover, the reference version of the code presents a scaling factor of 1.5, being comparable to the performance obtained for the calculation of the Fermi hole. On the other hand, the scaling factor of the accelerated code is close to ~ 1.9, suggesting that the latest version of the **KLD** program performs worse than the reference version when computing the $\Omega_c^{\sigma}(\vec{r}_1)$ function. Nevertheless, the calculated speedup Eq. (16) demonstrates that the new version of the code produces $\Omega_c^{\sigma}(\vec{r}_1)$ results faster in comparison to the reference version, being the acceleration close to 7–8 for small values of $\mathcal{N}$; however, the speedup is observed to decrease for $\mathcal{N} > 5000$.

The errors introduced in the calculation of $\Omega_c^{\sigma}(\vec{r}_1)$ due to the use of the pruned grid when ρ_{thresh} is employed are summarized in Table 3, where it is observed that the MAE values for the $\rho_{\text{thresh}} = 10^{-5}$ case are modest and close to a constant value of 5.0×10^{-7}. Clearly, the MAE values rise to almost ~ 2 orders of magnitud for the $\rho_{\text{thresh}} = 10^{-4}$ case. Although this increment in the errors observed, accurate enough results are also expected when computing $\Omega_c^{\sigma}(\vec{r}_1)$ by employing the $\rho_{\text{thresh}} = 10^{-4}$ condition in the latest version of the **KLD** program.

Table 3 Mean absolute errors (MAE) calculated between the Coulomb hole results obtained with the reference version of the program and the latest version considering different threshold densities(ρ_{thresh})

Molecule	NMO	NBS	MAE	
ρ_{thresh}			10^{-4}	10^{-5}
HF	33	51	1.15E-05	1.10E-07
H2O	38	58	1.93E-05	1.80E-07
NH3	43	65	2.59E-05	2.27E-07
H2S	42	82	0.000193	3.31E-07
CH4	48	72	3.28E-05	2.69E-07
N2	56	88	3.11E-05	2.42E-07
F2	56	88	1.04E-05	7.58E-07
CNH	61	95	4.62E-05	3.38E-07
(H2O)2	76	116	2.86E-05	2.36E-07
C2H4	76	116	0.000175	3.85E-07
Average			5.74E-05	3.08E-07

Table 4 Linear fit parameters of the plot presented in Fig. 4

Plot	Intercept	Slope	*R*-square (COD)
Reference version	−1.05218 ± 0.0854	1.49109 ± 0.0236	0.998000
$\rho_{\text{thresh}} = 10^{-4}$	−2.96531 ± 0.6050	1.87467 ± 0.1671	0.940022
$\rho_{\text{thresh}} = 10^{-5}$	−3.0862 ± 0 0.5122	1.8525 ± 0.1414	0.955402

Table 5 Speedup computed for the Coulomb hole calculation considering a lower and an upper bound as compared with the results obtained by using only a lower bound for the one-electron density ($\rho_{\text{thresh}} = 10^{-4}$)

Molecule ρ_{thresh}	NMO	NBS	Speedup 10^{-4}
HF	33	51	5.1
H2O	38	58	7.3
NH3	43	65	9.0
H2S	42	82	11.6
CH4	48	72	10.6
N2	56	88	11.7
F2	56	88	9.5
CNH	61	95	13.4
(H2O)2	76	116	12.9
C2H4	76	116	15.9

As a further acceleration strategy, the new version of the **KLD** program allows the user to define an upper bound for the one-electron density apart of the lower bound already commented above (see Appendix: Input file structure). This means that the $\Omega_c^{\sigma}(\vec{r}_1)$ function can be computed onto a constant value of $\rho(\vec{r}_1)$. This is particularly useful in the study of the strong non-covalent interactions (i.e., hydrogen bonds) present in stable molecular clusters. As reported in Ref. [58], $\Omega_c^{\sigma}(\vec{r}_1)$ computed for $[HF]_n$ clusters is observed to accumulate at the regions between the molecules and close to the critical point of the hydrogen bond. Therefore, it is reasonable to compute $\Omega_c^{\sigma}(\vec{r}_1)$ only at $\rho(\vec{r}_1)$ values associated to the borders of the individual molecules. As a result, a significant speedup is expected when the results of this strategy are compared with the cases where only a lower bound is considered. Table 5 contains the speedup results for each of the ten molecular systems employed to compute $\Omega_c^{\sigma}(\vec{r}_1)$. The data summarized in the table suggest a significant gain regarding the saving of computational time; moreover, the estimated speedup values are observed to increase with the problem size $\mathcal{N}$. The latter suggests that, for molecular systems characterized by a large number of values of NMO or NBS, this acceleration strategy is particularly useful because it enables the evaluation of many critical scenarios by projecting the Coulomb hole to a constant density value at a very low computational load.

Conclusions and future perspectives

Accurately predicting electron localization plots represents a major step to understanding the nature of the chemical bond. This paper described a significant advance towards making the calculation of electron localization from the information content of the conditional pair density computationally feasible. Two major improvements are presented here: first, the implementation of the Kullback–Leibler integral in a full GPU machinery, and second, the speedup by computing only non-negligible contributions to the integral. Future work will detail the efficacy of this method, which will unlock the ability to perform chemically accurate atomistic simulations. This research marks a significant advancement in the computational prediction of electron localization within molecular systems, crucial for understanding chemical bonding interactions. By leveraging the Kullback–Leibler divergence and advanced GPU acceleration, we have effectively streamlined the analysis of electron-pair densities, especially through the second-order reduced density matrix (2-RDM) and related functions. These enhancements enable rapid and accurate computations, utilizing modern computational frameworks to their full capacity. Our findings highlight the effectiveness of the computational strategy in independently analyzing Coulomb and Fermi holes, offering detailed insights into various physical phenomena within molecular systems. The adapted Becke's approach has been pivotal in refining our ability to analyze electron distributions swiftly across diverse molecular structures.

Furthermore, our approach not only elucidates the computational aspects but also underscores the separation of Coulomb and Fermi holes, facilitating an independent analysis of each component. This is crucial for isolating different physical phenomena within a molecular system, thereby providing a granular understanding of electron localization and molecular interactions. Future studies will delve deeper into the efficacy of these computational strategies, focusing on their impact on the speed and reliability of electron localization measurements. Our ongoing efforts aim to refine these techniques further, enhancing their application in chemically accurate atomistic simulations. This will open new avenues in molecular modeling, providing deeper insights into the electron behavior and interaction dynamics in complex molecular systems. Particularly, due to the interest in the use of pair density functional theory as a higher order DFT

[78–89], the present tool can be used to test some approximated expressions.

To further improve performance, especially for large molecular systems, future work could explore multi-GPU parallelization. This could involve distributing the computation across multiple GPUs, either within a single node or across multiple nodes in a cluster. Implementing efficient communication and load balancing strategies would be crucial for achieving good scalability. Future work could also explore integrating GPU-accelerated libraries with the **KLD** program to leverage their optimized routines for linear algebra, sparse matrix operations, and machine learning. This could help simplify the CUDA implementation and improve performance by using highly tuned library functions.

Appendix

KLD input file description

The input file of **KLD** is composed of three blocks of data:

- WFN
- CUBE
- PARAMETERS

WFN

The WFN block is used to define whether $\Omega_{\mathrm{xc}}^{\sigma}\left(\vec{r}_1\right)$ or $\Omega_{c}^{\sigma}\left(\vec{r}_1\right)$ is computed by indicating the number of wavefunctios to be used (see example below). Additionally, the name of each WFN file must be denoted.

NWFN	1 for $\Omega_{\mathrm{xc}}^{\sigma}\left(\vec{r}_1\right)$ (Fermi hole) calculation 2 for $\Omega_{c}^{\sigma}\left(\vec{r}_1\right)$ (Coulomb hole) calculation
FNWFN	name of WFN file 1 name of WFN file 2 (only for NWFN = 2)

CUBE

The first entry of the CUBE block indicates the name of the output file where the results of the calculation will be stored in the following format:

x_0	y_0	z_0	$\Omega_{\mathrm{xc}}^{\sigma}\left(\vec{r}_0\right)$ or $\Omega_{c}^{\sigma}\left(\vec{r}_0\right)$
x_1	y_1	z_1	$\Omega_{\mathrm{xc}}^{\sigma}\left(\vec{r}_1\right)$ or $\Omega_{c}^{\sigma}\left(\vec{r}_1\right)$
x_2	y_2	z_2	$\Omega_{\mathrm{xc}}^{\sigma}\left(\vec{r}_2\right)$ or $\Omega_{c}^{\sigma}\left(\vec{r}_2\right)$
⋮	⋮	⋮	⋮
x_i	y_i	z_i	$\Omega_{\mathrm{xc}}^{\sigma}\left(\vec{r}_i\right)$ or $\Omega_{c}^{\sigma}\left(\vec{r}_i\right)$
⋮	⋮	⋮	⋮
x_N	y_N	z_N	$\Omega_{\mathrm{xc}}^{\sigma}\left(\vec{r}_N\right)$ or $\Omega_{c}^{\sigma}\left(\vec{r}_N\right)$

The next line of the CUBE block contains the Cartesian coordinates of the electron 1 initial point $\vec{r}_0$ as well as the incremental distance step s. Finally, the third line of this block indicates the initial and final value of the N_x points that will be evaluated.

FNOUT	Name of output file
$\boldsymbol{x_0\ y_0\ z_0\ s}$	Initial position $\vec{r}_0$ and incremental distance step
$\boldsymbol{Nx_0\ Nx_F}$	Initial and final value of N_x

It is important to note that the $\boldsymbol{Nx_0\ Nx_F}$ values can be used to restart a **KLD** calculation interrupted by a computational crash or any other reason.

PARAMETERS

The final PARAMETERS block defines the values of different variables of control of the program including the ρ_{thresh} condition. These parameters are the following:

- Maximum weight: specifies a threshold for the set of weights ($\{\mathrm{w_i}\}$) of the Becke's algorithm that is employed for the numerical integration of either $\mathrm{KLD}_{\mathrm{xc}}\left(\vec{r}_1\right)$ or $\mathrm{KLD}_{c}\left(\vec{r}_1\right)$. A very small value must be set to avoid its effect on the calculation.
- Lebedev rule: controls the density of the angular points in the quadrature that defines the set of $\{\vec{\mathrm{r}}_\mathrm{i}\}$ positions.
- Number of bounds: defines the number of bounds with respect to the one-electron density to be employed during the evaluation of $\Omega_{\mathrm{xc}}^{\sigma}\left(\vec{r}_1\right)$ or $\Omega_{c}^{\sigma}\left(\vec{r}_1\right)$.
- Lower bound: defines the ρ_{thresh} value employed as lower bound.
- Upper bound: defines the ρ_{thresh} value employed as upper bound.

Note that the effect of the two first parameters are currently under investigation. Thus, caution is advised when using them.

MAXW	Threshold on the set of weights ($\{\mathrm{w_i}\}$)
LEBRUL	0 deactivate filter on angular points 1 activate filter on angular points
NBOUS	1 activate lower bound on one-electron density 2 activate lower and upper bounds on one-electron density
LOWRHOT	ρ_{thresh} value employed as lower bound
UPRHOT	ρ_{thresh} value employed as upper bound

In the following, a typical **KLD** input file is shown.

```
# WFN
1
'HF.wfn'
# CUBE
'out.txt'
-8.0 -8.0 -8.0 0.2
16 81
# PARAMETERS
0.00000000000000000001
0
1
0.001
```

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00894-024-06070-4.

Acknowledgements The present work has made use of the computational resources of the USFQ's and UR's High Performance Computing Systems.

Author contribution Valeria Bedoya: investigation, formal analysis, and writing original draft; Vladimir Rodríguez: software and validation; Luis Rincón: methodology, validation, writing, review, and editing; Cesar Zambrano: resources, funding acquisition, writing, review, and editing; Luis Seijas: investigation, formal analysis, and writing original draft; and F. Javier Torres: conceptualization, supervision, project administration, writing, review, and editing.

Funding This work was supported by the *2024 USFQ's POLIGRANTS program* (project ID: 24249) and the *2024 USFQ Collaboration Grants Program* (project ID: 23184).

Data availability No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

References

1. Nishimoto K (1966) An improvement in the calculation of electron repulsion integrals in Pariser-Parr-Pople theory-electron interactions in molecules. Theor Chim Acta 5:74–78. https://doi.org/10.1007/BF00527426/METRICS
2. Clementi E, Hofmann DWM (1994) Coulomb-Hole–Hartree–Fock functional. Int J Quantum Chem 52:849–865. https://doi.org/10.1002/QUA.560520413
3. Clementi E, Hofmann DWM (1995) Coulomb-hole-Hartree-Fock functional for molecular systems. J Mol Struct (Thoechem) 330:17–31. https://doi.org/10.1016/0166-1280(94)03814-2
4. Kaplan I (2020) Modern state of the Pauli exclusion principle and the problems of its theoretical foundation. Symmetry (Basel) 13:21
5. Rincon L, Torres FJ, Almeida R (2018) Is the Pauli exclusion principle the origin of electron localisation? Mol Phys 116:578–587
6. Buijse MA, Baerends EJ (1996) Fermi holes and Coulomb holes. In: Density Functional Theory of Molecules, Clusters, and Solids. Springer Netherlands, Dordrecht, pp 1–46
7. Boyd RJ, Coulson CA (1974) The Fermi hole in atoms. J Phys B: At Mol Phys 7:1805
8. Bultinck P, Cooper DL, Ponec R (2010) Influence of atoms-in-molecules methods on shared-electron distribution indices and domain-averaged fermi holes. J Phys Chem A 114:8754–8763. https://doi.org/10.1021/JP101707W
9. Bader RFW, Streitwieser A, Neuhaus A et al (1996) Electron delocalization and the Fermi hole. J Am Chem Soc 118:4959–4965
10. Shaik SS (2007) Is my chemical universe localized or delocalized? Is there a future for chemical concepts? New J Chem 31:2015
11. Gironés X, Ponec R (2006) Molecular quantum similarity measures from Fermi hole densities: modeling Hammett sigma constants. J Chem Inf Model 46:1388–1393
12. Ludeña EV, Ugalde JM, Lopez X et al (2004) A reinterpretation of the nature of the Fermi hole. J Chem Phys 120:540–547
13. Francisco E, Martín Pendás A, Blanco MA (2009) A connection between domain-averaged Fermi hole orbitals and electron number distribution functions in real space. J Chem Phys 131:124125
14. Bader RFW, Stephens ME (1975) Spatial localization of the electronic pair and number distributions in molecules. J Am Chem Soc 97:7391–7399
15. Bader RFW, Essén H (1984) The characterization of atomic interactions. J Chem Phys 80:1943–1960. https://doi.org/10.1063/1.446956
16. Kohout M (2004) A measure of electron localizability. Int J Quantum Chem 97:651–658. https://doi.org/10.1002/QUA.10768
17. Ponec R, Roithová J (2001) Domain-averaged Fermi holes–a new means of visualization of chemical bonds. Bonding in hypervalent molecules. Theor Chem Acc 105:383–392
18. Ayers PW (2005) Electron localization functions and local measures of the covariance. J Chem Sci 117:441–454. https://doi.org/10.1007/BF02708348
19. Mafra Lopes O, Braidä B, Causà M, Savin A (2011) Understanding maximum probability domains with simple models. In: Hoggan PE et al (eds) Advances in the Theory of Quantum Systems in Chemistry 173 and Physics, Progress in Theoretical Chemistry and Physics 22. Springer, pp 173–184
20. Janesko BG, Wiberg KB, Scalmani G, Frisch MJ (2016) Electron delocalization range in atoms and on molecular surfaces. J Chem Theory Comput 12:3185–3194. https://doi.org/10.1021/ACS.JCTC.6B00343
21. Becke AD, Edgecombe KE (1990) A simple measure of electron localization in atomic and molecular systems. J Chem Phys 92:5397–5403
22. Savin A, Becke AD, Flad J et al (1991) A new look at electron localization. Angew Chem, Int Ed Engl 30:409–412
23. Savin A, Jepsen O, Flad J et al (1992) Electron localization in solid-state structures of the elements: the diamond structure. Angew Chem, Int Ed Engl 31:187–188
24. Silvi B, Savin A (1994) Classification of chemical bonds based on topological analysis of electron localization functions. Nature 371:683–686
25. Savin A, Silvi B, Colonna F (1996) Topological analysis of the electron localization function applied to delocalized bonds. Can J Chem 74:1088–1096
26. Burnus T, Marques MAL, Gross EKU (2005) Time-dependent electron localization function. Phys Rev A 71:010501. https://doi.org/10.1103/PHYSREVA.71.010501
27. Matito E, Silvi B, Duran M, Sola M (2006) Electron localization function at the correlated level. J Chem Phys 125:024301
28. Reza FM (1994) An introduction to information theory. Dover Publications Inc., New York

29. Pratt LR, Hoffman GG, Harris RA (1988) Statistical theory of electron densities. J Chem Phys 88:1818–1823
30. He X, Lu T, Rong C et al (2023) Topological analysis of information-theoretic quantities in density functional theory. J Chem Phys 159:054112
31. Wilde MM (2013) Quantum information theory. Cambridge University Press
32. Nalewajski RF (2014) Quantum information approach to electronic equilibria: molecular fragments and non-equilibrium thermodynamic description. J Math Chem 52:1921–1948
33. Nalewajski RF (2014) Quantum information descriptors and communications in molecules. J Math Chem 52:1292–1323
34. Henderson L, Vedral V (2001) Classical, quantum and total correlations. J Phys A Math Gen 34:6899
35. Cerf NJ, Adami C (1997) Negative entropy and information in quantum mechanics. Phys Rev Lett 79:5194–5197. https://doi.org/10.1103/PHYSREVLETT.79.5194
36. Urbina AS, Torres FJ, Rincon L (2016) The electron localization as the information content of the conditional pair density. J Chem Phys 144:244104
37. Rincón L, Almeida R, Contreras PL, Torres FJ (2015) The information content of the conditional pair probability. Chem Phys Lett 635:116–119
38. Eguchi S, Copas J (2006) Interpreting Kullback-Leibler divergence with the Neyman-Pearson lemma. J Multivar Anal 97:2034–2040
39. Belov DI, Armstrong RD (2011) Distributions of the Kullback-Leibler divergence with applications. Br J Math Stat Psychol 64:291–309
40. Pérez-Cruz F (2008) Kullback-Leibler divergence estimation of continuous distributions. In: 2008 IEEE international symposium on information theory. IEEE, Toronto, pp 1666–1670
41. Liu S (2019) Identity for Kullback-Leibler divergence in density functional reactivity theory. J Chem Phys 151:141103
42. Burke K, Wagner LO (2013) DFT in a nutshell. Int J Quantum Chem 113:96–101
43 Gritsenko OV, Schipper PRT, Baerends EJ (1997) Exchange and correlation energy in density functional theory. Comparison of accurate DFT quantities with traditional Hartree-Fock based ones and generalized gradient approximations for the molecules Li2, N2, F2. J Chem Phys 107:5007
44. Himmetoglu B, Floris A, De Gironcoli S, Cococcioni M (2014) Hubbard-corrected DFT energy functionals: the LDA+ U description of correlated systems. Int J Quantum Chem 114:14–49
45. Slater JC (1951) A simplification of the Hartree-Fock method. Phys Rev 81:385–390. https://doi.org/10.1103/PHYSREV.81.385
46. Lykos P, Pratt GW (1963) Discussion on the Hartree-Fock approximation. Rev Mod Phys 35:496–501. https://doi.org/10.1103/REVMODPHYS.35.496
47. Scuseria GE, Staroverov VN (2005) Progress in the development of exchange-correlation functionals. Theory and applications of computational chemistry. Elsevier, pp 669–724
48. Koch W, Holthausen MC (2015) A chemist's guide to density functional theory. Wiley
49. Menconi G, Wilson PJ, Tozer DJ (2001) Emphasizing the exchange-correlation potential in functional development. J Chem Phys 114:3958–3967
50. Thanthiriwatte KS, Hohenstein EG, Burns LA, Sherrill CD (2011) Assessment of the performance of DFT and DFT-D methods for describing distance dependence of hydrogen-bonded interactions. J Chem Theory Comput 7:88–96. https://doi.org/10.1021/CT100469B
51. Karton A, Gruzman D, Martin JML (2009) Benchmark thermochemistry of the CnH2n+2 alkane isomers (n = 2–8) and performance of DFT and composite ab initio methods for dispersion-driven isomeric equilibria. J Phys Chem A 113:8434–8447. https://doi.org/10.1021/JP904369H
52. Kuta J, Patchkovskii S, Zgierski MZ, Kozlowski PM (2006) Performance of DFT in modeling electronic and structural properties of cobalamins. J Comput Chem 27:1429–1437. https://doi.org/10.1002/JCC.20454
53. Contreras-García J, Boto RA, Izquierdo-Ruiz F et al (2016) A benchmark for the non-covalent interaction (NCI) index or… is it really all in the geometry? Theor Chem Acc 135:1–14. https://doi.org/10.1007/S00214-016-1977-7
54. Su P, Li H (2009) Energy decomposition analysis of covalent bonds and intermolecular interactions. J Chem Phys 131:014102
55. Azar RJ, Head-Gordon M (2012) An energy decomposition analysis for intermolecular interactions from an absolutely localized molecular orbital reference at the coupled-cluster singles and doubles level. J Chem Phys 136:024103
56. Schneider WB, Bistoni G, Sparta M et al (2016) Decomposition of intermolecular interaction energies within the local pair natural orbital coupled cluster framework. J Chem Theory Comput 12:4778–4792
57. Fedorov DG, Kitaura K (2007) Pair interaction energy decomposition analysis. J Comput Chem 28:222–237
58. Rincon L, Javier Torres F, Becerra M et al (2019) On the separation of the information content of the Fermi and Coulomb holes and their influence on the electronic properties of molecular systems. Mol Phys 117:610–625. https://doi.org/10.1080/00268976.2018.1530462
59. Ayers PW, Davidson ER (2006) Necessary conditions for the N-representability of pair distribution functions. Int J Quantum Chem 106:1487–1498. https://doi.org/10.1002/QUA.20880
60. Ayers PW, Davidson ER (2007) Linear inequalities for diagonal elements of density matrices. Adv Chem Phys 134:443–483
61 Ayers PW (2006) Using classical many-body structure to determine electronic structure: an approach using k -electron distribution functions. Phys Rev A 74:042502. https://doi.org/10.1103/PHYSREVA.74.042502
62. Pistol ME (2004) N-representability of two-electron densities and density matrices and the application to the few-body problem. Chem Phys Lett 400(4–6):548–552
63. Golse F, Mouhot C, Paul T (2016) On the mean field and classical limits of quantum mechanics. Commun Math Phys 343:165–205. https://doi.org/10.1007/S00220-015-2485-7
64. Buijse MA, Baerends EJ (2002) An approximate exchange-correlation hole density as a functional of the natural orbitals. Mol Phys 100:401–421
65. Bader RFW, Johnson S, Tang T-H, Popelier PLA (1996) The electron pair. J Phys Chem 100:15398–15415
66. Jakobsen P, Jensen F (2020) Representing exact electron densities by a single Slater determinant in finite basis sets. J Chem Theory Comput 17:269–276
67 Löwdin P-O (1955) Quantum theory of many-particle systems. III. Extension of the Hartree-Fock scheme to include degenerate systems and correlation effects. Phys Rev 97:1509
68. Torres FJ, Rincón L, Zambrano C et al (2019) A review on the information content of the pair density as a tool for the description of the electronic properties in molecular systems. Int J Quantum Chem 119:e25763
69. Becke AD (1988) A multicenter numerical integration scheme for polyatomic molecules. J Chem Phys 88:2547–2553
70. Goede A, Preissner R, Frömmel C (1997) Voronoi cell: new method for allocation of space among atoms: elimination of avoidable errors in calculation of atomic volume and density. J Comput Chem 18:1113–1123
71. Becke AD, Dickson RM (1988) Numerical solution of Poisson's equation in polyatomic molecules. J Chem Phys 89:2993–2997

72. Gharibnejad H, Douguet N, Schneider BI et al (2021) A multi-center quadrature scheme for the molecular continuum. Comput Phys Commun 263:107889
73. Lebedev VI (1975) Values of the nodes and weights of ninth to seventeenth order Gauss-Markov quadrature formulae invariant under the octahedron group with inversion. USSR Comput Math Math Phys 15:44–51
74. Lebedev VI, Finogenov SA (1976) Utilization of ordered Chebyshev parameters in iterative methods. USSR Comput Math Math Phys 16:70–83
75. Forum MP (1994) MPI: a message-passing interface standard. University of Tennessee, USA
76. NVIDIA, Vingelmann P, Fitzek FHP (2020) CUDA, release: 10.2.89. https://developer.nvidia.com/cuda-toolkit. Accessed 13 Jan 2024
77. Frisch MJ, Trucks GW, Schlegel HB et al (2016) Gaussian 16, Revision C.01. Gaussian, Inc., Wallingford CT, 2016
78. Levy M, ZP, (2001) The pair density functional of the kinetic energy and its simple scaling property. J Chem Phys 115:9110–9112
79. Ziesche P (1996) Attempts toward a pair density functional theory. Int J Quantum Chem 60:1361–1374
80. Ziesche P (1994) Pair density functional theory—a generalized density functional theory. Phys Lett A 195:213–220
81. Ayers PW, Levy M (2005) Generalized density-functional theory: conquering the N-representability problem with exact functional for the electron pair density and the second-order reduced density matrix. J Chem Sci 117:507–514. https://doi.org/10.1007/BF02708356
82. Ayers PW, Levy M (2005) Using the Kohn-Sham formalism in pair density-functional theories. Chem Phys Lett 415:211–216
83. Ayers PW (2005) Generalized density functional theories using the *k*-electron densities: development of kinetic energy functionals. J Math Phys 46:062107
84. Higuchi K, Higuchi M (2010) Computational pair density functional theory: a proposal for the kinetic energy functional. Phys Rev B Condens Matter Mater Phys 82:155135. https://doi.org/10.1103/PHYSREVB.82.155135
85. Higuchi K, Higuchi M (2008) Pair density functional theory utilizing the noninteracting reference system: an effective initial theory. Phys Rev B Condens Matter Mater Phys 78:125101. https://doi.org/10.1103/PhysRevB.78.125101
86. Higuchi K, Higuchi M (2007) Pair density-functional theory by means of the correlated wave function. Phys Rev A 75:042510. https://doi.org/10.1103/PHYSREVA.75.042510
87. Nagy Á (2003) Pair density functional theory. In: Gidopoulos NI, Wilson S (eds) The fundamentals of electron density, density matrix and density functional theory in atoms, molecules and the solid state. Springer, Dordrecht, pp 79–87
88. Nagy A (2006) Spherically and system-averaged pair density functional theory. J Chem Phys 125:184104
89. Pistol ME (2009) Adiabatic connections and properties of coupling-integrated exchange–correlation holes and pair densities in density functional theory. Chem Phys Lett 480:136–139