Fractional Occupation Numbers and SIC-Scaling Methods with the Fermi-Löwdin Orbital SIC Approach

Fredy W. Aquino, Ravindra Shinde, Bryan M. Wong*

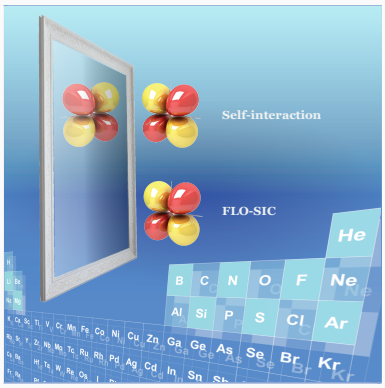
January 30, 2020

Abstract

We present a new assessment of the Fermi-Löwdin Orbital Self-Interaction Correction (FLO-SIC) approach with an emphasis on its performance for predicting energies as a function of fractional occupation numbers (FONs) for various multi-electron systems. Our approach in implemented in the massively-parallelized NWChem quantum chemistry software package and has been benchmarked on the prediction of total energies, atomization energies, and ionization potentials of small molecules and relatively large aromatic systems. Within our study, we also derive an alternate expression for the FLO-SIC energy gradient expressed in terms of gradients of the Fermi-orbital eigenvalues and revisit how the FLO-SIC methodology can be seen as a constrained unitary transformation of the canonical Kohn-Sham orbitals. Finally, we conclude with calculations of energies as a function of FONs using various SIC-scaling methods to test the limits of the FLO-SIC formalism on a variety of multi-electron systems. We find that these relatively simple scaling methods improve the prediction of total energies of atomic systems as well as enhance the accuracy of energies as a function of FONs for other multi-electron chemical species.

Keywords: self-interaction corrections, density functional theory, Fermi-Löwdin orbitals, fractional occupation numbers, ionization potentials. ■

*Department of Chemical & Environmental Engineering, Materials Science & Engineering Program, and Department of Physics & Astronomy, University of California-Riverside, Riverside, CA.
E-mail: bryan.wong@ucr.edu; Homepage: http://www.bmwong-group.com



Removing Self-Interaction Errors with the Fermi-Löwdin Orbital Self Interaction Correction (FLO-SIC) Approach

INTRODUCTION

The prediction of molecular properties using density functional theory (DFT)[1–3] continues to garner significant interest for studying large chemical/material systems (i.e., up to hundreds of atoms) due to its reasonable balance between computational efficiency and accuracy. DFT, in principle, is an exact theory for obtaining the ground state energy of a chemical/material system in terms of functionals of the electron density in an auxiliary system of non-interacting electrons (as opposed to other more computationally-expensive approaches that directly solve for the many-body wavefunction). However, the main practical limitation of DFT is its reliance on approximate exchange-correlation functionals that inherently introduce unphysical self-interactions between electrons (a notorious example of this is the dissociation of a H_2^+ molecule, where many common exchange-correlation functionals give unphysical results). These errors inherently arise from an incomplete cancellation of electronic interactions between the Coulomb and exchange-correlation term in approximate functionals. More concretely, for the case of one-electron densities, $\rho^{(1)}$, the expression for the energy, E_{xc}, of the exact (yet still unknown) exchange-correlation (xc) functional is given by

$$E_{\text{xc}}^{\text{exact}}[\rho^{(1)}] + J[\rho^{(1)}] = 0. \tag{1}$$

In other words, the self-Hartree repulsion energy exactly cancels out the self-exchange energy in an exact xc functional. However, for approximate xc functionals, this cancellation is not perfect,[4–6] and the error due to this spurious non-cancellation of energies is known as the self-interaction error (SIE):

$$E_{\text{xc}}^{\text{approx}}[\rho^{(1)}] + J[\rho^{(1)}] = E^{\text{SIE}}[\rho^{(1)}], \tag{2}$$

where $E^{\text{SIE}}[\rho^{(1)}]$ is the one-electron SIE. While Eq. 2 was written in terms of one-electron densities, SIE is present (and actually more deleterious) in many-electron systems and needs to be formally removed from all approximate exchange-correlation functionals. However, it is not straightforward to rectify these errors analytically for a given xc functional, and SIE corrections to the total energy must be carried out numerically in a systematic way. To further assess the effects of SIE in many-electron systems, an in-depth study of energies as a

function of fractional occupation numbers (and its deviation from linearity) can also provide additional insight into SIE, which is one of the primary thrusts of this current work.

In 1981, Perdew and Zunger proposed an orbital-dependent numerical scheme for the explicit orbital-by-orbital removal of SIE from the total energy.[4] Within this procedure, spatially-localized orbitals[7,8] are constructed for the minimization of the PZ self-interaction correction (SIC) to the total energy. In the present study, we compute the orbitals using the concept of the Fermi hole[9] to build spatially localized Fermi orbitals in one step. These orbitals are characterized by Fermi-orbital descriptors[10] (FODs) that can be interpreted as quasi-classical positions of electrons. These localized orbitals are then symmetrically orthonormalized using a Löwdin orthogonalization approach. A numerical optimization procedure (e.g., conjugate gradient,[11] pre-conditioned conjugate gradient,[12–14] or BFGS[15–18]) is then used to minimize the PZ-SIC energy as a function of Fermi-Löwdin orbital densities after an SCF DFT energy is obtained. Throughout the optimization process, the Fermi-orbital descriptors are updated, which redefines the Fermi-Löwdin orbitals associated to each electron in the system. Originally proposed by Pederson and co-workers,[10,19,20] the main advantage of this Fermi-Löwdin Orbital Self-Interaction Correction (FLO-SIC) approach is its use of a constrained unitary invariant transformation that also maintains size-extensivity.[20,21]

In the present paper we provide an analysis of our FLO-SIC implementation with the following new features and additions: (1) we re-visit the FLO-SIC methodology in the context of a constrained unitary transformation and present an alternate expression for the FLO-SIC energy gradient expressed in terms of gradients of the Fermi-orbital eigenvalues, (2) we examine the performance of our FLO-SIC methodology for the calculations of energies as a function of fractional occupation numbers (FONs) on representative chemical systems such as H_2, the carbon atom, and the diamine molecular cation, and (3) we test the accuracy of our FLO-SIC methodology in conjunction with various SIC-scaling methods. The latter two features are the most important results of our work since numerical tests of energies as a function of FONs (and their deviation from linearity) provides a stringent assessment of SIE, which has not previously been extensively studied with the FLO-SIC approach in many-electron systems. Finally, all of our FLO-SIC implementations are incorporated in the open-source, massively-parallelized NWChem quantum chemistry software package,[22] which

will be publicly available in the next release update.

THEORETICAL METHODS

Before proceeding to our detailed comparison of SIC approaches, we give a brief overview of the FLO-SIC formalism. Figure (1) depicts a simplified algorithmic flowchart of our FLO-SIC methodology implemented in the NWChem software package. The entire numerical approach consists of the following steps:

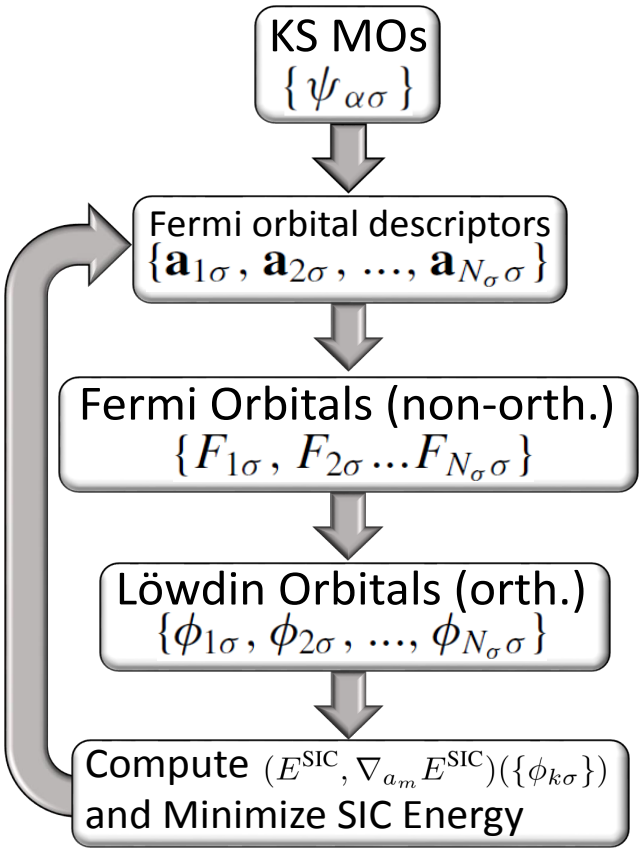


Figure 1: Algorithmic flowchart for the Fermi-Löwdin Orbital SIC (FLO-SIC) methodology implemented in the NWChem software package.

- **Step 1:** A converged set of Kohn-Sham molecular orbitals (KS MOs), $\{\psi_{\alpha\sigma}; \alpha = 1, 2, \cdots, N_\sigma\}$, are used as inputs, where N_σ is the number of electrons for each polar-

ization, σ.

- **Step 2:** Next, we determine a set of initial Fermi orbital descriptors (FODs), $\{\mathbf{a}_{1\sigma}, \mathbf{a}_{2\sigma}, \cdots, \mathbf{a}_{N_\sigma\sigma}\}$, using a Foster-Boys[23] localization algorithm. Note: the Foster-Boys localization is only performed at the first iteration to generate an initial guess, and subsequent iterations are carried out with a numerical optimization algorithm.

- **Step 3:** A single-step localization of orbitals is achieved by constructing Fermi orbitals, $\{F_{1\sigma}, F_{2\sigma}, \cdots, F_{N_\sigma\sigma}\}$, which are parametrized by FODs.

- **Step 4:** We construct localized orthonormal Löwdin orbitals, $\{\phi_{1\sigma}, \phi_{2\sigma}, \cdots, \phi_{N_\sigma\sigma}\}$, using the Löwdin symmetric orthogonalization method.[24]

- **Step 5:** The PZ-SIC energy, $E_\sigma^{\text{PZ-SIC}}[\{\rho_{k\sigma}\}]$, and the corresponding gradient of the PZ-SIC energy, $\nabla_{\mathbf{a}_{m\sigma}} E_\sigma^{\text{PZ-SIC}}[\{\rho_{k\sigma}\}]$, are computed using orbital densities, $\rho_{k\sigma} = \phi_{k\sigma}^* \phi_{k\sigma}$, which depend on the previously computed Löwdin orbitals, $\phi_{k\sigma}$.

- **Step 6:** A PZ-SIC energy minimization is carried out for each polarization using the L-BFGS-b[25–28] algorithm. Steps 2, 3, 4, and 5 are repeated to achieve convergence of the PZ-SIC energy subject to a pre-defined convergence threshold.

In the following subsections, we provide a brief derivation of the quantities required in the FLO-SIC approach.

Constructing Fermi-Löwdin Orbitals

As mentioned in the previous section, the Fermi orbitals, $F_{i\sigma}$, are constructed from a given set of Kohn-Sham orbitals, $\{\psi_{\alpha\sigma}; \alpha = 1, 2, \cdots, N_\sigma\}$, and FODs, $\{\mathbf{a}_{i\sigma}; i = 1, 2, \cdots, N_\sigma\}$. Mathematically, the Fermi orbital, $F_{i\sigma}$, can be written as

$$F_{i\sigma}(\mathbf{r}; \mathbf{a}_{\mathbf{i}\sigma}) = \frac{\rho_\sigma(\mathbf{a}_{i\sigma}, \mathbf{r})}{\sqrt{\rho_\sigma(\mathbf{a}_{i\sigma})}} = \frac{\sum_{\alpha=1}^{N_\sigma} n_{\alpha\sigma} \psi_{\alpha\sigma}^*(\mathbf{a}_{i\sigma}) \psi_{\alpha\sigma}(\mathbf{r})}{\sqrt{\rho_\sigma(\mathbf{a}_{i\sigma})}} = \sum_\alpha F_{i\alpha}^\sigma \psi_\alpha(\mathbf{r}), \tag{3}$$

where the Fermi orbital coefficient, $F_{i\alpha}^{\sigma}$, and the gradient of the Fermi orbital coefficient, $\nabla_{\mathbf{a}_{i\sigma}} F_{i\alpha}^{\sigma}$, are given by

$$F_{i\alpha}^{\sigma} = \frac{n_{\alpha\sigma}\psi_{\alpha\sigma}^{*}(\mathbf{a}_{i\sigma})}{\rho_{\sigma}(\mathbf{a}_{i\sigma})^{1/2}}, \tag{4}$$

$$\nabla_{\mathbf{a}_{i\sigma}} F_{i\alpha}^{\sigma} = F_{i\alpha}^{\sigma}\left[\frac{\nabla_{\mathbf{a}_{i\sigma}}\psi_{\alpha\sigma}^{*}(\mathbf{a}_{i\sigma})}{\psi_{\alpha\sigma}^{*}(\mathbf{a}_{i\sigma})} - \frac{\nabla_{\mathbf{a}_{i\sigma}}\rho_{\sigma}(\mathbf{a}_{i\sigma})}{2\rho_{\sigma}(\mathbf{a}_{i\sigma})}\right]. \tag{5}$$

Eq. 3 is an extension of a formal definition given by Pederson et al.[10,19] to account for fractional occupation numbers, $n_{\alpha\sigma}$, which are used in our study of many-electron molecular systems.

Assuming that the Kohn-Sham MOs, $\psi_{\alpha\sigma}$, are orthonormal, the Fermi orbital overlap, S_{ij}^{σ}, and gradient of the Fermi orbital overlap, $\nabla_{\mathbf{a}_{j\sigma}} S_{ij}^{\sigma}$, are given by

$$S_{ij}^{\sigma} = \langle F_{i\sigma} \,|\, F_{j\sigma}\rangle = \sum_{\alpha} F_{\alpha i}^{\sigma *} F_{j\alpha}^{\sigma} \tag{6}$$

$$\nabla_{\mathbf{a}_{j\sigma}} S_{ij}^{\sigma} = \left\langle F_{i\sigma} \,\middle|\, \nabla_{\mathbf{a}_{j\sigma}} F_{j\sigma}\right\rangle = \sum_{\alpha} F_{\alpha i}^{\sigma *} \nabla_{\mathbf{a}_{j\sigma}} F_{j\alpha}^{\sigma}. \tag{7}$$

Upon diagonalization of the Fermi orbital overlap matrix, we obtain the eigenvalues Q_{α}^{σ} and the corresponding eigenvectors $T_{\alpha j}^{\sigma}$:

$$\sum_{j} S_{ij}^{\sigma} T_{\alpha j}^{\sigma} = Q_{\alpha}^{\sigma} T_{\alpha i}^{\sigma}. \tag{8}$$

Using Eq. 8 and the intermediate Löwdin orbitals, $|T_{\alpha}^{\sigma}\rangle$,

$$|T_{\alpha}^{\sigma}\rangle = \sum_{j} T_{\alpha j}^{\sigma} \left|F_{j}^{\sigma}\right\rangle, \tag{9}$$

the Löwdin orbitals (LOs) are constructed from a Löwdin symmetric orthonormalization,[29] which gives[10,19]

$$|\phi_{k\sigma}\rangle = \sum_{n} \phi_{kn}^{\sigma} \left|F_{n}^{\sigma}\right\rangle, \text{where } \phi_{kn}^{\sigma} = \sum_{\alpha} \frac{T_{\alpha k}^{\sigma} T_{\alpha n}^{\sigma}}{\sqrt{Q_{\alpha}^{\sigma}}}. \tag{10}$$

Perdew-Zunger self-interaction correction, $E^{\text{PZ-SIC}}$

For a given exchange-correlation functional, the PZ-SIC expression for the energy is given by

$$E_{\sigma}^{\text{PZ-SIC}} = -\sum_{k}\left\{E_{xc}^{\text{approx}}[\rho_{k\sigma}, 0] + \frac{1}{2}\int\int d\mathbf{r} d\mathbf{r}' \frac{\rho_{k\sigma}(\mathbf{r})\rho_{k\sigma}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\right\}, \tag{11}$$

where $E_{xc}^{\text{approx}}[\rho_{k\sigma}, 0]$ is the energy obtained from an LDA or GGA xc functional, and the orbital densities, $\rho_{k\sigma}$, are given by

$$\rho_{k\sigma}(\mathbf{r}) \quad = \quad \phi^*_{k\sigma}(\mathbf{r})\phi_{k\sigma}(\mathbf{r}). \tag{12}$$

It is worth mentioning that since the FLO-SIC methodology uses a (constrained) unitary-invariant transformation of the Kohn-Sham MOs, the following expression holds:

$$\rho_\sigma(\mathbf{r}) = \sum_k \phi^*_{k\sigma}(\mathbf{r})\phi_{k\sigma}(\mathbf{r}) = \sum_k \psi^*_{k\sigma}(\mathbf{r})\psi_{k\sigma}(\mathbf{r}), \tag{13}$$

where $\psi_{k\sigma}(\mathbf{r})$ denotes a canonical Kohn-Sham MO, and $\phi_{k\sigma}(\mathbf{r})$ denotes a localized Löwdin MO. Most importantly, the expression in Eq. 13 implies that the ground state energy, including any other quantity dependent on the electron density, $\rho_\sigma(\mathbf{r})$, will be invariant under unitary transformations.

Gradient of $E_\sigma^{\text{PZ-SIC}}$

We present a new analytical expression for the FLO-SIC energy gradient expressed entirely in terms of gradients of the Fermi-orbital eigenvalues. In 2015, Pederson and co-workers presented expressions for the gradient of $E_\sigma^{\text{PZ-SIC}}$ with respect to the FODs:

$$\nabla_{\mathbf{a}_{m\sigma}} E_\sigma^{\text{PZ-SIC}} = \sum_{k=1}^{N_\sigma} \sum_{l=1,(l\neq k)}^{N_\sigma} \lambda_{kl}^{k\sigma}(\vec{\Delta}_{lk,m}^{1\sigma} + \vec{\Delta}_{lk,m}^{3\sigma}), \tag{14}$$

where $\vec{\Delta}_{lk,m}^{1\sigma}$ and $\vec{\Delta}_{lk,m}^{3\sigma}$ are vector quantities defined as

$$\vec{\Delta}_{lk,m}^{1\sigma} \quad = \quad \sum_{\alpha\beta n} (\nabla_{\mathbf{a}_{m\sigma}} S_{nm}^\sigma) T_{\alpha m}^\sigma T_{\beta n}^\sigma \frac{T_{\alpha k}^\sigma T_{\beta l}^\sigma - T_{\alpha l}^\sigma T_{\beta k}^\sigma}{\sqrt{Q_\alpha^\sigma Q_\beta^\sigma}}, \tag{15}$$

$$\vec{\Delta}_{lk,m}^{3\sigma} \quad = \quad -\frac{1}{2} \sum_{\alpha\beta n} \nabla_{\mathbf{a}_{m\sigma}} S_{nm}^\sigma (T_{\beta n}^\sigma T_{\alpha m}^\sigma + T_{\beta m}^\sigma T_{\alpha n}^\sigma)(T_{\alpha k}^\sigma T_{\beta l}^\sigma - T_{\alpha l}^\sigma T_{\beta k}^\sigma) \frac{\sqrt{Q_\beta^\sigma} - \sqrt{Q_\alpha^\sigma}}{\sqrt{Q_\alpha^\sigma Q_\beta^\sigma}(\sqrt{Q_\alpha^\sigma} + \sqrt{Q_\beta^\sigma})} \tag{16}$$

We have derived a new expression for Eq. 14 that involves gradients of Fermi orbital eigenvalues, summarized in Eqs. 17 - 22 below:

$$\nabla_{\mathbf{a}_{m\sigma}} E_\sigma^{\text{PZ-SIC}} \quad = \quad \sum_{k>l} (\lambda_{kl}^{k\sigma} - \lambda_{lk}^{l\sigma})(\vec{\Delta}_{lk,m}^{1\sigma} + \vec{\Delta}_{lk,m}^{3\sigma}), \tag{17}$$

with $\vec{\Delta}^{1\sigma}_{lk,m}$, $\vec{\Delta}^{3\sigma}_{lk,m}$, and $\lambda^{k\sigma}_{lk}$ given by

$$\vec{\Delta}^{1\sigma}_{lk,m} = \frac{1}{2}\phi^{\sigma}_{km}\sum_{\beta}\frac{\nabla_{\mathbf{a}_{m\sigma}}Q^{\sigma}_{\beta}}{\sqrt{Q^{\sigma}_{\beta}}}\frac{T^{\sigma}_{\beta l}}{T^{\sigma}_{\beta m}} - \frac{1}{2}\phi^{\sigma}_{lm}\sum_{\beta}\frac{\nabla_{\mathbf{a}_{m\sigma}}Q^{\sigma}_{\beta}}{\sqrt{Q^{\sigma}_{\beta}}}\frac{T^{\sigma}_{\beta k}}{T^{\sigma}_{\beta m}} \tag{18}$$

$$\vec{\Delta}^{3\sigma}_{lk,m} = -\sum_{\beta>\alpha}(T^{\sigma}_{\alpha k}T^{\sigma}_{\beta l} - T^{\sigma}_{\alpha l}T^{\sigma}_{\beta k})\frac{\sqrt{Q^{\sigma}_{\beta}} - \sqrt{Q^{\sigma}_{\alpha}}}{\sqrt{Q^{\sigma}_{\alpha}Q^{\sigma}_{\beta}}(\sqrt{Q^{\sigma}_{\alpha}} + \sqrt{Q^{\sigma}_{\beta}})}\left[T^{\sigma}_{\alpha m}\frac{\nabla_{\mathbf{a}_{m\sigma}}Q^{\sigma}_{\beta}}{2T^{\sigma}_{\beta m}} + T^{\sigma}_{\beta m}\frac{\nabla_{\mathbf{a}_{m\sigma}}Q^{\sigma}_{\alpha}}{2T^{\sigma}_{\alpha m}}\right] \tag{19}$$

$$\lambda^{k\sigma}_{lk} = \left\langle \phi_{l\sigma} \middle| V^{\mathrm{SIC}}_{k\sigma} \middle| \phi_{k\sigma} \right\rangle, \tag{20}$$

where ϕ^{σ}_{km} was previously defined in Eq. 10, and the SIC potential, $V^{\mathrm{SIC}}_{k\sigma}$, is given by

$$V^{\mathrm{SIC}}_{k\sigma} = \frac{\delta E^{\text{PZ-SIC}}}{\delta\rho_{k\sigma}} = -\frac{\delta E^{\mathrm{approx}}_{xc}[\rho_{k\sigma}, 0]}{\delta\rho_{k\sigma}} - \int d\mathbf{r}\frac{\rho_{k\sigma}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \tag{21}$$

and $\nabla_{\mathbf{a}_{m\sigma}}Q^{\sigma}_{\alpha}$ is

$$\nabla_{\mathbf{a}_{m\sigma}}Q^{\sigma}_{\alpha} = 2T^{\sigma}_{\alpha m}\sum_{j}(\nabla_{\mathbf{a}_{m\sigma}}S^{\sigma}_{jm})T^{\sigma}_{\alpha j}. \tag{22}$$

The most salient feature of this mathematical formulation is that $E^{\text{PZ-SIC}}_{\sigma}$ in Eq. 17 can be expressed in terms of gradients of Fermi orbital eigenvalues, $\nabla_{\mathbf{a}_{m\sigma}}Q^{\sigma}_{\alpha}$. In other words, one can, in principle, minimize the PZ-SIC energy via a minimization of the set of all Fermi orbital eigenvalues[30] $\{Q^{\sigma}_{\alpha}; \alpha = 1, 2, \cdots, N_{\sigma}\}$. This result is particularly interesting since the gradients of the Fermi-orbital eigenvalues do not depend on the computationally-expensive evaluation of two-electron integrals and, therefore, could possibly be used to accelerate further FLO-SIC calculations (which we save for future work). Nevertheless, it is worth mentioning that while this approach avoids the computation of two-electron integrals, the minimization of the set of all Fermi orbital eigenvalues is a multi-objective optimization problem[31–33] (compared to the original single-objective optimization problem of Eq. 17), which can pose additional numerical challenges.

Previous studies have used a variety of approaches for minimizing the SIC energy, which include gradients of the SIC energy with respect to (1) elements of a unitary transformation,[6,34] (2) Kohn-Sham orbital coefficients,[35] and (3) Fermi-orbital descriptor positions.[19] It is worth mentioning that the number of minimization parameters (N^2) in the first two approaches is greater than the number of parameters ($3N$) used in the FLO-SIC energy minimization, making them more computationally expensive. The connection between FLO-SIC and the other approaches is not trivial to establish. For a constrained unitary transformation

defined as $\mathbf{U}^\sigma\mathbf{C}^\sigma = (\mathbf{L}^\sigma\mathbf{F}^\sigma)\mathbf{C}^\sigma$, where $\mathbf{L}^\sigma$ and $\mathbf{F}^\sigma$ are the Löwdin and Fermi orbital coefficient matrices, Eq. 17 can be written as follows:

$$\frac{\partial E_\sigma^{\text{PZ-SIC}}}{\partial \mathbf{a}_{k\sigma}} = \sum_{ij} \frac{\partial E_\sigma^{\text{PZ-SIC}}}{\partial u_{ij}^\sigma} \frac{\partial u_{ij}^\sigma}{\partial \mathbf{a}_{k\sigma}} = 0, \tag{23}$$

with $\frac{\partial u_{ij}^\sigma}{\partial \mathbf{a}_{k\sigma}}$ given by

$$\frac{\partial u_{ij}^\sigma}{\partial \mathbf{a}_{k\sigma}} = \sum_{n} \frac{\partial \phi_{in}^\sigma}{\partial \mathbf{a}_{k\sigma}} F_{nj}^\sigma + \phi_{ik}^\sigma \frac{\partial F_{kj}^\sigma}{\partial \mathbf{a}_{k\sigma}}, \tag{24}$$

Eq. 23 shows the relationship between the FLO-SIC energy gradients, $\frac{\partial E_\sigma^{\text{PZ-SIC}}}{\partial \mathbf{a}_{k\sigma}}$, and the SIC energy gradient, which is dependent on the N^2 elments of the unitary transformation matrix. from previous studies. A particular solution of Eq. 23 is

$$\frac{\partial E_\sigma^{\text{PZ-SIC}}[u_{ij}^\sigma(\{\mathbf{a}_k\})]}{\partial u_{ij}^\sigma} = 0, \tag{25}$$

which results in N^2 nonlinear equations for $3N$ unknowns, $\{\mathbf{a}_k\}$. Nevertheless, Eq. 23 is a nonlinear system of $3N$ equations for $3N$ unknowns that may be satisfied by more than one single configuration of FODs, $\{\mathbf{a}_k\}$. In other words, this implies that several configurations of FODs, $\{\mathbf{a}_k\}$, can give the same SIC energy, which we and others have also found in previous work.[36]

Scaling factors for $E_\sigma^{\text{PZ-SIC}}$

Following Vydrov and co-workers,[35,37] we also tested the usage of scaling factors to remedy over-corrections introduced by the PZ-SIC energy expression for many-electron systems:

$$E_\sigma^{\text{PZ-SIC}} = -\sum_i X_{i\sigma}^k \left\{ E_{xc}^{\text{approx}}[\rho_{i\sigma}, 0] + \frac{1}{2}\int\int d\mathbf{r}d\mathbf{r}' \frac{\rho_{i\sigma}(\mathbf{r})\rho_{i\sigma}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \right\}. \tag{26}$$

The scaling factor, $X_{i\sigma}^k$, must equal unity for one-electron systems or satisfy $0 \leq X_{i\sigma}^k \leq 1$ for the multi-electron case. In this work, we tested two scaling factors, denoted by X^k and X^m that satisfy these criteria. The X^k scaling method uses an $X_{i\sigma}^k$ with the following form

$$X_{i\sigma}^k = \frac{1}{n_{i\sigma}} \int \left(\frac{\tau_\sigma^W}{\tau_\sigma}\right)^k \rho_{i\sigma}(\mathbf{r}) d\mathbf{r}, \tag{27}$$

where τ_σ is the non-interacting kinetic energy density with polarization σ:

$$\tau_\sigma(\mathbf{r}) = \frac{1}{2}\sum_i n_{i\sigma}|\nabla\phi_{i\sigma}(\mathbf{r})|^2, \tag{28}$$

and τ_σ^W is the von Weizsäcker kinetic energy density for polarization σ:

$$\tau_\sigma^W(\mathbf{r}) = \frac{|\nabla\rho_\sigma(\mathbf{r})|^2}{8\rho_\sigma(\mathbf{r})}, \tag{29}$$

where $n_{i\sigma}$ is the occupation number and k is a non-negative real number.

The X^m scaling method utilizes a $X_{i\sigma}^m$ of the form

$$X_{i\sigma}^m = \frac{1}{n_{i\sigma}}\int\left(\frac{\rho_{i\sigma}}{\rho_\sigma}\right)^m \rho_{i\sigma}d\mathbf{r}, \tag{30}$$

where m is a non-negative real number. Since the X^k method requires gradients of the charge densities, it is more computationally expensive than the X^m scaling method. It should be noted that most xc functionals do not need corrections in the uniform density limit; furthermore, the X^k scaling method vanishes for uniform densities, whereas the X^m scaling method is not guaranteed to vanish, making the former method more convenient. In this work, we test the performance of these scaling factors in the computation of relative energies of atoms ranging from He to Ar, as well as for energies as a function of FONs for the carbon atom and the diamine cation molecule, which was recently examined by us as a prototypical chemical system to test density functional methods.[38]

RESULTS AND DISCUSSION

To validate our custom implementation of FLO-SIC in the open-source, massively-parallelized NWChem software package, we first carried out a series of benchmark calculations on total energies, atomization energies, and ionization potentials of several small molecules. These calculations were compared against a different implementation of FLO-SIC in the literature, and both approaches were found to be in excellent agreement with each other. A detailed description of these calculations and benchmarks is given in the Supporting Information.

In the following subsections, we discuss the accuracy of the various SIC-scaling methods for the He – Ar atomic systems, which are compared against other benchmark calculations.[39]

We conclude our discussion with an analysis of fractional occupation numbers for many-electron systems such as H_2, the carbon atom, and the transition-state geometry of the diamine molecular cation to assess the performance of the FLO-SIC formalism in conjunction with the various scaling methods.

SIC Energies for Atomic Systems

In this section, we present total energies for atomic systems from Helium to Argon using the Perdew-Burke-Ernzerhof (PBE) xc functional[40,41] and various SIC corrections to this GGA reference state. Figure (2) plots the relative energy per electron, $(E - E_{\text{ref}})/Z$, where E is the total energy obtained with various methods (such as PBE, PBE/SCF-SIC, or PBE/FLO-SIC), E_{ref} are reference energies from highly-accurate benchmark values,[42,43] and Z is the atomic number. The fully self-consistent PBE/SCF-SIC calculations were computed with the ERKALE[6] software package, and the PBE and PBE/FLO-SIC energies were obtained from our own NWChem implementations and modifications. As an extra check on our results, we have verified that our NWChem-computed PBE energies coincide with the PBE energies obtained from the ERKALE software package. Although our FLO-SIC implementation was only carried out in a post-SCF mode, Figure 2 demonstrates that the FLO-SIC scheme still gives accurate results compared to the more computationally-expensive, self-consistent PZ-SIC approach. Since PBE/FLO-SIC progressively gets worse at predicting accurate total energies as Z increases, we tested the performance of various SIC-energy scaling factors (i.e., setting $k = 1/2$, 1, and 3 within the X^k scaling method) to understand their performance on softening these overcorrections. Interestingly, we find that the scaling factor has a more significant effect on systems with atomic numbers larger than 5, with the best overall results obtained with $k = 3$. Additional validation tests of this approach and corresponding plots with the X^m scaling method for $m = 1/2$, 1, and 3 are given in the Supporting Information.

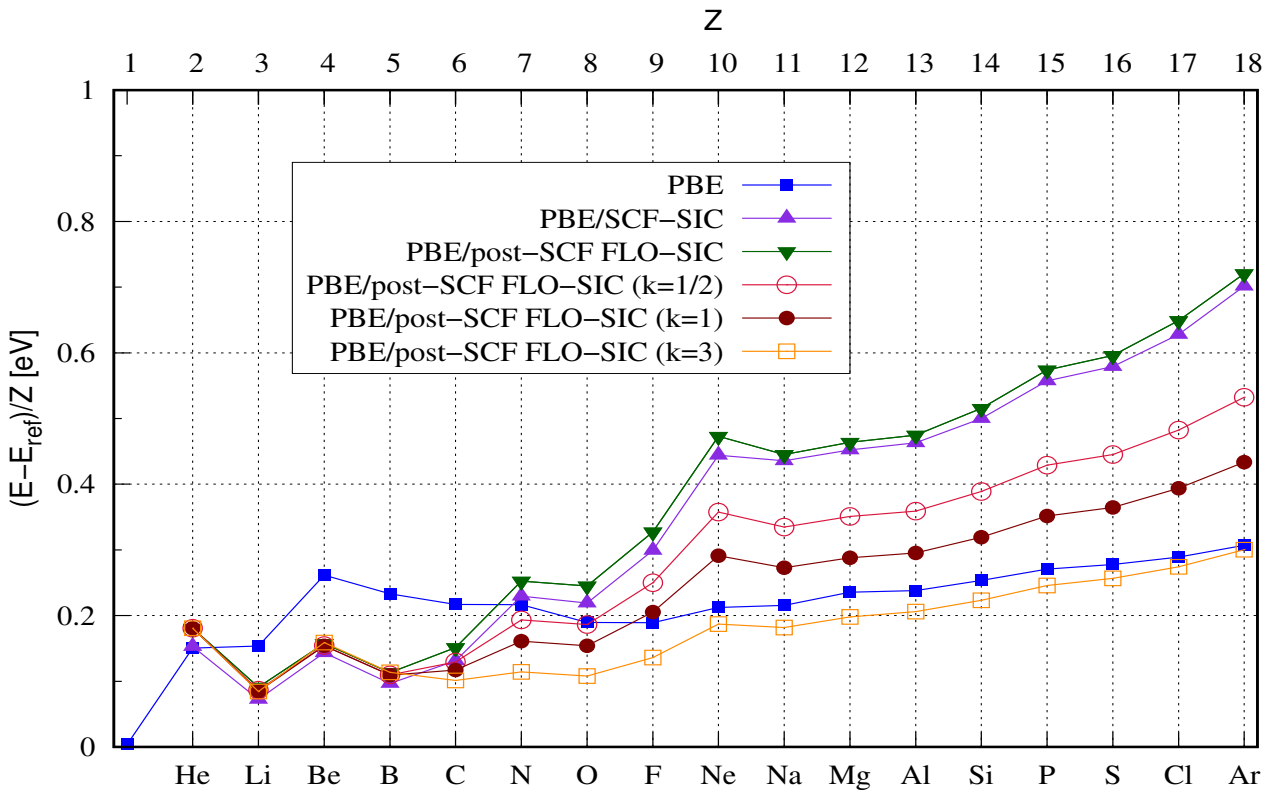


Figure 2: Relative total energies per electron for atomic systems ranging from He ($Z = 2$) to Ar ($Z = 18$) obtained with PBE, PBE/SCF-SIC, and various X^k-scaled energies for FLO-SIC. The aug-cc-pVTZ basis set was used in all of the calculations.

Energy vs. FONs for H_2, the carbon atom, and the diamine cation

One rigorous diagnostic to assess self-interaction errors (SIE) in an exchange-correlation functional is to check the energy linearity theorem for fractional occupations:[44–46]

$$E[N + \delta N] = (E[N + 1] - E[N])(\delta N - 1) + E[N + 1], \tag{31}$$

which shows that the total energy is linear as a function of the fractional occupations δN. In the interval $[N, N + 1]$ the slope $\partial E/\partial n$ is the negative of the electron affinity, $-$EA, which is the difference between the energy of an anion and a neutral system, $E^{N+1} - E^N$. Poorly behaved functionals such as LDA are "concave up" in the E vs. N plot, whereas the Hartree-Fock approach shows a "concave down" behavior. In contrast, range-separated functionals[47–52] show a straight line behavior,[47] which nearly satisfies Eq. 31 with a slope

that approximates the electron affinity. For our study on FONs, we also confirmed that the total number of electrons (including fractional numbers) were conserved during the FLO-SIC numerical procedure. Specifically, the total number of electrons is given by

$$N = \int \rho_\sigma(\mathbf{r}) d\mathbf{r} = \int \sum_{\mu\nu} P_{\mu\nu} \psi_\mu(\mathbf{r}) \psi^*_\nu(\mathbf{r}) d\mathbf{r}, \tag{32}$$

where $P_{\mu\nu}$ is the density matrix and $\psi_\mu(\mathbf{r})$ and $\psi_\nu(\mathbf{r})$ are Kohn-Sham orbitals. Substituting the Fermi-Löwdin density matrix, $P^{\mathrm{FL}}_{\mu\nu} = \sum_\alpha n_{\alpha\sigma} \left(C^{\mathrm{FL},\sigma}_{\alpha\mu}\right) \left(C^{\mathrm{FL},\sigma}_{\alpha\nu}\right)^*$ where $C_{\alpha\nu}$ are orbital coefficients, for $P_{\mu\nu}$ in the expression above gives

$$N = \sum_\alpha n_{\alpha\sigma} \int \sum_{\mu\nu} \left(C^{\mathrm{FL},\sigma}_{\alpha\mu}\right) \left(C^{\mathrm{FL},\sigma}_{\alpha\nu}\right)^* \psi_\mu(\mathbf{r}) \psi^*_\nu(\mathbf{r}) d\mathbf{r} = \sum_\alpha n_{\alpha\sigma}, \tag{33}$$

where we have used the fact that the Fermi-Löwdin orbitals are orthogonal in the last step; i.e., the total number of electrons, N, is the sumation of the occupation of each orbital.

Figure 3 shows our results for the relative energy of H_2 as a function of the fractional electron number $(n + \delta n)$ at various levels of theory. The benchmark reference is a straight line obtained from the highly accurate, wavefunction-based CCSD(T)/aug-cc-pVTZ calculations. The results obtained with LC-BLYP/aug-cc-pVTZ and LDA-FLO-SIC/aug-cc-pVTZ agree very well with the CCSD(T) results, which demonstrate an accurate removal of the SIE; in contrast, the LDA/aug-cc-pVTZ results show the expected "concave up" behavior. Figure 4 depicts various E vs. N curves for the carbon atom, where our benchmark reference is the dashed black line with a slope equal to $-$EA$= -1.26$ eV obtained from experiment.[53] The E vs. N curves obtained with LDA and PBE are "concave up," and the LC-BLYP functional yields a straight line with a slope slightly different from the experimental benchmark. For comparison, we have also included E vs. N plots obtained with various SIC approaches including LDA/FLOSIC, PBE/FLO-SIC, the fully self-consistent PBE/SCF-SIC from Vydrov et al.[35] Among these various "uncorrected" SIC methods, we find that LDA/FLO-SIC yields the best results, although it exhibits a curvature that is still slightly over-localized. In the same plot, we also tested the performance of the X^k scaling method (with $k = 1$) for LDA/FLO-SIC and PBE/FLO-SIC, which slightly improves the linearity and slope of their "uncorrected" counterparts.

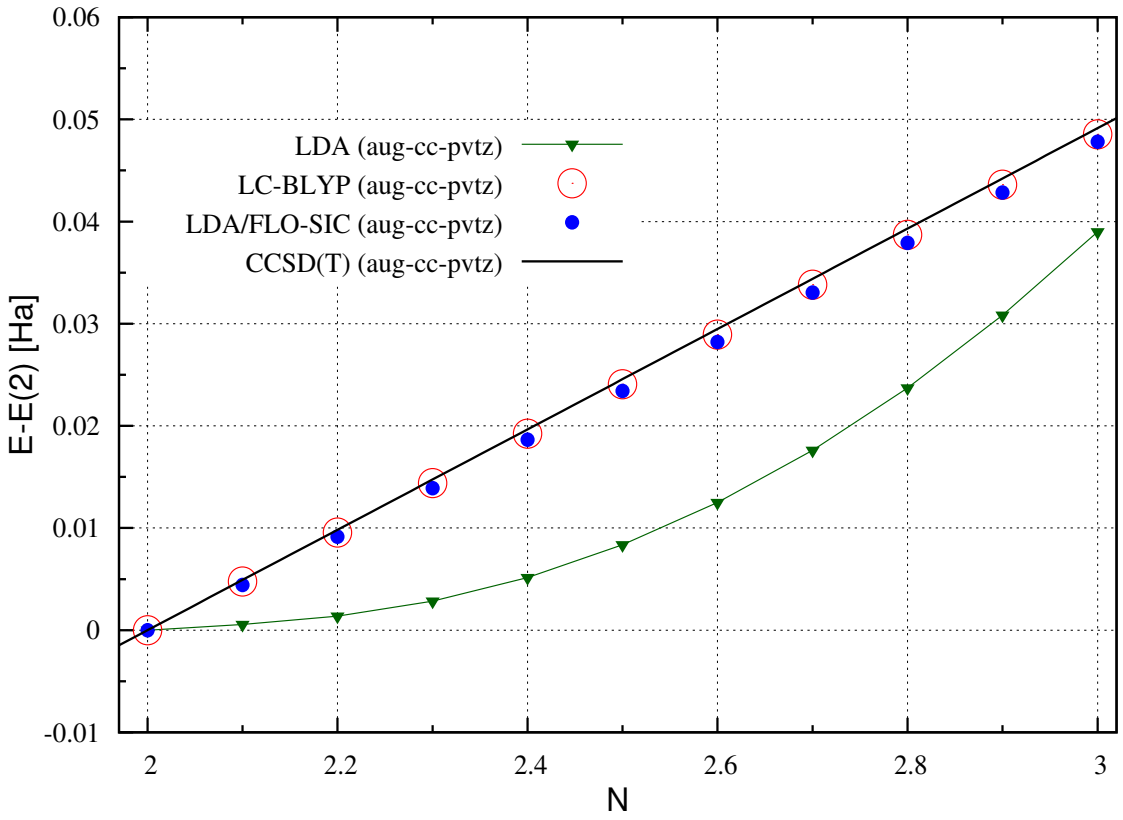


Figure 3: Relative total energies of H_2 as a function of the electron number N obtained using several levels of theory: LDA, LDA/FLO-SIC, and LC-BLYP. The reference dashed line was obtained with a single-point CCSD(T) calculation at $E[N]$ and $E[N+1]$ using the aug-cc-pVTZ basis set.

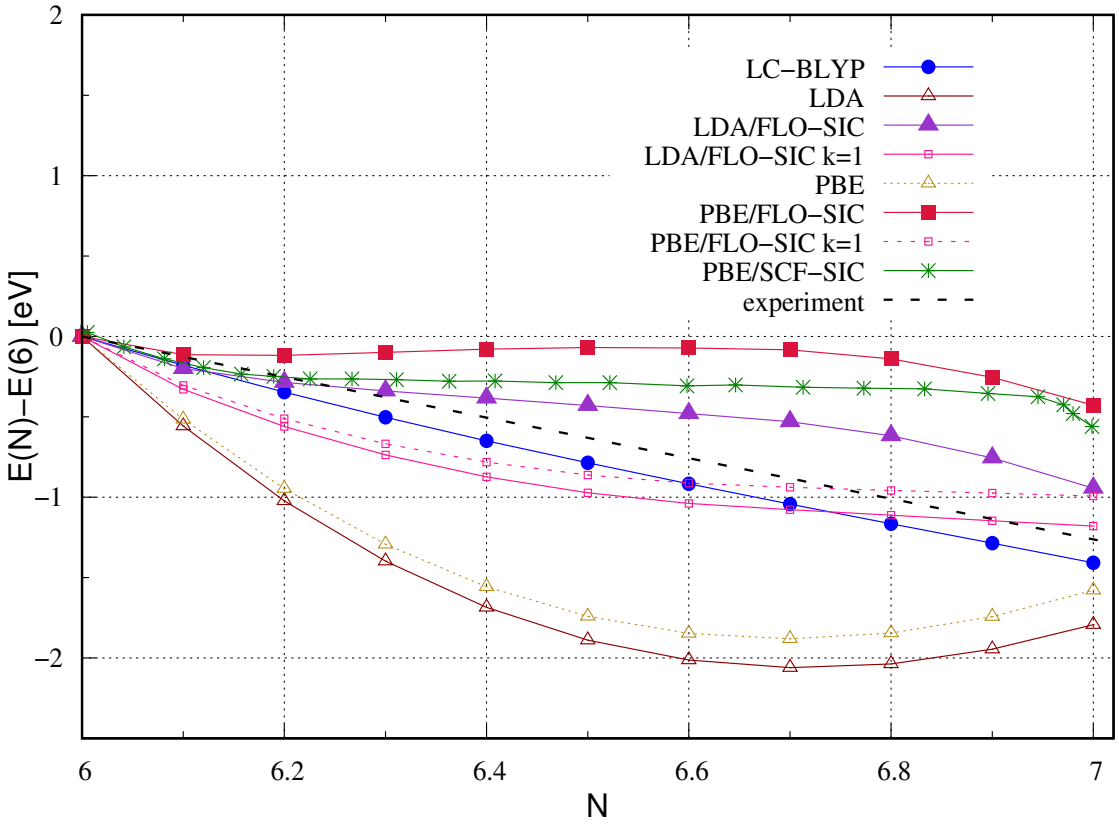


Figure 4: Relative total energies of the carbon atom as a function of the electron number N obtained using several functionals with the aug-cc-pVTZ basis set. For comparison, the PBE/SCF-SIC calculation of Vydrov et al.[35] is also shown. The dashed black line was obtained using the experimental electron affinity.

To analyze the FLO-SIC scaling trends more closely, the left panel of Figure 5 depicts deviations from linearity for PBE/FLO-SIC in conjunction with the X^k scaling method for k =0.2, 0.5, and 1. Comparing these various curves, we observe a slight improvement for $k = 0.2$. Similarly, the right panel of Figure 5 shows deviations from linearity for LDA/FLO-SIC with the X^k scaling method for k =0.1, 0.2, 0.5, and 1, and we observe the best results for $k = 0.1$. The Supporting Information contains additional E vs. N and deviation-from-linearity plots for the X^m scaling method for LDA, LDA/FLO-SIC, PBE, and PBE/FLO-SIC.

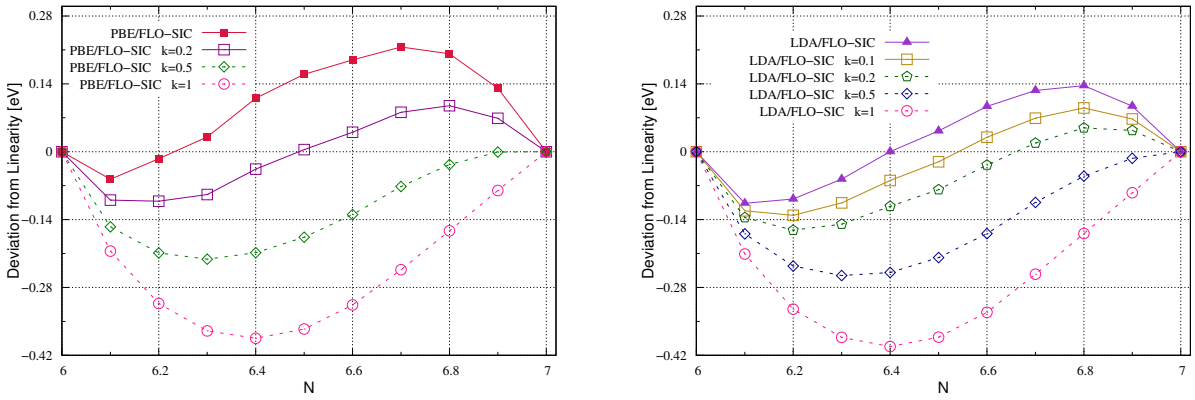


Figure 5: Deviations from linearity as a function of FONs for the carbon atom using PBE/FLO-SIC and X^k-scaled PBE/FLO-SIC (left panel) and LDA/FLO-SIC and X^k-scaled LDA/FLO-SIC (right panel).

As a final application of our FLO-SIC approach for energies as a function of FONs, we analyze the electronic structure of the transition state of the N,N'-dimethylpiperazine (DMP) diamine molecular cation (DMP-TS), which was recently examined by us as a prototypical chemical system to test density functional methods.[38] Figure 6 shows E vs. N plots for the CCSD-optimized transition-state geometry of DMP-TS for different levels of theory with the aug-cc-pVDZ basis set. Our benchmark reference is the dashed line whose slope is obtained from CCSD(T) energies for the anion and neutral molecule, E^{N+1} and E^N, respectively. The range-separated LC-BLYP is almost indistinguishable from the reference, whereas the PBE functional shows a slight "concave up" curve. The PBE/FLO-SIC curve exhibits a good linearity between 63.4 and 64; however, in the interval between 63 and 63.4, we encountered convergence problems during the SIC optimization, which may be due to an improper choice of initial FODs obtained with our Foster-Boys localization algorithm. While these convergence problems manifest themselves as a "bump" in the energy vs. FON curve, it is important to note that the FLO-SIC energy gradients for those problematic points are still extremely small. Specifically, the inset of Figure 6 depicts a plot of the FLO-SIC gradient as a function of FONs, and all values of the gradient are extremely small and within 10^{-3} - 10^{-5}. In addition, we also examined PBE/FLO-SIC in conjunction with the X^k scaling method (using $k = 1$), which shows a slight improvement over the uncorrected PBE/FLO-SIC values (i.e., the curve within the 63.8 - 64.0 interval moves closer to the reference CCSD(T) straight

line. Figure 7 shows the corresponding deviation-from-linearity plots for PBE, PBE/FLO-SIC, and various X^k scaling corrections. The Supporting Information contains additional plots for the LDA and LDA/FLO-SIC functionals.

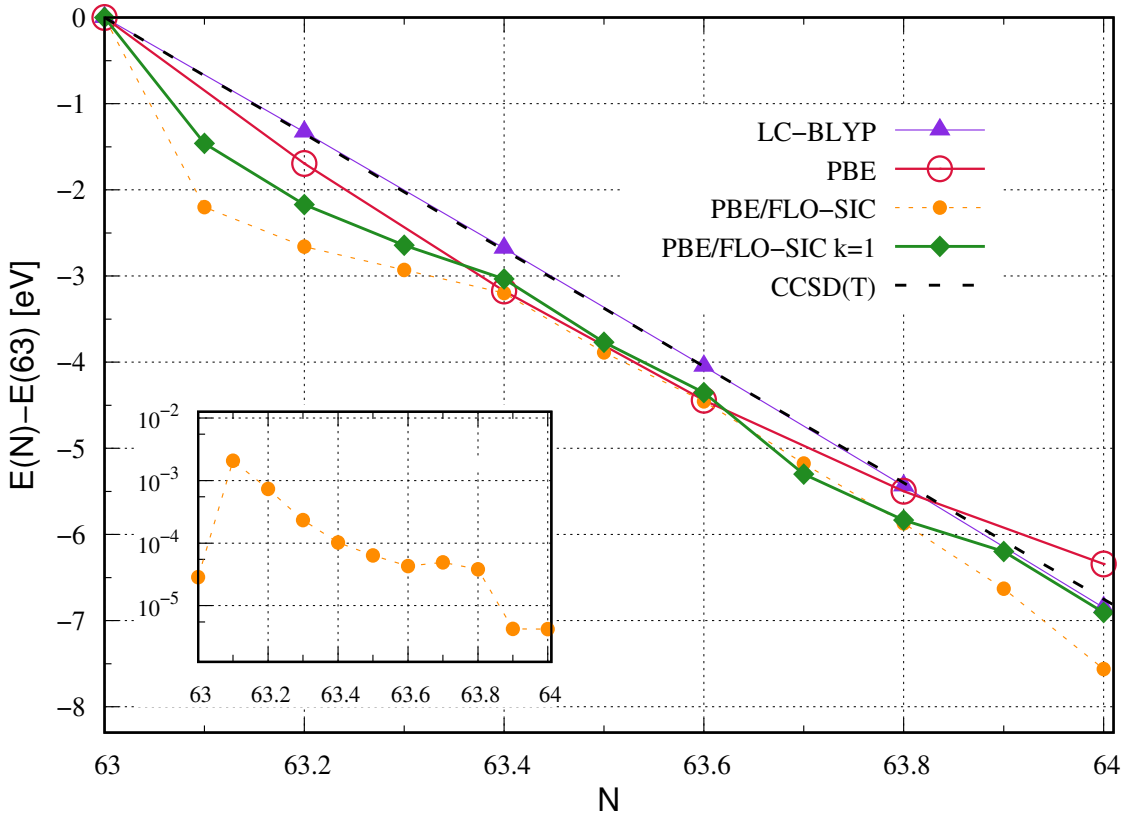


Figure 6: Relative energies of DMP-TS as a function of the fractional electron number N obtained using LC-BLYP, PBE, PBE/FLO-SIC, and PBE/FLO-SIC with the $X^k(k = 1)$ scaling method. All functionals utilized the aug-cc-pVDZ basis set. The benchmark reference is the black dashed line whose slope is computed using CCSD(T)/aug-cc-pVDZ calculations for E^{N+1} and E^N. The inset shows a logarithmic plot of the converged gradient of E^{SIC} vs. N.

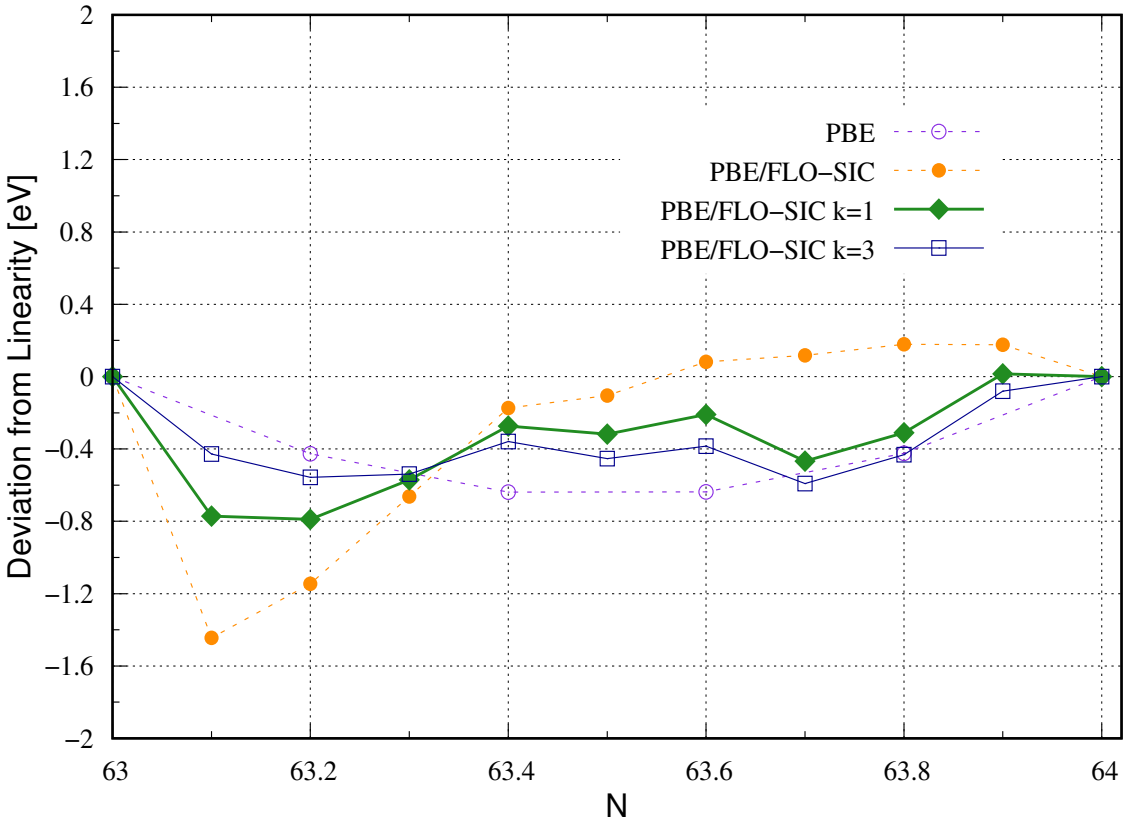


Figure 7: Deviations from linearity as a function of FONs for DMP-TS using PBE, PBE/FLO-SIC, and PBE/FLO-SIC with the X^k scaling method for k =1 and 3.

CONCLUSIONS

In conclusion, we have provided a new assessment of our FLO-SIC implementation with an emphasis on its performance for predicting energies as a function of fractional occupation numbers (FONs) of various multi-electron chemical systems. Within our analysis, we provided an alternate expression for the FLO-SIC energy gradient expressed in terms of gradients of Fermi-orbital eigenvalues. This new expression is particularly interesting and insightful since the gradients of the Fermi-orbital eigenvalues do not depend on the computationally-expensive evaluation of two-electron integrals and, therefore, could possibly be used to accelerate future FLO-SIC implementations. To validate our implementation, we carried out benchmark calculations on total energies, atomization energies, and ionization potentials for various atomic and molecular systems. Finally, we calculated energies

as a function of FONs with various SIC-scaling methods to test the limits of the FLO-SIC formalism on various multi-electron chemical systems, which have not been systematically examined in previous studies. We find that these relatively simple scaling methods improve the prediction of total energies of atomic systems as well as enhance the accuracy of energies as a function of FONs for the carbon atom and diamine molecular cation. Finally, all of our FLO-SIC implementations are incorporated in the open-source, massively-parallelized NWChem quantum chemistry software package,[22] which will be publicly available in the next release update.

ACKNOWLEDGMENTS

We gratefully acknowledge Torsten Han, Zeng-hui Yang, Tunna Baruah, and Susi Lehtola for helpful conversations and providing computational details for the FLO-SIC approach. This work was supported by the U.S. Department of Energy, Office of Science, Early Career Research Program under Award No. DE-SC0016269.

References

1. P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

2. W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

3. A. D. Becke, J. Chem. Phys. **140**, 18A301 (2014).

4. J. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

5. S. Klüpfel, P. Klüpfel, and H. Jónsson, J. Chem. Phys. **137**, 124102 (2012).

6. S. Lehtola and H. Jonsson, J. Chem. Theory Comput. **10**, 5324 (2014).

7. F. Jensen, *Introduction to Computational Chemistry* (John wiley & sons, 2017).

8. S. Lehtola and H. Jónsson, J. Chem. Theory Comput. **9**, 5365 (2013).

9. W. L. Luken and J. C. Culberson, Theor. Chem. Acc. **61**, 265 (1982).

10. M. Pederson and T. Baruah, Adv. Atom. Mol. Opt. Phy. **64**, 153 (2015).

11. J. C. Gilbert and J. Nocedal, SIAM J. Opt. **2**, 21 (1992).

12. A. V. Knyazev and I. Lashuk, SIAM J. Matrix Anal. Appl. **29**, 1267 (2007).

13. J. L. Nazareth, SIAM Rev. **28**, 501 (1986).

14. Y. Notay, SIAM J. Sci. Comput. **22**, 1444 (2000).

15. C. Broyden, IMA J. Appl. Math. **6**, 76 (1970).

16. R. Fletcher, Comput. J **13**, 317 (1970).

17. D. Goldfarb, Math. Comput. **24**, 23 (1970).

18. D. Shano, Mathematics of Computation **24**, 647 (1970).

19. M. Pederson, J. Chem. Phys. **142**, 064112 (2015).

20. M. Pederson, A. Ruzsinszky, and J. Perdew, J. Chem. Phys. **140**, 121103 (2014).

21. J. P. Perdew, Adv. Quantum Chem **21**, 113 (1990).

22. E. J. Bylaska, W. A. de Jong, N. Govind, K. Kowalski, T. P. Straatsma, M. Valiev, J. J. van Dam, D. Wang, E. Apra, T. L. Windus, and Others, Pacific Northwest National Laboratory: Richland, WA (2011).

23. J. Foster and S. Boys, Rev. Mod. Phys. **32**, 300 (1960).

24. P.-O. Löwdin, Reviews of Modern Physics **34**, 520 (1962).

25. R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, SIAM J. Sci. Comput. **16**, 1190 (1995).

26. J. L. Morales and J. Nocedal, ACM Trans. Math. Softw. **38**, 7 (2011).

27. C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal, ACM T. Math. Software **23**, 550 (1997).

28. J. Nocedal, Math. Comput. **35**, 773 (1980).

29. P.-O. Löwdin, J. Chem. Phys. **18**, 365 (1950).

30. C. Vogelbusch, *Self-interaction correction to electronic structure calculations with density functional theory: Analytical second order derivatives in FLO SIC*, Bachelor thesis, Technische Universität Bergakademie Freiberg, Faculty of Chemistry and Physics, Institute of Theoretical Physics (2016).

31. V. Chankong and Y. Y. Haimes, *Multiobjective Decision Making: Theory and Methodology* (Courier Dover Publications, 2008).

32. K. Miettinen, *Nonlinear Multiobjective Optimization*, Vol. 12 (Springer Science & Business Media, 2012).

33. K. Deb, *Multi-objective Optimization Using Evolutionary Algorithms*, Vol. 16 (John Wiley & Sons, 2001).

34. S. Klüpfel, *Implementation and Reassessment of the Perdew-Zunger Self-interaction Correction*, Ph.D. thesis, Faculty of Physical Sciences, School of Engineering and Natural Sciences, University of Iceland (2012).

35. O. A. Vydrov, *Correcting the Self-Interaction Error of Approximate Density Functionals*, Ph.D. thesis, Department of Chemistry, Rice University (2007).

36. F. W. Aquino and B. M. Wong, J. Phys. Chem. Lett. **9**, 6456 (2018).

37. O. A. Vydrov, G. E. Scuseria, J. P. Perdew, A. Ruzsinszky, and G. I. Csonka, J. Chem. Phys. **124**, 094108 (2006).

38. Z. A. Ali, F. W. Aquino, and B. M. Wong, Nat. Commun. **9**, 4733 (2018).

39. S. Lehtola, M. Hakala, A. Sakko, and K. Hämäläinen, J. Comput. Chem. **33**, 1572 (2012).

40. J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

41. J. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **78**, 1396 (1997).

42. E. R. Davidson, S. A. Hagstrom, S. J. Chakravorty, V. M. Umar, and C. F. Fischer, Phys. Rev. A **44**, 7071 (1991).

43. S. J. Chakravorty and E. R. Davidson, J. Phys. Chem. **100**, 6167 (1996).

44. T. Tsuneda and K. Hirao, J. Chem. Phys. **140**, 18A513 (2014).

45. J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz Jr, Phys. Rev. Lett. **49**, 1691 (1982).

46. W. Yang, Y. Zhang, and P. W. Ayers, Phys. Rev. Lett. **84**, 5172 (2000).

47. L. N. Anderson, M. B. Oviedo, and B. M. Wong, J. Chem. Theory Comput. **13**, 1656 (2017).

48. A. E. Raeber and B. M. Wong, J. Chem. Theory Comput. **11**, 2199 (2015).

49. M. E. Foster and B. M. Wong, J. Chem. Theory Comput. **8**, 2682 (2012).

50. B. M. Wong and T. H. Hsieh, J. Chem. Theory Comput. **6**, 3704 (2010).

51. B. M. Wong, M. Piacenza, and F. Della Sala, Phys. Chem. Chem. Phys. **11**, 4498 (2009).

52. B. M. Wong and J. G. Cordaro, J. Chem. Phys. **129**, 214703 (2008).

53. D. Bresteau, C. Drag, and C. Blondel, Phys. Rev. A **93**, 013414 (2016).