Chapter 14
Effect of Molecular and Electronic Geometries on the Electronic Density in FLO-SIC

Simon Liebing, Kai Trepte, and Sebastian Schwalbe

Abstract Recently, Trepte et al. [J. Chem. Phys., vol. 155, 2021] pointed out the importance of analyzing dipole moments in the Fermi-Löwdin orbital (FLO) self-interaction correction (SIC) for cyclic, planar molecules. In this manuscript, the effect of the molecular and electronic geometries on dipole moments and polarizabilities is discussed for non-cyclic molecules. Computed values are presented for water, formaldehyde, and nitromethane. Continuing the work of Schwalbe et al. [J. Chem. Phys. vol. 153, (2020)], we reconfirm that systematic numerical parameter studies are essential to obtain consistent results in density functional theory (DFT) and SIC. In agreement with Trepte et al. [J. Chem. Phys., vol. 155, 2021], DFT agrees well with experiment for dipole moments, while SIC slightly overestimates them. A Linnett double-quartet electronic geometry is found to be energetically preferred for nitromethane.

14.1 Introduction/Motivation

Electronic structure methods have become more important over recent years [1, 2]. These methods can be used to verify experimental observations [3–7]. However, the role of electronic structure methods has changed significantly over the years, as they allow to determine properties which are not easily accessible by experiments [8–11]. Screening for novel materials utilizing purely theoretical and/or computational frameworks saves time, work and money [12–16]. The leading methodology is Kohn-Sham (KS) density functional theory (DFT) [17], based on its suitable accuracy and reasonable numerical effort. Machine learning (ML) strategies are used to speed-

S. Liebing (✉)
JINR Dubna, Bogoliubov Laboratory of Theoretical Physics, 141980 Dubna, Russia
e-mail: science@liebing.cc

S. Liebing · S. Schwalbe
Institute of Theoretical Physics, TU Bergakademie Freiberg, 09599 Freiberg, Germany

K. Trepte
SUNCAT Center for Interface Science and Catalysis, Stanford University, Menlo Park, CA 94025, USA

D. Blaschke et al. (eds.), *Optics and Its Applications*, Springer Proceedings in Physics 281, https://doi.org/10.1007/978-3-031-11287-4_14

up DFT [18] even more or to find novel density functional approximations (DFAs) [19, 20]. The accuracy of novel DFAs [20–22] is getting closer to chemical accuracy.

Some remaining limitations of DFT can be attributed to the so-called self-interaction error (SIE), describing artificial interactions of electrons. The Perdew-Zunger self-interaction correction (PZ-SIC) [23] approximately removes the one-electron SIE. It has a long history of successes and failures [24]. In PZ-SIC, the choice of orbitals is important. Lehtola et al. [25] showed that PZ-SIC suffers from the local minima problem. A recent formulation of PZ-SIC utilizes so-called Fermi-Löwdin orbitals (FLO-SIC) [26–30]. FLO-SIC depends on Fermi-orbital descriptors (FODs) [29] to construct the localized orbitals used for PZ-SIC [31]. These FODs can be imagined as semi-classical electron positions which form an electronic geometry. Recently, Trepte et al. [31] showed that one can guide and classify local minima in PZ-SIC with the help of special sets of FODs that reflect chemical bonding theories of Lewis [32] and Linnett [33, 34]. The latter is known as Linnett's double-quartet (LDQ) theory. While typically one is interested in the *variational* total energy of the system, Trepte et al. [31] proposed to additionally monitor the dipole moment to classify PZ-SIC solutions. The dipole moment is one of the most simple descriptor for the electronic density—the key property in any DFA.

In this work we investigate the influence of molecular and electronic geometries as well as a properly chosen parameter space on the quality of density-related properties in DFT and FLO-SIC. We show that numerical parameters need to be optimized not only for the total energy but also for, e.g., the electric dipole moments and/or polarizabilities. We discuss the results based on small, illustrative and educationally-valuable molecules.

The manuscript is structured as follows. In the first two sections we outline the theoretical background and the computational details, after which we present the major results. In the last section we summarize and conclude our findings.

14.2 Theoretical Background

KS-DFT, see Fig. 14.1, is an approximation to solve the Schrödinger equation.

The total energy of a system is expressed as a functional of the electron density

$$E_{\mathrm{KS}}[n^\alpha, n^\beta] = T_{\mathrm{s}}[n^\alpha, n^\beta] + V[n] + J[n] + K_{\mathrm{XC}}[n^\alpha, n^\beta], \tag{14.1}$$

where $T_{\mathrm{s}}[n^\alpha, n^\beta]$ is the kinetic energy of the non-interacting system, $V[n]$ is the external potential energy, $J[n]$ is the Coulomb functional, $K_{\mathrm{XC}}[n^\alpha, n^\beta]$ is the exchange-correlation (XC) functional, n is the electron density, and α and β indicate spin channels.

To compute the XC functional

$$K_{\mathrm{XC}}[n^\alpha, n^\beta] = \int \varepsilon_{\mathrm{XC}}^{\mathrm{hom}}[n^\alpha, n^\beta] n(\boldsymbol{r}) F_{\mathrm{XC}} \, \mathrm{d}\boldsymbol{r}, \tag{14.2}$$

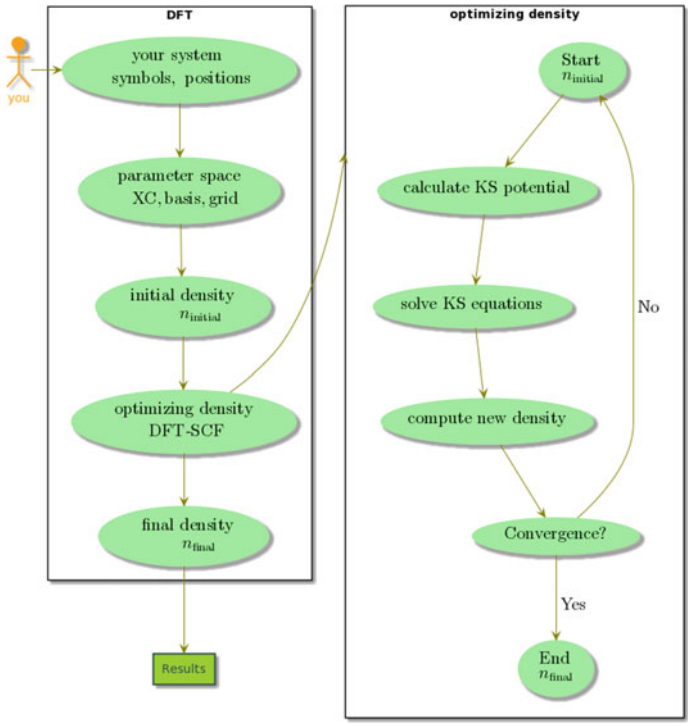


Fig. 14.1 Simplified overview of a DFT calculation

one needs to evaluate an explicit density integral using a numerical quadrature, see Sect. 14.3. Here, $\varepsilon_{\text{XC}}^{\text{hom}}[n^\alpha, n^\beta]$ is the XC energy-density of the homogeneous electron gas and F_{XC} is an XC enhancement factor.

Several approximations exist for the XC enhancement factor, many of which are available in LIBXC [35]. These approximations lead to artificial interactions of electrons with themselves; this is called self-interaction (SI). The corresponding SI energy comes from an incomplete cancellation of the exchange-correlation energy and the Coloumb energy for one-electron densities n_1^σ

$$E_{\text{SI}}[n_1^\sigma] = K_{\text{XC}}[n_1^\sigma, 0] + J[n_1^\sigma]. \tag{14.3}$$

In PZ-SIC, the total E_{KS} is corrected orbital-by-orbital as

$$E_{\text{PZ}} = E_{\text{KS}}[n^\alpha, n^\beta] + E_{\text{SIC}} = E_{\text{KS}}[n^\alpha, n^\beta] - \sum_\sigma \sum_{i=1}^{N^\sigma} E_{\text{SI}}[n_i^\sigma]. \tag{14.4}$$

A novel flavor of PZ-SIC is FLO-SIC. This formulation utilizes FODs to construct Fermi orbitals (FO). These FOs are then orthogonalized to become FLOs. The FODs

need to be optimized in the employed numerical parameter space using the respective analytical gradients [36]. With energy and gradient expressions at hand one can study, e.g., ionization potentials, atomization energies or barrier heights. However, guided by Trepte et al. [31] our focus is not only on energies, i.e., Eqs. (14.1) and (14.4), but on properties characterizing the density. Thus, having introduced energy expressions for DFT and PZ-SIC, we continue to discuss dipole moments and polarizabilities as fingerprints of the electron density.

Density-related properties can be analyzed using small applied electric fields. The total energy of a system under an external electrical field $\boldsymbol{\varepsilon}$ can be written as

$$E(\boldsymbol{\varepsilon}) = E_0 + \sum_i \mu_i \varepsilon_i + \sum_{ij} \alpha_{ij} \varepsilon_i \varepsilon_j + \mathcal{O}(\boldsymbol{\varepsilon}^3). \tag{14.5}$$

Here, E_0 refers to a ground state energy, e.g., KS-DFT E_{KS} (see Eq. (14.1)) or PZ-SIC E_{PZ} (see Eq. (14.4)). From this energy expression we can derive the electric dipole moment as

$$\mu_i = \left(\frac{\partial E(\boldsymbol{\varepsilon})}{\partial \varepsilon_i} \right)_{\boldsymbol{\varepsilon}=0}. \tag{14.6}$$

Commonly, the dipole moment is directly calculated from the electronic density

$$\boldsymbol{\mu} = \sum_A Z_A \boldsymbol{R}_A - \int d\boldsymbol{r}\, n(\boldsymbol{r})\, \boldsymbol{r}, \tag{14.7}$$

where Z_A, $\boldsymbol{R}_A$, and $n(\boldsymbol{r})$ are nuclear charges and positions and the total electron density, respectively. Note, we only discuss the electric dipole moment in this work and refer to it simply as *dipole moment*. The dipole moment is a measure for the polarity of a system and tells us about the charge separation in this system.

As an example, the molecular electrostatic potential [39] (MEP) and the dipole moment of the H_2O molecule is visualized in Fig. 14.2. The dipole points from the more electronegative O atom to the less electronegative H atoms. This can be directly seen in the coloring scheme of the MEP.

Having the possibility to calculate the total energy of the system under the influence of an external electric field allows to study dipole moments using, e.g., a 2-point finite difference (FD) stencil

$$\mu_{\mathrm{FD},i} = \frac{E(+\varepsilon_i) - E(-\varepsilon_i)}{2\varepsilon_i}. \tag{14.8}$$

One can also derive the electric polarizability α_{ij} from Eq. (14.5) as

$$\alpha_{ij} = \left(\frac{\partial E(\boldsymbol{\varepsilon})}{\partial \varepsilon_i \varepsilon_j} \right)_{\boldsymbol{\varepsilon}=0}. \tag{14.9}$$

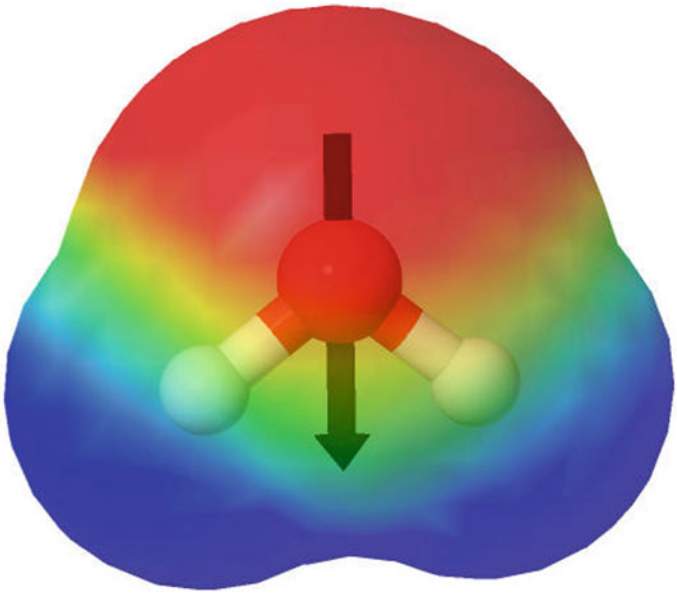


Fig. 14.2 Illustration of the dipole moment $\boldsymbol{\mu}$ in the H_2O molecule using the geometry [37] from CCCBDB [38]. The molecular electrostatic potential (MEP) [39] was calculated using DFT with LDA-PW, aug-pc-3, and a grid=(200,1454) in PySCF. The visualization was done using *Jmol* [40]

The electric polarizability can be calculated using analytical approaches, e.g., solving the coupled perturbed Hartree-Fock (CPHF) equation [41–43]. It describes the tendency of a system to acquire an induced dipole moment in the presence of an external electric field. Similar to the dipole moment we will refer to the electric polarizability as *polarizability*.

With the dipole moment at hand, i.e., Eqs. (14.7) or (14.8), one can calculate the directional components of the polarizability tensor as vector components

$$[\alpha_{\mathrm{FD},ix}, \alpha_{\mathrm{FD},iy}, \alpha_{\mathrm{FD},iz}] = \frac{\boldsymbol{\mu}(+\varepsilon_i) - \boldsymbol{\mu}(-\varepsilon_i)}{2\varepsilon_i}, \tag{14.10}$$

which is a row in Eq. (14.11).

To further simplify the characterization of the density, we introduce scalar values. The vectorial dipole moment $\boldsymbol{\mu} = (\mu_x, \mu_y, \mu_z)$ will be represented as $\mu = |\boldsymbol{\mu}|$, while the tensorial polarizability

$$\alpha_{ij} = \begin{bmatrix} \alpha_{xx} & \alpha_{xy} & \alpha_{xz} \\ \alpha_{yx} & \alpha_{yy} & \alpha_{yz} \\ \alpha_{zx} & \alpha_{zy} & \alpha_{zz} \end{bmatrix} \tag{14.11}$$

will be represented as $\alpha = \mathrm{Tr}(\alpha_{ij})/3$. As shown by Trepte et al. [31], dipole moments are sensitive to the chemical bonding described by the electronic geometry. Moving

FODs changes the dipole moment in FLO-SIC. Dipole moments in FLO-SIC provide insights into whether the electronic density respects the molecular symmetry or not [31].

In practical FLO-SIC calculations, the quality of the density is determined by the used basis set, numerical quadrature, molecular and electronic geometry. Computational details and the used electronic structure codes are discussed next.

14.3 Computational Details

All scripts to produce the data presented in the manuscript are available at https://gitlab.com/opensic/dippo [44]. The calculations were performed with the all-electron Gaussian-type orbital (GTO) codes PySCF [42] and PyFLOSIC2. PyFLOSIC2, see https://gitlab.com/opensic/pyflosic2, is the successor of PyFLOSIC [30]. It offers a cleaner and more modular code structure and can now easily be installed via the Python package manager *pip*. For calculations of real- and complex-valued SIC, i.e., RSIC and CSIC [25, 45, 46], we used the ERKALE code [47]. In previous studies [30, 31] we observed that the pc-n basis sets [48, 49] perform well for DFT as well as SIC calculations. Therefore, for all calculations in this work we use pc-n basis set variants. All codes use the Libxc [35] library, offering access to a vast variety of exchange-correlation functionals. From this library we access LDA-PW [50], PBEsol [51], and r^{2}SCAN [22]. The used codes are Open-Source codes, meaning they are freely available to anyone [52, 53]. Open-Source codes enable faster code development, re-usable concepts, and versatile tool-boxes.

PySCF and PyFLOSIC2 are written in Python, where only numerically demanding parts in PySCF are written in C. Python is simple and elegant, has a friendly and helpful community, and provides various well-maintained libraries. These are only some reasons why it is easy for students or non-programmers to start coding with Python. This allows to solve even non-trivial tasks, like writing a DFT code from scratch [54] in the limited time of a master thesis when guided and educated with novel strategies [55].

A numerical quadrature [56] is needed to evaluate XC properties in DFT and SIC, see Eq. (14.2) in Sect. 14.2. We will refer to it simply as *grid*. A typical grid consists of a radial and an angular part. Its size is given as a pair of numbers, i.e., the number of radial shells and the number of angular points. SIC requires significantly finer grids than DFT [45, 46, 57]. In analogy to [30, 31], we prune the used grids neither for DFT nor for SIC. This is done as the orbital densities evaluated in SIC are not as smooth as the total density used in DFT [58], thus requiring a finer resolution [46].

FLO-SIC has two major variational degrees of freedom, the density matrix (DM) and the FODs. All FLO-SIC calculations in this work are realized with a two-step FLO-SIC SCF cycle, which follows the idea proposed by [59]. In FLO-SIC, the initial DM and initial molecular coefficients are typically the ones from a DFT calculation. The initial FODs can be generated with various procedures, e.g., Python-based center of mass PyCOM, Fermi-orbital descriptor Monte-Carlo FODMC, or other so-called

FOD generators [29]. All initial FOD configurations used in this work were generated using the FODMC.

Within a full FLO-SIC calculation, the FODs are fully optimized in an inner FOD loop for a given DM. All SIC properties are calculated for this DM and the respective optimized FODs. Based on this, the unified Hamiltonian [30, 60–62] is updated which then provides the next DM in the outer DM loop. The initial FODs in the inner FOD loop are optimized using the *SciPy* L-BFGS-B [63–70] optimization algorithm with a maximum force component threshold of $f_{\max,\mathrm{tol}} = 2 \cdot 10^{-4}\ E_\mathrm{h}/a_0$. This two-step procedure is repeated until the FOD forces reach $f_{\max,\mathrm{tol}}$ and the DM is not changing anymore.

The computational methods introduced in this section have been applied to calculate dipole moments and polarizabilities. The results for water, formaldehyde, and nitromethane are discussed in the next section.

14.4 Results

14.4.1 Sisyphus Rock: The Importance of Grid and Basis Set Size

For the correct description of density-related properties, it is important to converge the used numerical parameters space consisting of grid and basis set. Systematic parameters studies have been performed utilizing CCCBDB molecular geometries [38], and FODs generated with the FODMC in the case of FLO-SIC to exemplify this. The determined trends and optimal values should be transferable to other molecular geometries and other FOD arrangements.

We carried out systematic grid convergence tests varying the number of radial shells with a fixed number of angular points and vice versa. Detailed information can be found at https://gitlab.com/opensic/dippo [44]. The DFT values converge at smaller grids than the respective SIC values, see Table 14.1. For LDA-PW, a value of $N_\mathrm{rad} = 200$ for the radial shells gives converged results for both DFT and SIC. The angular dependency for DFT as well as for FLO-SIC is converged at $N_\mathrm{ang} = 590$. However, we use a value of $N_\mathrm{ang} = 1454$ to resolve all one-electron and total densities accurately.

Note that we investigated the convergence of the density in DFT using the dipole moment as well as the polarizability. Both density fingerprints deliver the same optimal parameters for the tested molecules. Thus, for FLO-SIC we only used the dipole moment to determine optimal parameters.

Having an optimal grid of (200,1454) enabled us to determine a suitable basis set. The convergence of the basis set for water is shown in Fig. 14.3 using DFT and in Fig. 14.4 using FLO-SIC. The aug-pc-3 basis set shows convergence w.r.t. the basis set size. Those parameters are optimal for water, formaldehyde, and nitromethane. However, such convergence checks need to be done for any molecule—a Sisyphus

Table 14.1 Optimal numerical parameters for DFT and FLO-SIC regarding the molecular geometries reported in CCCBDB [38] for H_2O [37], CH_2O [71] and CH_3NO_2 [72]

System	DFT			FLO-SIC		
	n_{rad}	n_{ang}	basis set	n_{rad}	n_{ang}	basis set
H_2O	100	302	aug-pc-3	150	590	aug-pc-3
CH_2O	100	302	aug-pc-3	200	590	aug-pc-3
CH_3NO_2	150	590	aug-pc-3	200	590	aug-pc-3

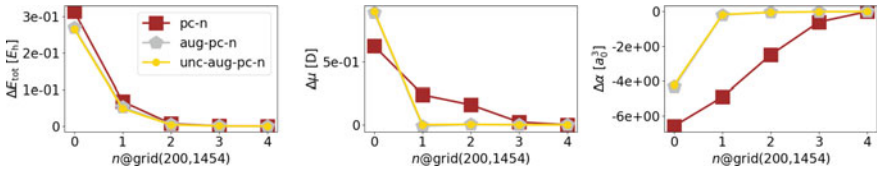


Fig. 14.3 H_2O (DFT, LDA-PW): Convergence of the total energy E_{tot}, the dipole moment μ and the polarizability α w.r.t. increasing basis set size for LDA-PW DFT using PySCF. We used pc-n, aug-pc-n, and unc-aug-pc-n with n=0-4 [48, 49]. Each plot shows the difference to the largest used basis set

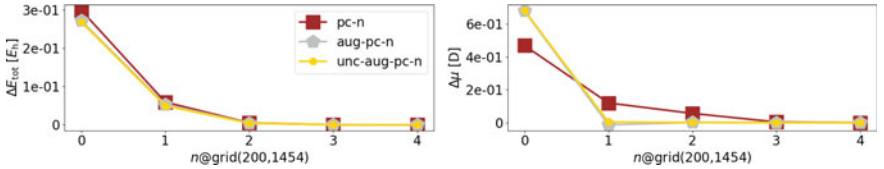


Fig. 14.4 H_2O (FLO-SIC, LDA-PW): Convergence of the total energy E_{tot} and the dipole moment μ w.r.t. increasing basis set size for LDA-PW FLO-SIC using PyFLOSIC2. We used pc-n, aug-pc-n, and unc-aug-pc-n with n=0-4 [48, 49]. Each plot shows the difference to the largest used basis set. Only the density matrix was optimized, while the FODs were fixed

work. Otherwise, the meaning of absolute values for dipole moments or polarizabilities are questionable.

Having established a suitable numerical parameter space, we continue to discuss the influence of molecular geometries on dipole moments and polarizabilities.

14.4.2 Pandora's Box: The Quality of Molecular Geometries Matters

A molecular geometry is needed to perform electronic structure theory calculations; in case of DFT see Fig. 14.1. While such geometries can be optimized within most theories, it is not uncommon to use a fixed molecular geometry to be comparable to other approximations or to simply save computational time.

Table 14.2 Molecular information for H_2O, CH_2O, and CH_3NO_2 with number of atoms, electrons, α and β electrons N_{nuc}, N_{elec}, N_α and N_β. Experimental references for the dipole moments [79] and polarizabilities [38] are provided

System	N_{nuc}	N_{elec}	N_α	N_β	μ_{REF} [D]	α_{REF} [a_0^3]
Water (H_2O)	3	10	5	5	1.85	10.13
Formaldehyde (CH_2O)	4	16	8	8	2.33	18.69
Nitromethane (CH_3NO_2)	7	32	16	16	3.46	32.39

For quantum chemical calculations there exist a vast a variety of seemingly promising databases, such as CCCBDB [38], ChemSpider [73, 74], PubChem [75], and many more. However, the quality of the geometries in these databases can vary [76]. CCCBDB offers access to a vast variety of molecular geometries. In this work, we used the CCCBDB geometries which can be found in the experimental section. Note that those geometries are not necessarily experimental ones. For example, for water [37] the geometry is derived semi-empirically utilizing experimental reference values. PubChem provides molecular geometries calculated using the MMFF94s [77] force field. For other databases such as ChemSpider it is even not that trivial to find the quality of the geometries.

The question of the quality of the molecular geometry might be important for other fields as well. For example, machine learning [78] models might be trained on low quality geometries, which could affect the predictability of the resulting models. For SIC calculations, the quality of the molecular geometry is of great importance, as the orbital densities are sensitive to the underlying molecular geometry. Molecular geometry optimizations are a standard task for commonly used approaches like DFT. However, for more computational demanding methods like FLO-SIC, full geometry optimization require high computational effort. This is caused by the coupled degrees of freedoms of nuclei and FODs [31].

The following results are based on the small, educational systems H_2O, CH_2O, and CH_3NO_2. We summarized some essential molecular information for those molecules in Table 14.2, including experimental reference values for dipole moments and polarizabilities.

The effect of molecular geometries on *density* fingerprints, i.e., dipole moments and polarizabilities, is significant even at the DFT level (see Fig. 14.5). For the used test systems the molecular geometries from CCCBDB provide the best molecular geometries; the dipole moments and polarizabilities are close to the values obtain from a DFT optimized geometry. The basis set size affects the density-related properties significantly, and only the aug-pc-3 basis set provides converged results. The molecular geometries from PubChem and ChemSpider should be used with care, as the dipole moments deviate significantly from values obtained from a DFT optimized geometry.

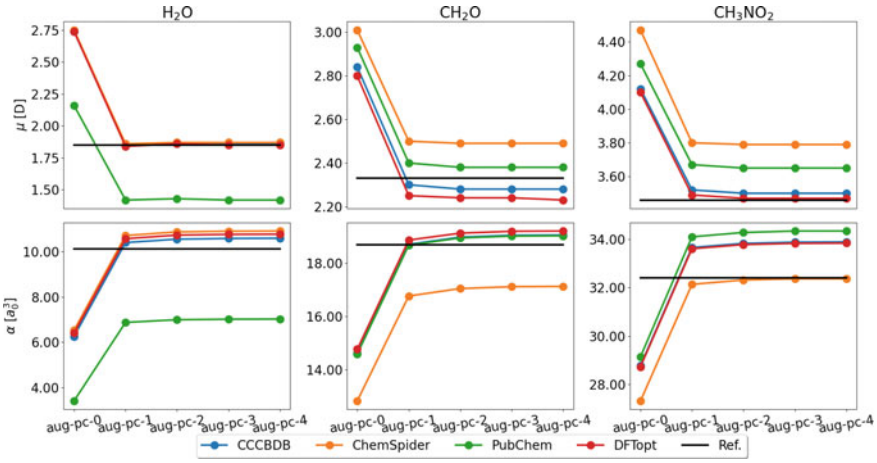


Fig. 14.5 Influence of molecular geometries on density fingerprints, e.g., dipole moments and polarizabilities. For these DFT calculations, LDA-PW was used with a grid=(200,1454) in PySCF. Molecular geometries are taken from common chemical databases. In addition, DFT optimized geometries were used. These DFT geometry optimizations were started from the CCCBDB geometries utilizing ERKALE with aug-pc-3 and grid=(200,1454). The used reference values are provided in Table 14.2

Table 14.3 Finite-difference (FD) errors of polarizabilities in a_0^3 for the chosen step size of $\varepsilon = 10^{-7}$ a.u. utilizing the 2-point FD approximation and LDA-PW DFT. The analytical α as calculated with PySCF were used as reference

Database	aug-pc-0	aug-pc-1	aug-pc-2	aug-pc-3	aug-pc-4
CCCBDB	0.00	0.29	−0.01	−0.01	−0.01
ChemSpider	0.00	−0.08	0.00	−0.02	−0.00
PubChem	−0.01	0.02	0.02	−0.01	0.33
DFTopt	0.00	0.00	−0.03	−0.01	0.01

For all calculations, the finite difference approximations, see Eqs. (14.8) and (14.10), agree well with the analytical results. The mean error for the dipole moments is 0.00 D, while the respective mean errors for the polarizabilities are given in Table 14.3.

Accordingly, for the investigated systems and the employed method the chosen value of 10^{-7} a.u. for the magnitude of $\boldsymbol{\varepsilon}$ regarding the 2-point finite difference approximation delivers reliable numerical results. Note that this finding might not be reproducible for other systems using the same value. Having examined the dependence on the molecular geometry, in Sect. 14.4.3 we proceed to investigate electronic degrees of freedom in FLO-SIC, i.e., DM and FODs.

14.4.3 The Sword of Damocles: Curse and Blessing of Approximations

Approximations are often needed and can be useful to enable the treatment or computation of a specific property at a certain level of theory. However, each approximation needs to be carefully investigated regarding the limits of its predictive power.

Effect of initial FODs Continuing the work of Trepte et al. [31], we show for CH_3NO_2 that it is possible to find several FOD configurations following chemical bonding theories. One possible FOD configuration can be based on Lewis theory of bonding. There, the FODs form one double N=O and one single N-O bond in the -NO_2 group, see Fig. 14.6. Clearly, there exist two identical Lewis configuration where the double and single bonds are exchanged with each other. These Lewis configuration have N-O bond orders of 2 and 1, respectively. Regarding LDQ theory, 2 FODs of one spin channel and 1 FOD of the other spin channel are placed between the N and the O. This leads to bond orders of 1.5 in both N-O bonds. Besides FOD configuration which follow chemical bond theories, other FOD configuration are possible. For example, one can generate a configuration with an over-binding N atom, placing two N=O double bonds in the molecule. We denote this FOD configuration as *other*. Changing the N-O bond order affects the local chemical environment, and with that the resulting SIC solution.

In electronic structure theories there exist several possibilities to treat the spin of the system. In FLO-SIC, we can do restricted calculations where all electrons are paired, $N_\alpha = N_\beta$, and unrestricted calculations where N_α and N_β can vary. In restricted FLO-SIC only one set of FODs is needed, while in unrestricted FLO-SIC two sets of FODs are required.

Given that the calculations for CH_3NO_2 are more computational demanding, we use a grid=(200,1454) and the aug-pc-2 basis set. As seen in Sect. 14.4.1, using this basis set comes with errors in the order of mE_h and mD w.r.t. to the basis set limit. However, the energy difference of the considered FOD configurations, see Fig. 14.6, are in the order of 10^{-2} E_h. The differences in the dipole are in the order of 10^{-1} D. Thus, using the aug-pc-2 basis set should deliver reliable trends.

We performed restricted FLO-SIC calculations for the Lewis configuration of CH_3NO_2. The FODs converge to an electronic geometry which does not follow any bonding theory. The double bond FODs are not lying on the N-O bond axis. Instead, they are distorted towards the respective O atoms. Given this non-symmetric arrangement of FODs, the density of CH_3NO_2 becomes non-symmetric. This leads to an energy of $E_\mathrm{final} = -245.80939$ E_h and a non-symmetric dipole moment with an absolute value of $\mu = 4.06$ D.

The effect of various unrestricted FOD configurations for CH_3NO_2 is shown in Fig. 14.6. The differences in these configurations can already be seen for the initial FODs. Only optimizing the density shows significantly different dipole moments, and only the LDQ value is close to the experimental value. The energy for the initial LDQ arrangement is also the lowest. Upon full optimization of the density

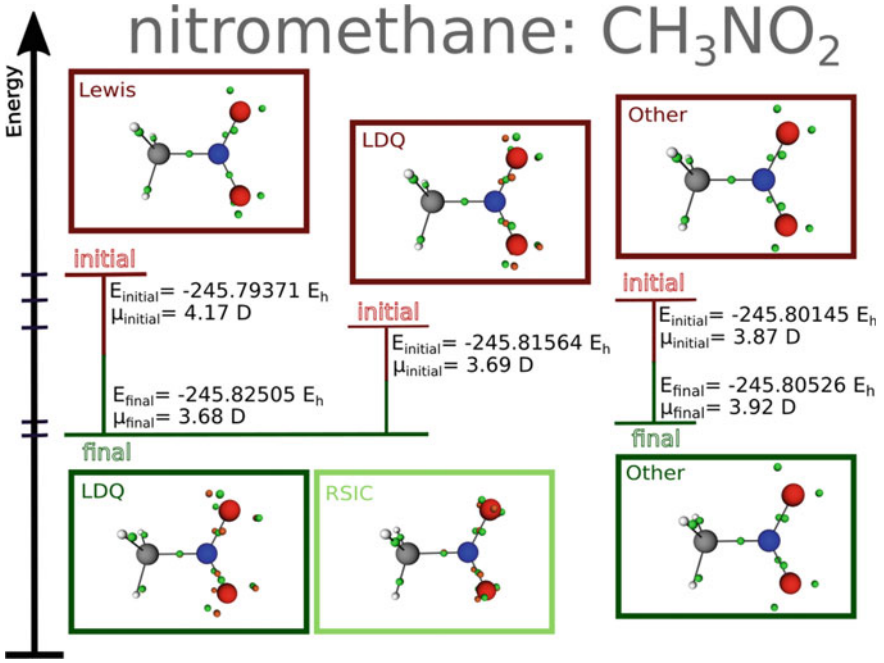


Fig. 14.6 Total energies and dipole moments for various unrestricted FOD configurations for CH_3NO_2, evaluated at the initial and the final FODs. *Other* represents a structure with two N=O double bonds. Any additionally tested structure converged into the LDQ solution. Calculations were performed using PyFLOSIC2 employing LDA-PW, the aug-pc-2 basis set and a grid=(200,1454). In addition, the RSIC center of mass (COM) obtained from ERKALE are shown in the light green box as comparison to the final LDQ FODs. Color code: C-gray, H-white, O-red, N-blue, spin-up FOD/COM - green, spin-down FOD/COM - red. There are situations where spin-up and spin-down FODs/COMs have the same position. Then, only one color is seen

matrix as well as the FODs, the configuration based on LDQ deliver the lowest energy. In the unrestricted calculations, both the initial LDQ and the initial Lewis configuration converge to a final FOD arrangement which can be characterized via LDQ. Interestingly, the *other* FODs stay in the their configuration, but the final energy is higher than for the LDQ FODs.

LDQ not only gives the lowest energy but also the best SIC dipole moment for CH_3NO_2. As a highlight, the center of mass (COM) for the optimal RSIC localized orbitals using ERKALE also reflect the LDQ chemical bonding motif, see the RSIC box in Fig. 14.6. In this section, we reconfirm that the choice of the initial FODs significantly influences dipole moments in FLO-SIC. Next, we describe how the FOD optimization itself can influence dipole moments.

Effect of FOD optimization We continue the discussion for H_2O and CH_2O. These two molecules only have one meaningful FOD configuration, see Fig. 14.7. Having one FOD configuration may allow finding the same local minima in a reproducible

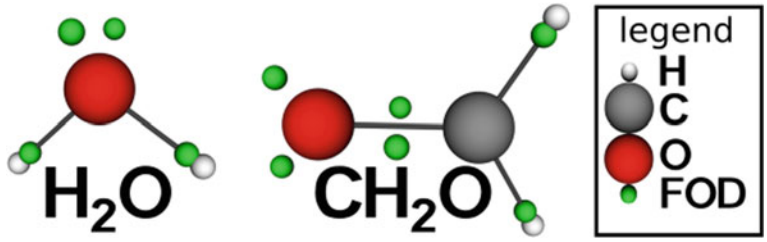
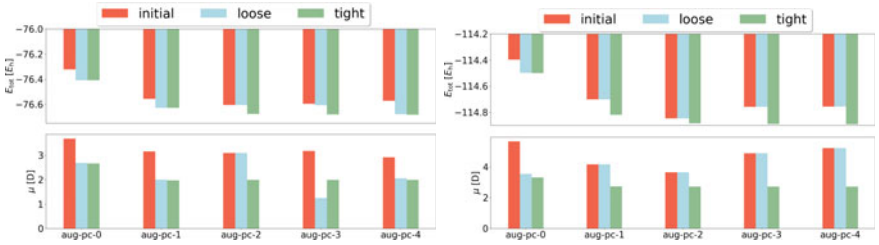


Fig. 14.7 Displaying both the molecular and electronic geometries for H_2O and CH_2O. The picture is generated with the PyFLOSIC2 graphical user interface (GUI). Note, only H_2O and CH_2O are shown as they have only one trivial initial FOD configuration, whereas CH_3NO_2 is more complex; see Fig. 14.6



(a) H_2O (b) CH_2O

Fig. 14.8 Effect of FOD optimization vs. basis set size using the LDA-PW XC functional for (**a**) H_2O and (**b**) CH_2O. The tag *initial* represents FODs generated by the FODMC, *loose* represents FODs with a maximal force criterion of $f_{\max,\mathrm{tol}} = 5 \cdot 10^{-3}\ E_\mathrm{h}/a_0$, and *tight* refers to optimized FODs with $f_{\max,\mathrm{tol}} = 2 \cdot 10^{-4}\ E_\mathrm{h}/a_0$. A grid=(200,1454) was used. Note that when *initial* and *loose* give the same result, the initialized FODs are already at a force threshold of $5 \cdot 10^{-3}\ E_\mathrm{h}/a_0$

fashion. Thus, those molecules are promising candidates for systematic FOD convergence studies.

In FLO-SIC it is rather easy to make various approximations. A common approximation is to use fixed FODs which are not optimized for the numerical parameter space of the respective calculation. Optimizing the FODs is, however, critical to obtain reasonable FLO-SIC solutions. Such optimizations are carried out until a specific threshold for the maximum FOD force, $f_{\max,\mathrm{tol}}$, is reached. The influence of the FOD optimization on the total energy and the dipole moment for water and formaldehyde is shown in Fig. 14.8.

FLO-SIC values for the total energy and the dipole moment change drastically when going from an initial set of FODs to optimized FODs. Here, optimized FODs are characterized by $f_{\max,\mathrm{tol}} = 2 \cdot 10^{-4}\ E_\mathrm{h}/a_0$. Stopping the optimization too early, i.e., at $f_{\max,\mathrm{tol}} = 5 \cdot 10^{-3}\ E_\mathrm{h}/a_0$ can lead to insufficiently converged densities and energies. This furthermore leads to an incorrect prediction of trends; increasing the basis set size should smoothly converge the dipole moments, see Fig. 14.5. However,

Table 14.4 Dipole moments for DFT and FLO-SIC using three XC functionals. Using the aug-pc-3 basis set and a grid=(200,1454) in PySCF for DFT and PyFLOSIC2 for FLO-SIC

XC functional	H_2O		CH_2O	
	μ_{DFT}	$\mu_{FLO\text{-}SIC}$	μ_{DFT}	$\mu_{FLO\text{-}SIC}$
LDA-PW	1.86	1.99	2.28	2.72
PBEsol	1.82	1.95	2.23	2.65
r^2SCAN	1.83	1.96	2.31	2.62

with insufficiently optimized FODs these trends can be predicted incorrectly, as can be seen in Fig. 14.8.

Accordingly, global statements and generalization of trends are only valid for optimized FODs in combination with a sufficient basis set and grid. In the tested cases, the aug-pc-3 basis set with a grid=(200,1454) deliver converged results, in analogy to Sects. 14.4.1 and 14.4.2.

As proposed by the authors [31], monitoring the dipole moment is important to classify and analyze SIC solutions. We showed here that converging the total energy is necessary, but might not be sufficient when one aims to study density-related properties, i.e., dipole moments or polarizabilities. All calculations in the previous sections utilized the LDA-PW functional. Next, we discuss how changing the exchange-correlation functional influences dipole moments.

Effect of exchange-correlation functional For H_2O and CH_2O, we find that already LDA-PW describes dipole moments qualitatively correctly for DFT and SIC. However, comparing three exchange-correlation functionals the absolute dipole moments differ, see Table 14.4. Note that for CH_3NO_2 this comparison was not carried, see *Effect of initial FODs* for more information.

All three pure DFAs agree quite well with the experimental dipole reference values, see Table 14.2. FLO-SIC tends to overshot the dipole moment in order of 10^{-1} D w.r.t. the DFT values. This trend has also been observed in the literature, see [31]. For DFT, LDA-PW performs best for H_2O while r^2SCAN agrees the most for CH_2O. In case of FLO-SIC, the best dipoles are given by PBEsol for H_2O and by r^2SCAN for CH_2O.

In the previous sections we have discussed results for FLO-SIC. To investigate the influence of a specific flavor of SIC on dipole moments, we compare FLO-SIC, RSIC, and CSIC in the next section.

Effect of SIC methods To verify the used numerical parameter space, i.e., the aug-pc-3 basis set and a grid=(200,1454), we carried out RSIC and CSIC calculations using ERKALE. There is no significant difference between FLO-SIC and RSIC values, see Table 14.5. This is noteworthy as the results are calculated with two independent electronic structure codes. Thus, the used numerical parameter space is sufficient to deliver reproducible results. SIC suffers from the multiple local-

Table 14.5 Comparison of μ in D using the aug-pc-3 basis set, a grid=(200,1454) and the LDA-PW functional. Molecular structures from the CCCBDB database [38] were used, coming from the following sources: H_2O [37], CH_2O [71] and CH_3NO_2 [72]. DFT values are obtained from PySCF, FLO-SIC values are obtained from PyFLOSIC2, and RSIC and CSIC values come from ERKALE

Molecule	μ_{DFT}	$\mu_{\text{FLO-SIC}}$	μ_{RSIC}	μ_{CSIC}
H_2O	1.86	1.99	1.99	2.06
CH_2O	2.28	2.72	2.72	2.70
CH_3NO_2	3.50	3.67	3.70	3.90

minima problem [25], see Sect. 14.1. Accordingly, we recommend to verify FLO-SIC results with independent SIC methods such as RSIC. For example for nitromethane, see *Effect of initial FODs*, only the LDQ FLO-SIC dipole agrees with the RSIC solution. FLO-SIC, RSIC and CSIC together deliver a consistent SIC description. The advantage of FLO-SIC is the access to bonding information. This information allows to easily classify and further analyze PZ-SIC solutions [31].

In the next section we summarize and conclude our findings.

14.5 Summary and Conclusion

As shown in Sect. 14.4.1, and already stated in earlier works [30, 31, 45], SIC needs finer numerical quadrature meshes in comparison to DFT calculations. Global statements about the predictive power of SIC are only meaningful using very accurate numerical parameter spaces.

In this work we show how density-related properties, i.e., dipole moments and polarizabilities, can help to determine an appropriate numerical parameter space for DFT and SIC. For our investigated molecules, the aug-pc-3 basis set and a numerical grid of (200,1454) is such an appropriate parameter space. Note, even trends can clearly change using other numerical parameters. While it is mandatory to converge the energy it is not necessarily sufficient for the study of density-related properties.

Furthermore, using water, formaldehyde, and nitromethane we show that these density-related properties are not only sensitive to the used numerical parameter space. They are also significantly influenced by the used molecular geometry, see Sect. 14.4.2. Molecular geometries from common chemical database, i.e., CCCBDB, PubChem, or ChemSpider, can deliver very different dipole moments and polarizabilities. Only molecular geometries from the CCCBDB database deliver reasonable trends in comparison to optimized geometries.

Continuing the work of [31], for FLO-SIC we showed that density-fingerprints are also sensitive to the chemical bonding situation introduced by FODs. We demonstrate that the numerical quality of the FODs, represented by their gradients, as well as the choice of initial FODs clearly influences the dipole moments. For molecules with non-trivial bonding situations, e.g., nitromethane, it is highly recommended to

use various initial FOD configurations. This allows to reasonably sample the FOD configuration space to determine the most reasonable FOD configuration. Such FOD configurations typically follow chemical bonding theories [31], i.e., Lewis and LDQ. In case where several FOD configurations are possible, FODs based on LDQ are often superior in FLO-SIC. This has been shown in [31] and verified here again for nitromethane in Sect. 14.4.3.

When computational time matters, the simplest density functional approximation, i.e., a local density approximation like LDA-PW, provides reasonable trends for dipole moments in FLO-SIC, as seen in Sect. 14.4.3. If computational time is not limited we recommend to use higher rung functional like PBEsol or r^{2}SCAN to verify and further analyze determined trends.

We want to emphasis the importance of Open-Science and Open-Source developments, as this work would not be possible without them.

Acknowledgements S. Schwalbe has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project ID 421663657—KO 1924/9-2. S. Liebing wants to express his gratitude for Prof. Jens Kortus enabling the authors to work on this topic. All authors thank M. Sc. Wanja T. Schulze for his contributions to the PyFLOSIC2 code. We thank both referees, as their comments significantly improved the content and shape of our manuscript. This work is part of our OpenSIC project, and we thank all members for fruitful discussions. We thank the ZIH Dresden for computational time and support.

References

1. Becke, A.D.: Perspective: Fifty years of density-functional theory in chemical physics. J. Chem. Phys. **140**(18), 18A301 (2014)
2. Verma, P., Truhlar, D.G.: Status and challenges of density functional theory. Trends Chem. **2**(4), 302–318 (2020)
3. Förster, S., Hahn, T., Loose, C., Röder, C., Liebing, S., Seichter, W., Eißmann, F., Kortus, J., Weber, E.: Synthesis and characterization of new derivatives of azulene, including experimental and theoretical studies of electronic and spectroscopic behavior. J. Phys. Org. Chem. **25**(10), 856 (2012)
4. Pfaff, U., Hildebrandt, A., Schaarschmidt, D., Hahn, T., Liebing, S., Kortus, J., Lang, H.: Di- and triferrocenyl (hetero) aromatics: Synthesis, characterization,(spectro-) electrochemistry, and calculations. Organometallics **31**(19), 6761 (2012)
5. Seidel, N., Hahn, T., Liebing, S., Seichter, W., Kortus, J., Weber, E.: Synthesis and properties of new 9, 10-anthraquinone derived compounds for molecular electronics. New J. Chem. **37**(3), 601 (2013)
6. Trepte, K., Schaber, J., Schwalbe, S., Drache, F., Senkovska, I., Kaskel, S., Kortus, J., Brunner, E., Seifert, G.: The origin of the measured chemical shift of ^{129}Xe in UiO-66 and UiO-67 revealed by DFT investigations. Phys. Chem. Chem. Phys. **19**, 10020–10027 (2017)
7. Trepte, K., Schwalbe, S., Schaber, J., Krause, S., Senkovska, I., Kaskel, S., Brunner, E., Kortus, J., Seifert, G.: Theoretical and experimental investigations of ^{129}Xe NMR chemical shift isotherms in metal-organic frameworks. Phys. Chem. Chem. Phys. **20**, 25039–25043 (2018)
8. Trepte, K., Schwalbe, S., Seifert, G.: Electronic and magnetic properties of DUT-8(Ni). Phys. Chem. Chem. Phys. **17**, 17122–17129 (2015)

9. Rühlig, K., Abylaikhan, A., Aliabadi, A., Kataev, V., Liebing, S., Schwalbe, S., Trepte, K., Ludt, C., Kortus, J., Büchner, B., et al.: Ni^{II} formate complexes with bi-and tridentate nitrogen-donor ligands: synthesis, characterization, and magnetic and thermal properties. Dalton Trans. **46**(12), 3963–3979 (2017)
10. Taubert, F., Schwalbe, S., Seidel, J., Hüttl, R., Gruber, T., Janot, R., Bobnar, M., Gumeniuk, R., Mertens, F., Kortus, J.: Thermodynamic characterization of lithium monosilicide (LiSi) by means of calorimetry and DFT-calculations. Int. J. Mater. Res. **108**(11), 942–958 (2017)
11. Schwalbe, S., Gruber, T., Trepte, K., Taubert, F., Mertens, F., Kortus, J.: Mechanical, elastic and thermodynamic properties of crystalline lithium silicides. Comput. Mater. Sci. **134**, 48–57 (2017)
12. Schwalbe, S., Trepte, K., Seifert, G., Kortus, J.: Screening for high-spin metal organic frameworks (MOFs): density functional theory study on DUT-8(M_1, M_2) (with M_i = V,..., Cu). Phys. Chem. Chem. Phys. **18**, 8075–8080 (2016)
13. Friedrich, R., Usanmaz, D., Oses, C., Supka, A., Fornari, M., Nardelli, M.B., Toher, C., Curtarolo, S.: Coordination corrected ab initio formation enthalpies. NPJ Comput. Mater. **5**(1), 1–12 (2019)
14. Mehl, M.J., Ronquillo, M., Hicks, D., Esters, M., Oses, C., Friedrich, R., Smolyanyuk, A., Gossett, E., Finkenstadt, D., Curtarolo, S.: Tin-pest problem as a test of density functionals using high-throughput calculations. Phys. Rev. Mater. **5**(8), 083608 (2021)
15. Trepte, K., Schwalbe, S.: porE: A code for deterministic and systematic analyses of porosities. J. Comput. Chem. **42**(9), 630–643 (2021)
16. Friedrich, R., Ghorbani-Asl, M., Curtarolo, S., Krasheninnikov, A.V.: Data-driven quest for two-dimensional non-van der Waals materials. Nano Lett. **22**, 989 (2022)
17. Kohn, W., Sham, L.J.: Self-consistent equations including exchange and correlation effects. Phys. Rev. **140**(4A), A1133 (1965)
18. Ellis, J.A., Fiedler, L., Popoola, G.A., Modine, N.A., Stephens, J.A., Thompson, A.P., Cangi, A., Rajamanickam, S.: Accelerating finite-temperature Kohn-Sham density functional theory with deep neural networks. Phys. Rev. B **104**(3), 035120 (2021)
19. Brown, K., Maimaiti, Y., Trepte, K., Bligaard, T., Voss, J.: MCML: Combining physical constraints with experimental data for a multi-purpose meta-generalized gradient approximation. J. Comput. Chem. **42**, 2004–2013 (2021)
20. Kirkpatrick, J., McMorrow, B., Turban, D.H., Gaunt, A.L., Spencer, J.S., Matthews, A.G., Obika, A., Thiry, L., Fortunato, M., Pfau, D., et al.: Pushing the frontiers of density functionals by solving the fractional electron problem. Science **374**(6573), 1385–1389 (2021)
21. Sun, J., Ruzsinszky, A., Perdew, J.P.: Strongly constrained and appropriately normed semilocal density functional. Phys. Rev. Lett. **115**(3), 036402 (2015)
22. Furness, J.W., Kaplan, A.D., Ning, J., Perdew, J.P., Sun, J.: Accurate and numerically efficient r^2SCAN meta-generalized gradient approximation. J. Phys. Chem. Lett. **11**, 8208–8215 (2020)
23. Perdew, J.P., Zunger, A.: Self-interaction correction to density-functional approximations for many-electron systems. Phys. Rev. B **23**(10), 5048 (1981)
24. Perdew, J.P., Ruzsinszky, A., Sun, J., Pederson, M.R.: Paradox of self-interaction correction: How can anything so right be so wrong? Adv. At. Mol. Opt. Phys. **64**, 1–14 (2015)
25. Lehtola, S., Head-Gordon, M., Jónsson, H.: Complex orbitals, multiple local minima, and symmetry breaking in Perdew-Zunger self-interaction corrected density functional theory calculations. J. Chem. Theory Comput. **12**(7), 3195–3207 (2016)
26. Pederson, M.R., Ruzsinszky, A., Perdew, J.P.: Communication: Self-interaction correction with unitary invariance in density functional theory. J. Chem. Phys. **140**(12), 121103 (2014)
27. Pederson, M.R., Baruah, T.: Self-interaction corrections within the Fermi-orbital-based formalism. Adv. At. Mol. Opt. Phys. **64**, 153 (2015)
28. Yang, Z.-H., Pederson, M.R., Perdew, J.P.: Full self-consistency in the Fermi-orbital self-interaction correction. Phys. Rev. A **95**, 052505 (2017)
29. Schwalbe, S., Trepte, K., Fiedler, L., Johnson, A.I., Kraus, J., Hahn, T., Peralta, J.E., Jackson, K.A., Kortus, J.: Interpretation and automatic generation of Fermi-orbital descriptors. J. Comput. Chem. **40**(32), 2843–2857 (2019)

30. Schwalbe, S., Fiedler, L., Kraus, J., Kortus, J., Trepte, K., Lehtola, S.: PyFLOSIC: Python-based Fermi-Löwdin orbital self-interaction correction. J. Chem. Phys. **153**(8), 084104 (2020)
31. Trepte, K., Schwalbe, S., Liebing, S., Schulze, W.T., Kortus, J., Myneni, H., Ivanov, A.V., Lehtola, S.: Chemical bonding theories as guides for self-interaction corrected solutions: multiple local minima and symmetry breaking. J. Chem. Phys. **155**(22), 224109 (2021)
32. Lewis, G.N.: The atom and the molecule. J. Am. Chem. Soc. **38**(4), 762–785 (1916)
33. Linnett, J.W.: Valence-bond structures: A new proposal. Nature **187**(4740), 859 (1960)
34. Linnett, J.W.: A modification of the Lewis-Langmuir octet rule. J. Am. Chem. Soc. **83**(12), 2643–2653 (1961)
35. Lehtola, S., Steigemann, C., Oliveira, M.J.T., Marques, M.A.L.: Recent developments in LIBXC–A comprehensive library of functionals for density functional theory. SoftwareX **7**, 1 (2018)
36. Pederson, M.R.: Fermi orbital derivatives in self-interaction corrected density functional theory: applications to closed shell atoms. J. Chem. Phys. **142**(6), 064112 (2015)
37. Hoy, A.R., Bunker, P.R.: A precise solution of the rotation bending Schrödinger equation for a triatomic molecule with application to the water molecule. J. Mol. Spectrosc. **74**(1), 1–8 (1979)
38. Johnson, R.D.: NIST Computational Chemistry Comparison and Benchmark Database. NIST Standard Reference Database Number 101 (2020)
39. Scrocco, E., Tomasi, J.: Electronic molecular structure, reactivity and intermolecular forces: an euristic interpretation by means of electrostatic molecular potentials. Adv. Quantum Chem. **11**, 115–193 (Elsevier, 1978)
40. Jmol: an open-source java viewer for chemical structures in 3d. http://www.jmol.org/ (2022)
41. Handy, N.C., Schaefer, H.F., III.: On the evaluation of analytic energy derivatives for correlated wave functions. J. Chem. Phys. **81**(11), 5031–5033 (1984)
42. Sun, Q., Zhang, X., Banerjee, S., Bao, P., Barbry, M., Blunt, N.S., Bogdanov, N.A., Booth, G.H., Chen, J., Cui, Z.-H., Eriksen, J.J., Gao, Y., Guo, S., Hermann, J., Hermes, M.R., Koh, K., Koval, P., Lehtola, S., Li, Z., Liu, J., Mardirossian, N., McClain, J.D., Motta, M., Mussard, B., Pham, H.Q., Pulkin, A., Purwanto, W., Robinson, P.J., Ronca, E., Sayfutyarova, E.R., Scheurer, M., Schurkus, H.F., Smith, J.E.T., Sun, C., Sun, S.-N., Upadhyay, S., Wagner, L.K., Wang, X., White, A., Whitfield, J.D., Williamson, M.J., Wouters, S., Yang, J., Yu, J.M., Zhu, T., Berkelbach, T.C., Sharma, S., Sokolov, A.Y., Chan, G.K.-L.: Recent developments in the PySCF program package. J. Chem. Phys. **153**(2), 024109 (2020)
43. Smith, J.E.T.: Modern Multireference Electronic Structure Theory. PhD thesis, https://scholar.colorado.edu/concern/graduate_thesis_or_dissertations/9306t035k. Accessed Feb 2022 (2020)
44. Liebing, S., Trepte, K., Schwalbe, S.: Dippo-(dip)oles and (po)larizabilities. https://doi.org/10.5281/zenodo.6246152 (2022)
45. Lehtola, S., Jónsson, H.: Variational, self-consistent implementation of the Perdew-Zunger self-interaction correction with complex optimal orbitals. J. Chem. Theory Comput. **10**(12), 5324 (2014)
46. Lehtola, S., Jónsson, E.Ö., Jónsson, H.: Effect of complex-valued optimal orbitals on atomization energies with the Perdew-Zunger self-interaction correction to density functional theory. J. Chem. Theory Comput. **12**, 4296–4302 (2016)
47. Lehtola, J., Hakala, M., Sakko, A., Hämäläinen, K.: ERKALE–A flexible program package for X-ray properties of atoms and molecules. J. Comput. Chem. **33**(18), 1572 (2012)
48. Jensen, F.: Polarization consistent basis sets: principles. J. Chem. Phys. **115**, 9113 (2001)
49. Jensen, F.: Polarization consistent basis sets II: estimating the Kohn-Sham basis set limit. J. Chem. Phys. **116**(17), 7372 (2002)
50. Perdew, J.P., Wang, Y.: Accurate and simple analytic representation of the electron-gas correlation energy. Phys. Rev. B **45**, 13244–13249 (1992)
51. Perdew, J.P., Ruzsinszky, A., Csonka, G.I., Vydrov, O.A., Scuseria, G.E., Constantin, L.A., Zhou, X., Burke, K.: Restoring the density-gradient expansion for exchange in solids and surfaces. Phys. Rev. Lett. **100**, 136406 (2008)

52. Oliveira, M.J.T., Papior, N., Pouillon, Y., Blum, V., Artacho, E., Caliste, D., Corsetti, F., de Gironcoli, S., Elena, A.M., García, A., García-Suárez, V.M., Genovese, L., Huhn, W.P., Huhs, G., Kokott, S., Küçükbenli, E., Larsen, A.H., Lazzaro, A., Lebedeva, I.V., Li, Y., López-Durán, D., López-Tarifa, P., Lüders, M., Marques, M.A.L., Minar, J., Mohr, S., Mostofi, A.A., O'Cais, A., Payne, M.C., Ruh, T., Smith, D.G.A., Soler, J.M., Strubbe, D.A., Tancogne-Dejean, N., Tildesley, D., Torrent, M., Yu, V.W.: The CECAM electronic structure library and the modular software development paradigm. J. Chem. Phys. **153**(2), 024117 (2020)
53. Lehtola, S., Karttunen, A.: Free and open source software for computational chemistry education. ChemRxiv, https://doi.org/10.33774/chemrxiv-2021-hr1r0-v2 (2021)
54. Schulze, W.: Domain-averaged fermi holes: a self-interaction correction perspective. Master Thesis, https://doi.org/10.13140/RG.2.2.27958.42568/2 (2021)
55. Ismail-Beigi, S., Arias, T.: New algebraic formulation of density functional calculation. Comput. Phys. Commun. **128**(1–2), 1–45 (2000)
56. Lebedev, V.I., Laikov, D.N.: A quadrature formula for the sphere of the 131st algebraic order of accuracy. Dokl. Math. **59**(3), 477–481 (1999)
57. Vydrov, O.A., Scuseria, G.E.: Effect of the Perdew-Zunger self-interaction correction on the thermochemical performance of approximate density functionals. J. Chem. Phys. **121**(17), 8187–8193 (2004)
58. Shahi, C., Bhattarai, P., Wagle, K., Santra, B., Schwalbe, S., Hahn, T., Kortus, J., Jackson, K.A., Peralta, J.E., Trepte, K., Lehtola, S., Nepal, N.K., Myneni, H., Neupane, B., Adhikari, S., Ruzsinszky, A., Yamamoto, Y., Baruah, T., Zope, R.R., Perdew, J.P.: Stretched or noded orbital densities and self-interaction correction in density functional theory. J. Chem. Phys. **150**(17), 174102 (2019)
59. Karanovich, A., Yamamoto, Y., Jackson, K.A., Park, K.: Electronic structure of mononuclear Cu-based molecule from density-functional theory with self-interaction correction. J. Chem. Phys. **155**(1), 014106 (2021)
60. Harrison, J.G., Heaton, R., Lin, C.C.: Self-interaction correction to the local density Hartree-Fock atomic calculations of excited and ground states. J. Phys. B-At. Mol. **16**(12), 2079 (1983)
61. Pederson, M.R., Heaton, R.A., Lin, C.C.: Local-density Hartree-Fock theory of electronic states of molecules with self-interaction correction. J. Chem. Phys. **80**, 1972 (1984)
62. Lehtola, S., Jónsson, H.: Unitary optimization of localized molecular orbitals. J. Chem. Theory Comput. **9**(12), 5365–5372 (2013)
63. Broyden, C.G.: The convergence of a class of double-rank minimization algorithms 1 general considerations. IMA J. Appl. Math. **6**(1), 76 (1970)
64. Fletcher, R.: A new approach to variable metric algorithms. Comput. J. **13**(3), 317 (1970)
65. Goldfarb, D.: A family of variable-metric methods derived by variational means. Math. Comput. **24**(109), 23 (1970)
66. Shanno, D.F.: Conditioning of quasi-Newton methods for function minimization. Math. Comput. **24**(111), 647 (1970)
67. Nocedal, J.: Updating quasi-Newton matrices with limited storage. Math. Comput. **35**(151), 773 (1980)
68. Liu, D.C., Nocedal, J.: On the limited memory BFGS method for large scale optimization. Math. Program. **45**(1–3), 503 (1989)
69. Byrd, R.H., Lu, P., Nocedal, J., Zhu, C.: A limited memory algorithm for bound constrained optimization. SIAM J. Sci. Comput. **16**(5), 1190 (1995)
70. Zhu, C., Byrd, R.H., Lu, P., Nocedal, J.: Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. ACM Trans. Math. Softw. **23**(4), 550 (1997)
71. Gurvich, L.V., Veits, I.V., Alcock, C.B.: Thermodynamics Properties of Individual Substances, vol 1 and 116. Hemisphere Pub. Co. (1989)
72. Hellwege, K.H., Hellwege, A.M.: Structure Data of Free Polyatomic Molecules. Springer Verlag (1976)
73. ChemSpider. https://www.chemspider.com/. Accessed Jan 2022
74. Pence, H.E., Williams, A.: Chemspider: an online chemical information resource. J. Chem. Educ. **87**(11), 1123–1124 (2010)

75. Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., Li, Q., Shoemaker, B.A., Thiessen, P.A., Yu, B., Zaslavsky, L., Zhang, J., Bolton, E.E.: PubChem in 2021: new data content and improved web interfaces. Nucleic Acids Res. **49**, D1388–D1395 (2020)
76. Williams, A.J., Ekins, S., Tkachenko, V.: Towards a gold standard: regarding quality in public domain chemistry databases and approaches to improving the situation. Drug Discov. Today **17**(13–14), 685–701 (2012)
77. Halgren, T.A.: MMFF VI. MMFF94s option for energy minimization studies. J. Comput. Chem. **20**(7), 720–729 (1999)
78. Fiedler, L., Shah, K., Bussmann, M., Cangi, A.: A deep dive into machine learning density functional theory for materials science and chemistry. arXiv:2110.00997 (2021)
79. Nelson, R.D., Jr., Lide, D.R., Jr., Maryott, A.A.: Selected values of electric dipole moments for molecules in the gas phase. Tech. Rep, National Standard Reference Data System (1967)