



PCCP

Local self-interaction correction method with a simple scaling factor

Journal:	<i>Physical Chemistry Chemical Physics</i>
Manuscript ID	CP-ART-12-2020-006282.R1
Article Type:	Paper
Date Submitted by the Author:	23-Dec-2020
Complete List of Authors:	Romero, Selim; The University of Texas at El Paso Yamamoto, Yoh; University of Texas at El Paso, Physics Baruah, Tunna; University of Texas at El Paso, Physics Zope, Rajendra; University of Texas at El Paso, Physics

SCHOLARONE™
Manuscripts

Cite this: DOI: 00.0000/xxxxxxxxxx

Local self-interaction correction method with a simple scaling factor

Selim Romero,^{ab} Yoh Yamamoto,^a Tunna Baruah,^{ab} and Rajendra R. Zope^{ab*}Received Date
Accepted Date

DOI: 00.0000/xxxxxxxxxx

A recently proposed local self-interaction correction (LSIC) method [Zope *et al.* J. Chem. Phys., 2019, 151, 214108] when applied to the simplest local density approximation provides a significant improvement over standard Perdew-Zunger SIC (PZSIC) for both equilibrium properties such as total or atomization energies as well as properties involving stretched bond such as barrier heights. The method uses an iso-orbital indicator to identify the single-electron regions. To demonstrate the LSIC method, Zope *et al.* used the ratio z_σ of von Weizsäcker τ_σ^W and total kinetic energy densities τ_σ , ($z_\sigma = \tau_\sigma^W / \tau_\sigma$) as a scaling factor to scale the self-interaction correction. The present work further explores the LSIC method using a simpler scaling factor as a ratio of orbital and spin densities in place of the ratio of kinetic energy densities. We compute a wide array of both, equilibrium and non-equilibrium properties using LSIC and orbital scaling methods using this simple scaling factor and compare them with previously reported results. Our study shows that LSIC with the simple scaling factor performs better than PZSIC, with results comparable to those obtained by LSIC(z_σ) for most properties, but has slightly larger errors than LSIC(z_σ). Furthermore, we study the binding energies of small water clusters using both scaling factors. Our results show that LSIC with z_σ has limitations in predicting the cluster binding energies of weakly bonded systems due to the inability of z_σ to distinguish weakly bonded regions from slowly varying density regions. LSIC when used with the density ratio as a scaling factor, on the other hand, provides a good description of water cluster binding energies, thus highlighting the appropriate choice of the iso-orbital indicator.

1 Introduction

The Kohn-Sham (KS) formulation of density functional theory (DFT)^{1–3} is widely used to study electronic structures of atoms, molecules, and solids because of its low computational cost and its availability in easy to use software packages. The practical application of DFT requires an approximation to the exchange-correlation (XC) functional. The simplest form of the XC functional is the local spin density approximation (LSDA)^{1,4} which belongs to the lowest rung of Jacob's ladder of XC functionals⁵. The higher rungs of the ladder contain more complex and more accurate functionals—generalized gradient approximation (GGA), meta-GGA, hybrid functionals, and functionals that include virtual orbitals. The majority of the density functional approximations (DFAs) suffer from self-interaction errors (SIE) though the magnitude of error can vary from one class of functionals to another or from one parameterization to another in a given class of functionals. The SIE occurs as a result of incomplete cancel-

lation of self-Coulomb energy by the self-exchange energy of the approximate XC functional.

Many failures of DFAs have been attributed to SIE. SIE causes the potential to decay asymptotically as $-\exp(-r)$ instead of the correct $-1/r$ decay for finite neutral systems. As a result, DFAs produce errors such as overly shallow eigenvalues of valence orbitals, inaccurate chemical reaction barriers, electron delocalization errors, incorrect charges on dissociated fragments, incorrect binding energies for anions, etc.^{4,6–8} The $-1/r$ asymptotic behavior is also important for the computation of electronic properties that are sensitive to virtual orbitals and long-range density such as excited states, for example.

Several approaches to remove SIEs have been proposed.^{9–20} Early approaches^{9,10} used orbitalwise schemes to eliminate SIE but used functionals related to Slater's $X\alpha$ method²¹. The most widely used approach to remove SIE is the one proposed by Perdew and Zunger (PZ)⁴. Their approach is commonly referred to as PZ self-interaction correction (PZSIC) where the one-electron SIE due to both exchange and correlation are removed from a DFA calculation on an orbital-by-orbital basis. PZSIC provides exact cancellation for one-electron self-interaction (SI),

^aDepartment of Physics, University of Texas at El Paso, El Paso, Texas, 79968, USA

^bComputational Science Program, University of Texas at El Paso, El Paso, Texas, 79968, USA

*rzope@utep.edu

but not necessarily for many-electron SI²². It has been applied to study properties of atoms, molecules, clusters, and solids.^{4,8,12,15,16,23–54,54–101}

PZSIC is an orbital dependent theory and when used with KS orbitals results in the size-extensivity problem. In PZSIC, local orbitals are used to keep the corrections size-extensive. Traditionally, PZSIC requires solving the so-called Pederson or localization equations (LE)^{23,25} to find the set of local orbitals that minimizes the total energy. Solving the LE and finding the optimal orbitals compliant with these conditions is computationally expensive since it requires solving the LE for each pair of orbitals. Pederson *et al.* in 2014 used Fermi-Löwdin orbitals^{102,103} (FLOs) to solve the PZSIC equations. This approach is known as FLO-SIC^{104,105}. FLOs are a Löwdin orthogonalized set of Fermi orbitals (FOs)^{102,103} that can be obtained from the KS orbitals. The FOs depend on the density matrix and spin density. The FLOs are the local orbitals that make FLOSIC total energy unitarily invariant. For the construction of FLOs, Fermi orbital descriptor (FOD) positions are used as $3N$ parameters in space that can be optimized analogously to the optimization of atomic positions in molecular structure optimization. Unlike the traditional PZSIC implementation which requires the determination of N^2 parameters, the FLOSIC method requires the determination of only $3N$ parameters.

Earlier applications of FLO-SIC with LSDA showed significant improvements in atomic and molecular properties over SI-uncorrected LSDA performance^{51,74,106,107}. Naturally, FLOSIC was later also applied to more sophisticated XC functionals, such as Perdew–Burke–Ernzerhof (PBE) and Strongly Constrained and Appropriately Normed (SCAN), to see whether SIC improves the performance of those functionals in the higher rungs^{17,52,66,67,69,70,75–78,108–113}. PZSIC when applied to semi-local functionals such as PBE, GGA, and SCAN meta-GGA provides a good description of stretched bond situations and predicts bound atomic anions, but this improvement occurs at the expense of worsening^{36,109,110,114–116} the performance for properties where SI-uncorrected DFA performs well. Shahi *et al.*¹⁰⁹ recently attributed the poor performance of PZSIC with GGAs and higher rung functionals to the nodality of the local orbital densities. The use of complex localized orbitals with nodeless densities in PZSIC calculations by Klüpfel, Klüpfel and Jónsson¹¹⁵ and by Lethola *et al.*¹¹⁷ show that complex orbital densities ameliorate the worsening of atomization energies when used with PBE. This conflicting performance of PZSIC is called the paradox of SIC by Perdew and coworkers¹¹⁸. The worsening of energetics pertaining to equilibrium regions primarily is a result of the overcorrecting tendency of PZSIC. A few methods have been proposed to mitigate this overcorrection by scaling down the SIC contribution. Jónsson's group simply scaled the SIC by a constant scaling factor¹¹⁴. In a similar spirit, Vydrov *et al.* proposed a method to scale down the SIC according to an orbital dependent scaling factor (OSIC)³⁸. This method however does not provide a significant improvement for all properties. It improved over PZSIC atomization energies but worsened barrier heights. Moreover, the scaling approach by Vydrov *et al.* results in a worsening of the asymptotic description of the effective potential, causing atomic

anions to be unbound. Ruzsinszky *et al.*¹¹⁹ found that many-electron SIE and fractional-charge dissociation behavior of positively charged dimers reappear in the OSIC of Vydrov *et al.* Yamamoto and coworkers¹¹² implemented a new 'selective OSIC' method (SOSIC), that selectively scales down the SIC in many-electron regions. SOSIC overcomes the deficiencies of the OSIC method and predicts stable atomic anions as well as improved total atomic energies. It also improves the barrier heights over the OSIC method. Very recently, Zope *et al.*¹⁷ proposed a new SIC method which identifies single-electron regions using iso-orbital indicators and corrects for SIE in a pointwise fashion by scaling down the SIC. The iso-orbital indicator serves as a weight in numerical integration and identifies both the single-orbital regions where full correction is needed and the uniform density regions where the DFAs are already accurate and correction is not needed. They called the new SIC method local-SIC (LSIC)¹⁷ and assessed its performance for a wide array of properties using LSDA. Unlike PZSIC, LSIC performed remarkably for both equilibrium properties like atomization energies and stretched bond situations that occur in barrier height calculation.

The LSIC method makes use of an iso-orbital indicator to identify one-electron regions. It offers an additional degree of freedom in that a suitable iso-orbital can be used or designed to identify one-electron regions or tune the SIC contribution in a pointwise manner. In the original LSIC work, Zope *et al.* used a ratio of von Weizsäcker and total kinetic energy densities as a choice for the local scaling factor. This iso-orbital indicator has been used in the construction of self-correlation free meta-GGAs, in the regional SIC¹⁵ and also in local hybrid functionals^{120,121}. Several different choices for the local scaling factors are already available in the literature. Alternatively, new iso-orbital indicators particularly suitable for LSIC can be constructed. In this work, we explore the performance of the LSIC method using a simple ratio of the orbital density and spin density as the weight of SIC correction at a given point in space. This is the same scaling factor used by Slater to average the Hartree-Fock exchange potential in his classic work that introduced the Hartree-Fock-Slater method²¹. We refer to this choice of scaling factor as LSIC(w) for the remainder of this manuscript and use LSIC(z) to refer to the first LSIC application where the scaling factor is the ratio of von Weizsäcker kinetic energy and kinetic energy densities. We investigate the performance of LSIC(w) for a few atomic properties: total energy, ionization potentials, and electron affinities. For molecules, we calculate the total energies, atomization energies, and dissociation energies of a few selected systems. We find that LSIC(w) provides comparable results to LSIC(z). We also show a case where LSIC(w) performs better than the original LSIC(z). Additionally, we examine the performance of the scaling factor w based on the density ratio with the OSIC scheme.

In the following section, brief descriptions of the PZSIC, OSIC, and LSIC methods are presented. These methods are implemented using FLOs. Therefore, very brief definitions pertaining to FLOs are also presented. The results and discussion are presented in the next sections.

2 Theory and computational method

2.1 Perdew-Zunger and Fermi-Löwdin Self-Interaction Correction

In the PZSIC method⁴, SIE is removed on an orbital-by-orbital basis from the DFA energy as

$$E^{PZSIC-DFA} = E^{DFA}[\rho_{\uparrow}, \rho_{\downarrow}] - \sum_{i\sigma}^{occ} \left\{ U[\rho_{i\sigma}] + E_{XC}^{DFA}[\rho_{i\sigma}, 0] \right\}, \quad (1)$$

where i is the orbital index, σ is the spin index, ρ ($\rho_{i\sigma}$) is the electron density (local orbital density), $U[\rho_{i\sigma}]$ is the exact self-Coulomb energy and $E_{XC}^{DFA}[\rho_{i\sigma}, 0]$ is the self-exchange-correlation energy for a given DFA XC functional. Perdew and Zunger applied this scheme to atoms using KS orbitals. For larger systems, KS orbitals can be delocalized which would result in a violation of size extensivity. Therefore, local orbitals are required. This was recognized long ago by Slater and Wood¹²² in 1971 and was also emphasized by Gopinathan¹⁰ in the context of self-interaction-correction of the Hartree-Slater method and later by Perdew and Zunger in the context of approximate Kohn-Sham calculations. Subsequent PZSIC calculations by the Wisconsin group^{23,24,24,25,80} used local orbitals in a variational implementation. It was shown by Pederson and coworkers^{23,25} that local orbitals used in Eq. (1) must satisfy LEs for variational minimization of energy. The LE for the orbitals $\phi_{i\sigma}$ is a pairwise condition and is given as

$$\langle \phi_{i\sigma} | V_{i\sigma}^{SIC} - V_{j\sigma}^{SIC} | \phi_{j\sigma} \rangle = 0. \quad (2)$$

These equations are sometimes called symmetry conditions⁵⁹. In the FLOSIC approach, FLOs are used instead of directly solving Eq. (2). First, FOs ϕ^{FO} are constructed with the density matrix and spin density at special positions in space called Fermi orbital descriptor (FOD) positions as

$$\phi_i^{FO}(\vec{r}) = \frac{\sum_j^{N_{occ}} \psi_j(\vec{a}_i) \psi_j(\vec{r})}{\sqrt{\rho_i(\vec{a}_i)}}. \quad (3)$$

Here i and j are the orbital indexes, ψ is the KS orbital, ρ_i is the electron spin density and \vec{a}_i is the FOD position. The FOs are then orthogonalized with Löwdin's scheme to form the FLOs. The FLOs are used for the calculation of the SIC energy and potential. In this method, the optimal set of FLOs is obtained by minimizing the total SIC energy by adjusting the corresponding FODs. The optimization of FOD positions is similar to the optimization of atomic(ionic) positions in standard structure optimization. We emphasize that FLOs can be used in all three SIC (PZSIC, OSIC, and LSIC) methods.

2.2 Orbitalwise scaling of SIC

As mentioned in Sec. 1, PZSIC tends to overcorrect the DFA energies, and several modifications to PZSIC were proposed to *scale down* the PZSIC correction. In the OSIC method of Vydrov *et al.*³⁸

mentioned in the Introduction, Eq. (1) is modified to

$$E^{OSIC-DFA} = E_{XC}^{DFA}[\rho_{\uparrow}, \rho_{\downarrow}] - \sum_{i\sigma}^{occ} X_{i\sigma}^k \left(U[\rho_{i\sigma}] + E_{XC}^{DFA}[\rho_{i\sigma}, 0] \right), \quad (4)$$

where each local orbitalwise scaling factor $X_{i\sigma}^k$ is defined as

$$X_{i\sigma}^k = \int z_{\sigma}^k(\vec{r}) \rho_{i\sigma}(\vec{r}) d^3\vec{r}. \quad (5)$$

Here, i indicates the orbital index, σ is the spin index, z_{σ} is the iso-orbital indicator and k is an integer. The quantity z_{σ} is used to interpolate the single-electron regions ($z_{\sigma} = 1$) and uniform density regions ($z_{\sigma} = 0$). In their original work, Vydrov *et al.* used $z_{\sigma} = \tau_{\sigma}^W / \tau_{\sigma}$ to study the performance of OSIC with various XC functionals, where $\tau_{\sigma}^W(\vec{r}) = |\vec{\nabla} \rho_{\sigma}(\vec{r})|^2 / (8\rho_{\sigma}(\vec{r}))$ is the von Weizsäcker kinetic energy density and $\tau_{\sigma}(\vec{r}) = \frac{1}{2} \sum_i |\vec{\nabla} \psi_{i\sigma}(\vec{r})|^2$ is the non-interacting kinetic energy density. Satisfying the gradient expansion in ρ requires $k \geq 1$ for LSDA, $k \geq 2$ for GGAs, and $k \geq 3$ for meta-GGA. Vydrov *et al.*, however, used various values of k to study its effect on the OSIC performance.

In their subsequent work, Vydrov and Scuseria¹⁶ used

$$w_{i\sigma}^k(\vec{r}) = \left(\frac{\rho_{i\sigma}(\vec{r})}{\rho_{\sigma}(\vec{r})} \right)^k, \quad (6)$$

the weight used by Slater²¹ in averaging Hartree-Fock potential, as a scaling factor instead of kinetic energy ratio. They repeated the OSIC calculations using $w_{i\sigma}$ in place of z_{σ} in Eq. (5). Notice that Eq. (6) contains a local orbital index, this weight is thus an orbital dependent quantity. $w_{i\sigma}$ approaches unity at single orbital regions since $\rho_{\sigma}(\vec{r}) = \rho_{i\sigma}(\vec{r})$ at this limit. Similarly, $w_{i\sigma}$ approaches zero at the many-electron region since $\rho_{\sigma}(\vec{r}) \gg \rho_{i\sigma}(\vec{r})$ at this condition. It was reported by Vydrov and Scuseria¹⁶ that OSIC with Eq. (6) showed comparable performance as $z_{\sigma} = \tau_{\sigma}^W / \tau_{\sigma}$ despite its simpler form.

2.3 LSIC

Though OSIC had some success in improving the performance with SIC, the approach leads to parameter k dependent performance. Also, it gives $-X_{HO}/r$ asymptotic potential instead of $-1/r$ for finite neutral systems and results in an inaccurate description of dissociation behavior²². In addition, many-electron SIE and fractional-charge dissociation behavior of positively charged dimers reemerge with OSIC¹¹⁹. The recent LSIC method by Zope *et al.*¹²³ applies the SIC in a different way than OSIC³⁸ and PZSIC⁴. It retains desirable beneficial features of PZSIC. In LSIC, the SIC energy density is scaled down *locally* as follows,

$$E_{XC}^{LSIC-DFA} = E_{XC}^{DFA}[\rho_{\uparrow}, \rho_{\downarrow}] - \sum_{i\sigma}^{occ} \left(U^{LSIC}[\rho_{i\sigma}] + E_{XC}^{LSIC}[\rho_{i\sigma}, 0] \right), \quad (7)$$

where

$$U^{LSIC}[\rho_{i\sigma}] = \frac{1}{2} \int d^3\vec{r} z_{\sigma}(\vec{r})^k \rho_{i\sigma}(\vec{r}) \int d^3\vec{r}' \frac{\rho_{i\sigma}(\vec{r}')}{|\vec{r}-\vec{r}'|}, \quad (8)$$

$$E_{XC}^{LSIC}[\rho_{i\sigma}, 0] = \int d^3\vec{r} z_{\sigma}(\vec{r})^k \rho_{i\sigma}(\vec{r}) \epsilon_{XC}^{DFA}([\rho_{i\sigma}, 0], \vec{r}). \quad (9)$$

LSIC uses an iso-orbital indicator to apply SIC pointwise in space. An ideal choice of the iso-orbital indicator should be such that LSIC reduces to DFA in the uniform gas limit and reduces to PZSIC in the pure one-electron limit. To demonstrate the LSIC concept, Zope *et al.*¹⁷ used $z_{\sigma} = \tau_{\sigma}^W / \tau_{\sigma}$ as an iso-orbital indicator. In this study, however, we use $w_{i\sigma}(\vec{r}) = \rho_{i\sigma}(\vec{r}) / \rho_{\sigma}(\vec{r})$ in place of z_{σ} in Eqs. (8) and (9).

2.4 Computational details

All calculations were performed using the developmental version of the FLOSIC code^{107,124}, a software based on the UTEP-NRLMOL code. The PZSIC, OSIC, and LSIC methods using FLOs are implemented in this code. For brevity, hereafter we will refer FLO-PZSIC, FLO-OSIC, and FLO-LSIC calculations as PZSIC, OSIC, and LSIC. The FLOSIC/NRLMOL code uses Gaussian type orbitals¹²⁵. We used the NRLMOL default basis sets throughout our calculations. For calculations of atomic anions, long range s, p, and d single Gaussian orbitals are added to give a better description of the extended nature of anions. The exponents β of these added single Gaussians were obtained using the relation $\beta(N+1) = \beta(N)^2 / \beta(N-1)$, where N is the N -th exponent. The FLOSIC code uses a variational integration mesh¹²⁶ that provides accurate numerical integration.

In this work, we use LSIC with the PW92 LSDA functional¹²⁷. LSIC applied to LSDA is free from the gauge problem¹²³ unlike GGAs and meta-GGAs where a gauge transformation is needed since their XC potentials are not in the Hartree gauge. The geometries used in our calculations are the same as in the respective databases and no further optimizations were performed. We used the SCF energy convergence criteria of $10^{-6} E_h$ for the total energy and an FOD force tolerance of $10^{-3} E_h/\text{bohr}$ for FOD optimizations^{105,128} in FLOSIC calculations. The self-consistency in the PZSIC calculations is obtained using Jacobi-like iterative procedure¹²⁹. For OSIC and LSIC calculations, we used the respective PZSIC densities and FODs as a starting point and performed one-shot (non-self-consistent) calculation of energy on the PZSIC densities. The self-interaction correction effects are included in these densities at the PZSIC level. To assess the effect of self-consistency we have calculated the atomization energies, barrier heights, etc. by performing self-consistent OSIC(w) calculations. We summarize these results in Table S1 of the supplementary information. The results from the table show that mean absolute errors in total energies of atoms, atomization energies, and barrier heights in the OSIC($w, k=1$) method changed only slightly (0.1–0.7 kcal/mol) after self-consistency. We expect similar differences in the LSIC method and therefore think that the present work provides the general picture of the performance of LSIC(w) method. The present results can alternatively be viewed as the performance of LSIC(w) energy functional in the spirit of density

corrected DFT by Kim, Sim, and Burke¹³⁰. Several values for the scaling power k are used in the LSIC(w) and OSIC(w) calculations. The additional computational cost of the scaling factor in OSIC and LSIC is small compared to standard PZSIC calculation.

3 Results and discussion

We assessed the performance of LSIC(w) vis-a-vis LSIC(z) and OSIC(w) using the wide array of electronic properties. We consider total energies, ionization potentials, and electron affinities for atoms and atomization energies, reaction barrier heights and dissociation energies for molecules.

3.1 Atoms

In this section, we present our results for total energies, ionization potentials, and electron affinities for atoms.

3.1.1 Total energy of atoms

We compared the total atomic energies of the atoms $Z = 1 - 18$ against accurate non-relativistic values reported by Chakravorty *et al.*¹³¹. Various integer values of k were used for LSIC(w) and OSIC(w). The differences between our calculated total energies with $k = 1$ and the reference values are plotted in Fig. 1. The plot clearly shows the effect of scaling on the total energies of atoms. Consistent with reported results, the LSDA total energies are too high compared to accurate reference values¹³¹ whereas PZSIC consistently underestimates the total energies due to its over correcting tendency. The LSIC method, where both scaling factors perform similarly, provides the total energies closer to the reference values than LSDA and the PZSIC. Likewise, the OSIC method also reduces the overcorrection, bringing the total energies into close agreement with the reference values. The mean absolute errors (MAEs) in total energy with respect to the reference for various k values are shown in Table 1. The MAE of PZSIC is $0.381 E_h$ whereas LSIC(w) and OSIC(w) show MAEs of 0.061 and $0.074 E_h$, respectively, with $k = 1$. LSIC(z) shows a better performance than OSIC(w) and LSIC(w). The LSIC(w) MAE is in the same order of magnitude as the earlier reported MAE of LSIC(z) of $0.041 E_h$ ¹⁷. As the value of k increases, the magnitude of SI-correction is reduced. This results in MAEs become larger for $k > 1$ eventually approaching the LSDA numbers.

For $k = 0$ the scaled methods correctly produce the PZSIC results. The scaling is optimal for $k = 1$ which results in an optimal magnitude of SI-correction for LSIC(w) and an almost right magnitude for OSIC(w). The magnitude of the SIC energy of each orbital is compared among different methods. It is found that the SIC correction in LSIC(w) is larger (i.e. less scaling down) for the core orbitals than in LSIC(z). This trend is reversed for the valence orbitals (cf. Table 2). It can be seen from Table 2 that total SIC energy in both methods is essentially similar in magnitude. However, the way scaling factors behave affects the orbitalwise contribution to the total SIC energy. This changes the SIC potentials and results in the two methods performing differently for cations and anions. For OSIC(w), we find the smallest MAE for $k = 2$ of $0.070 E_h$, a value slightly smaller than that for $k = 1$.

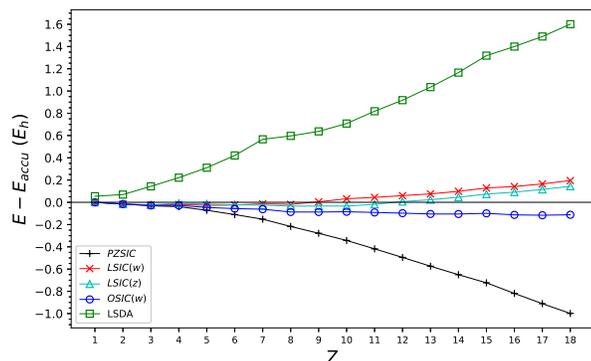


Fig. 1 Total energy difference (in E_h) of atoms $Z = 1 - 18$ with respect to accurate nonrelativistic estimates¹³¹.

Table 1 Mean absolute error of the total atomic energy (in E_h) for atoms $Z = 1 - 18$ with respect to accurate nonrelativistic estimates¹³¹.

Method	MAE
PZSIC	0.381
LSIC($z, k = 1$)	0.041
LSIC($w, k = 1$)	0.061
LSIC($w, k = 2$)	0.196
LSIC($w, k = 3$)	0.277
LSIC($w, k = 4$)	0.332
OSIC($w, k = 1$)	0.074
OSIC($w, k = 2$)	0.070
OSIC($w, k = 3$)	0.135

Table 2 Magnitude of SIC energy (in E_h) per orbital type in Ar atom for each method with $k = 1$.

Orbital	PZSIC	LSIC(z)	LSIC(w)	OSIC(w)
1s	-0.741	-0.387	-0.490	-0.584
2sp ³	-0.126	-0.070	-0.050	-0.062
3sp ³	-0.016	-0.017	-0.006	-0.008
Total SIC	-2.616	-1.473	-1.421	-1.729

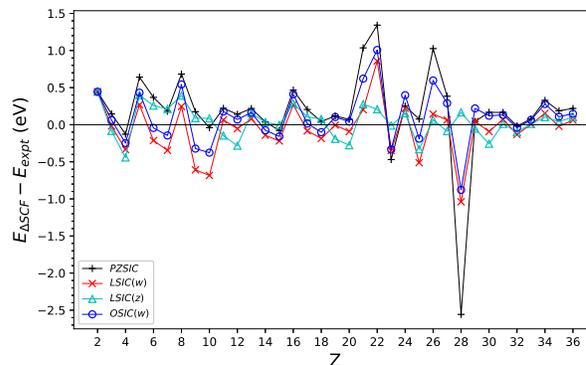


Fig. 2 Energy difference in ionization potential (in eV) for a set of atoms $Z = 2 - 36$ with respect to experimental values¹³².

3.1.2 Ionization potential

The ionization potential (IP) is the energy required to remove an electron from the outermost orbital. Since electron removal energy is related to the asymptotic shape of the potential, one can expect SIC plays an important role in determining IPs. We calculated the IPs using the Δ SCF method defined as

$$E_{IP} = E_{cat} - E_{neut} \quad (10)$$

where E_{cat} is the total energy in the cationic state and E_{neut} is the total energy at the neutral state. The calculations were performed for atoms from helium to krypton, and we compared the computed IPs against experimental ionization energies¹³². FODs were relaxed both for neutral atoms and for their cations. Fig. 2 shows the difference of calculated IPs with respect to the reference values. MAEs with different methods are shown in Table 3 for a subset $Z = 2 - 18$ as well as for the entire set $Z = 2 - 36$ to facilitate comparison against literature. For the smaller subset, $Z = 2 - 18$, the MAEs are 0.248 and 0.206 eV for PZSIC and LSIC(z), respectively. The MAE for OSIC($w, k = 1$) is 0.223 eV, showing an improvement over PZSIC. LSIC($w, k = 1$) shows MAE of 0.251 eV, a comparable error with PZSIC. MAEs increase for LSIC($w, k \geq 2$) and OSIC($w, k \geq 2$) in comparison to their respective $k = 1$ MAEs. Interestingly, however, when we considered the entire set of atoms ($Z = 2 - 36$), LSIC(w) has MAEs of 0.238 and 0.216 eV for $k = 1$ and $k = 2$ respectively showing smaller errors than PZSIC (MAE, 0.364 eV), but LSIC(w) falls short of LSIC(z) which has the smallest error (MAE, 0.170 eV). For this case, OSIC($w, k = 1 - 3$) shows better performance than PZSIC but not as well as LSIC(w) for a given k . LSIC(z) performs better than both LSIC(w) and OSIC(w). The difference in performance between LSIC(z) and LSIC(w) implies that scaling of SIC for the cationic states is more sensitive to the choice of a local scaling factor than for the neutral atoms.

3.1.3 Electron affinity

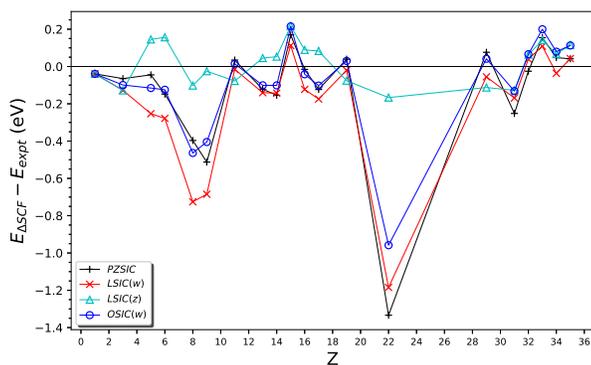
The electron affinity (EA) is the energy released when an electron is added to the system. We studied EAs for 20 atoms that are experimentally found to bind an electron¹³³. They are H, Li, B,

Table 3 Mean absolute error of ionization potentials (in eV) for set of atoms $Z = 2 - 18$ and $Z = 2 - 36$ with respect to experiment¹³².

Method	Z=2-18 (17-IPs)	Z=2-36 (35-IPs)
PZSIC	0.248	0.364
LSIC($z, k = 1$)	0.206	0.170
LSIC($w, k = 1$)	0.251	0.238
LSIC($w, k = 2$)	0.271	0.216
LSIC($w, k = 3$)	0.297	0.247
LSIC($w, k = 4$)	0.324	0.284
OSIC($w, k = 1$)	0.223	0.267
OSIC($w, k = 2$)	0.247	0.247
OSIC($w, k = 3$)	0.255	0.259

Table 4 Mean absolute error in electron affinities (in eV) for 12 EAs and 20 EAs set of atoms with respect to experiment¹³³.

Method	(12 EAs) MAE	(20 EAs) MAE
PZSIC	0.152	0.190
LSIC($z, k = 1$)	0.097	0.102
LSIC($w, k = 1$)	0.235	0.224
LSIC($w, k = 2$)	0.229	0.205
LSIC($w, k = 3$)	0.215	0.189
LSIC($w, k = 4$)	0.202	0.176
OSIC($w, k = 1$)	0.152	0.172
OSIC($w, k = 2$)	0.150	0.164
OSIC($w, k = 3$)	0.145	0.155

**Fig. 3** Electron affinity (eV) difference for atoms $Z = 2 - 36$ with respect to experiment¹³³.

C, O, F, Na, Al, Si, P, S, Cl, K, Ti, Cu, Ga, Ge, As, Se, and Br atoms. The EAs were calculated using the Δ SCF method $E_{EA} = E_{neut} - E_{anion}$ and values were compared against the experimental EAs¹³³.

Fig. 3 shows the deviation of EA from reference experimental values for various methods. The MAEs are summarized in Table 4. We have presented the MAEs in two sets, the smaller subset which contains hydrogen through chlorine (12 EAs) and for the complete set, hydrogen to bromine (20 EAs).

For 12 EAs, MAEs for PZSIC and LSIC(z) are 0.152 and 0.097 eV, respectively. OSIC(w) shows MAE of 0.152 eV for $k = 1$, the same performance as PZSIC. LSIC(w), however, does not perform as well as PZSIC, giving the MAEs of 0.235 eV for $k = 1$. In both cases, the error decreases slightly for $k \geq 2$ but there is no significant impact on their performance.

For 20 EAs, a similar trend persists. PZSIC and LSIC(z) have MAEs of 0.190 and 0.102 eV, respectively. The MAEs of LSIC(w) are in the range 0.176–0.224 eV for $k = 1 - 4$ and those of OSIC(w) are between 0.155–0.172 eV for $k = 1 - 3$. Again, a decrease in error is observed as k increases. In particular, a larger discrepancy between LSIC($w, k = 1$) and experiment is seen for O, F, and Ti atoms. This is due to LSIC(w) raising the anion energies more than their neutral state energies.

3.2 Atomization energy

To study the performance of LSIC(w) for molecules, first, we calculated the atomization energies (AEs) of 37 selected molecules. Many of these molecules are a subset of the G2/97 test set¹³⁴. The 37 molecules include systems from the AE6 set¹³⁵, a small but good representative subset of the main group atomization energy (MGAE109) set¹³⁶. The AEs were calculated by taking the energy difference of fragment atoms and the complex, that is, $AE = \sum_i^{N_{atom}} E_i - E_{mol} > 0$. E_i is the total energy of an atom, E_{mol} is the total energy of the molecule and N_{atom} is the number of atoms in the molecule. The calculated AEs were compared to the non-spin-orbit coupling reference values¹³⁶ for the AE6 set and to experimental values¹³³ for the entire set of 37 molecules. The percentage errors obtained through various methods are shown in Fig. 4. The overestimation of AEs with PZSIC-LSDA due to overcorrection is rectified in LSIC(w). We have summarized MAEs and mean absolute percentage errors (MAPEs) of AE6 and 37 molecules from the G2 set in Table 5. For the AE6 set, MAEs for PZSIC, LSIC(z), LSIC($w, k = 1$), and OSIC($w, k = 1$) are 57.9, 9.9, 13.8, and 33.7 kcal/mol respectively. The MAE in LSIC($w, k = 1$) is about 4 kcal/mol larger than LSIC(z) but substantially better than the PZSICs or OSIC(w). For the larger k in LSIC(w), however, the performance starts to degrade with a consistent increase in the MAE of 33.5 kcal/mol for $k = 4$. This is in contrast to OSIC where the performance improves for $k = 2$ and 3 compared to $k = 1$. The scaling thus differently affects the two methods. OSIC($w, k = 1$) tends to slightly underestimate total energies. By increasing k , total energies shift toward the LSDA total energies, and performance is improved for a moderate increase in k . On the contrary, total energies are slightly overestimated for LSIC($w, k = 1$), and increasing k makes the energies deviate away from the accurate estimates. OSIC($w, k = 3$) and LSIC($w, k = 1$) have a similar core orbital SIC energy. In their study of OSIC(w), Vydrov and Scuseria¹⁶ used values of k up to 5 and found the smallest error of $k = 5$ (MAE, 11.5 kcal/mol). But we expect the OSIC performance to degrade eventually for large k since an increase in k results in an increase in quenching of the SIC, thus the results will eventually approach those of DFA, in this LSDA case. For the full set of 37 molecules, PZSIC, LSIC(z), LSIC($w, k = 1$), and OSIC($w, k = 1$) show the MAPEs of 13.4, 6.9, 9.5 and 11.9%, respectively. OSIC(w) shows a slight improvement in MAPE for $k = 2$ and

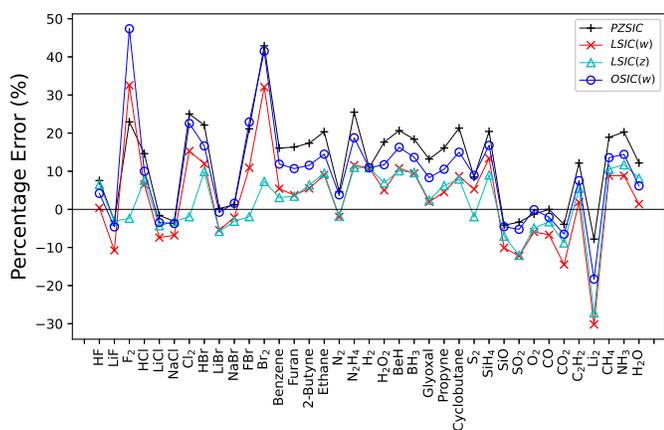


Fig. 4 Percentage errors of atomization energy (%) for a set of 37 molecules with respect to experimental values¹³³ using different scaling methods.

Table 5 Mean absolute error (in kcal/mol) and mean absolute percentage error (in %) of atomization energy for AE6 set of molecules¹³⁶ and for the set of 37 molecules from G2 set with respect to experiment¹³³.

Method	AE6 MAE (kcal/mol)	AE6 MAPE (%)	37 molecules MAPE (%)
PZSIC	57.9	9.4	13.4
LSIC($z, k = 1$)	9.9	3.2	6.9
LSIC($w, k = 1$)	13.8	3.9	9.5
LSIC($w, k = 2$)	18.6	4.8	9.1
LSIC($w, k = 3$)	26.9	5.8	9.2
LSIC($w, k = 4$)	33.5	6.7	9.7
OSIC($w, k = 1$)	33.7	6.3	11.9
OSIC($w, k = 2$)	24.1	5.1	11.3
OSIC($w, k = 3$)	17.8	4.3	10.9

3. For the larger set, LSIC(w) consistently shows smaller MAPEs than OSIC(w) for $k = 1 - 3$. All four values of k with the LSIC(w) in this study showed better performance than PZSIC for the 37 molecules set.

3.3 Barrier heights

An accurate description of chemical reaction barriers is challenging for DFAs since it involves the calculation of energies in non-equilibrium situations. In most cases, the saddle point energies are underestimated since DFAs do not perform well for a non-equilibrium state that involves a stretched bond. This shortcoming of DFAs in a stretched bond case arises from SIE; when an electron is shared and stretched out, SIE incorrectly lowers the energy of transition states. SIC handles the stretched bond states accurately and provides a correct picture of chemical reaction paths. We studied the reaction barriers using the BH6¹³⁵ set of molecules for the LSIC(w) method. BH6 is a representative subset of the larger BH24¹³⁷ set consisting of three reactions $\text{OH} + \text{CH}_4 \rightarrow \text{CH}_3 + \text{H}_2\text{O}$, $\text{H} + \text{OH} \rightarrow \text{H}_2 + \text{O}$, and $\text{H} + \text{H}_2\text{S} \rightarrow \text{H}_2 + \text{HS}$. We calculated the total energies of the left- and right-hand sides and at the saddle point of these chemical reactions. The

Table 6 Mean error (in kcal/mol) and mean absolute error (in kcal/mol) of BH6 sets of chemical reactions¹³⁵.

Method	ME (kcal/mol)	MAE (kcal/mol)
PZSIC	-4.8	4.8
LSIC($z, k = 1$)	0.7	1.3
LSIC($w, k = 1$)	-1.0	3.6
LSIC($w, k = 2$)	-0.1	4.6
LSIC($w, k = 3$)	0.3	5.0
LSIC($w, k = 4$)	0.6	5.5
OSIC($w, k = 1$)	-3.4	3.6
OSIC($w, k = 2$)	-3.1	4.1
OSIC($w, k = 3$)	-3.0	4.6

barrier heights for the forward (f) and reverse (r) reactions were obtained by taking the energy differences of their corresponding reaction states.

The mean errors (MEs) and MAEs of computed barrier heights are compared against the reference values¹³⁵ in Table 6. MAEs for PZSIC, LSIC(z), LSIC($w, k = 1$), and OSIC($w, k = 1$) are 4.8, 1.3, 3.6, and 3.6 kcal/mol, respectively. PZSIC significantly improves MAE compared to LSDA (MAE, 17.6 kcal/mol), LSIC($w, k = 1$) further reduces the error from PZSIC. Its ME and MAE indicate that there is no systematic underestimation or overestimation. LSIC($w, k = 1$) also further improves upon PZSIC, but not to the same extent as LSIC(z). For $k \geq 2$, MAEs increases systematically for LSIC($w, k \geq 2$) though small MEs are seen for LSIC($w, k = 2, 3$). The performance deteriorates for $k > 2$ beyond that of PZSIC. OSIC(w) shows marginally better performance than PZSIC. Vydrov and Scuseria¹⁶ showed that the best performance is achieved with $k = 1$ (MAE, 3.5 kcal/mol). The performance improvement with OSIC is not as dramatic as that with LSIC where the rather large MEs are seen with OSIC. Overall LSIC(w) performs better than OSIC(w) for barrier heights.

3.4 Dissociation and reaction energies

A pronounced manifestation of SIE is seen in the dissociation of positively charged dimers X_2^+ . SIE causes the system to dissociate into two fractionally charged cations instead of X and X^+ . Here we use the SIE4x4¹³⁸ and SIE11¹³⁹ sets to study the performance of LSIC(w) and OSIC(w) in correcting the SIEs. The SIE4x4 set consists of dissociation energy calculations of four positively charged dimers at varying bond distances R from their equilibrium distance R_e such that $R/R_e = 1.0, 1.25, 1.5, \text{ and } 1.75$. The dissociation energy E_D is calculated as

$$E_D = E(X) + E(X^+) - E(X_2^+). \quad (11)$$

The SIE11 set consists of eleven reaction energy calculations: five cationic reactions and six neutral reactions. These two sets are commonly used for studying SIE related problems. The calculated dissociation and reaction energies are compared against the CCSD(T) reference values^{138,139}, and MAEs are summarized in Table 7. For the SIE4x4 set, PZSIC, LSIC(z), LSIC($w, k = 1$) and OSIC($w, k = 1$) show MAEs of 3.0, 2.6, 4.7 and 5.2 kcal/mol.

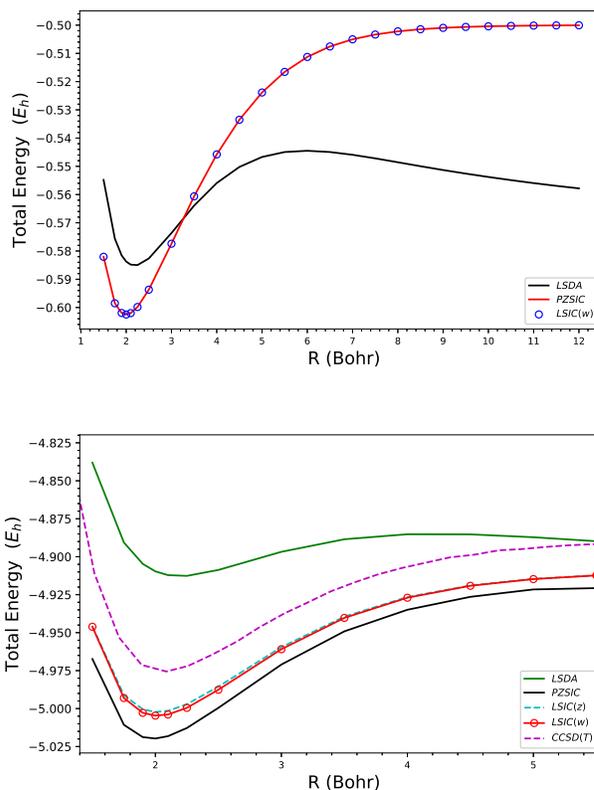
Table 7 Mean absolute error for dissociation and reaction energies (in kcal/mol) of SIE4x4 and SIE11 sets of chemical reactions with respect to CCSD(T)^{138,139}.

Reaction	SIE4x4	SIE11	SIE11	
			5 cationic	6 neutral
PZSIC	3.0	11.5	14.9	8.7
LSIC(z)	2.6	4.5	2.3	6.3
LSIC(w) ($k=1$)	4.7	8.3	8.6	8.0
LSIC(w) ($k=2$)	5.5	8.3	8.3	8.3
LSIC(w) ($k=3$)	5.8	8.8	8.2	9.3
LSIC(w) ($k=4$)	5.9	9.3	8.2	10.2
OSIC(w) ($k=1$)	5.2	11.1	13.7	9.0
OSIC(w) ($k=2$)	6.0	11.0	13.5	9.0
OSIC(w) ($k=3$)	6.4	10.9	13.3	8.8

LSIC(z) provides a small improvement in equilibrium energies while keeping the accurate behavior of PZSIC at the dissociation limit, resulting in marginally better performance. LSIC(w) shows errors a few kcal/mol larger than PZSIC. This increase in error arises because LSIC(w) alters the $(\text{NH}_3)_2^+$ and $(\text{H}_2\text{O})_2^+$ dissociation curves. In LSIC(z) the scaling of SIC occurs mostly for the core orbitals (Cf. Table 2), whereas LSIC(w) also includes some noticeable scaling down effects from valence orbitals. This different scaling behavior seems to contribute to different dissociation curves. Lastly, OSIC(w) has a slightly larger error than LSIC(w).

For the SIE11 set, MAEs are 11.5, 4.5, 8.3, and 11.1 kcal/mol for PZSIC, LSIC(z), LSIC($w, k=1$), and OSIC($w, k=1$), respectively. All scaled-down approaches we considered, LSIC(z) and LSIC(w) and OSIC(w) showed a performance improvement over PZSIC. LSIC(z) shows the largest error reduction by 60%, while LSIC($w, k=1$) shows 28% decrease in error with respect to PZSIC. OSIC(w) with $k=1-3$ has slightly smaller MAEs within 1 kcal/mol of PZSIC. The LSIC(z) method improves cationic reactions more than neutral reactions with respect to PZSIC. An increase in k beyond 2 results in over suppression of SIC and leads to an increase in error for LSIC($w, k \geq 2$). LSIC(w) yielded consistently smaller MAEs than OSIC(w), but larger than LSIC(z) for all SIE11 reactions.

Finally, we show the ground-state dissociation curves for H_2^+ and He_2^+ in Fig. 5. As previously discussed in literature¹⁴⁰, DFAs at large separation cause the complexes to dissociate into two ionic fragments. PZSIC restores the correct dissociation behavior at large separations. When LSIC is applied, the behavior of PZSIC at the dissociation limit is preserved in both LSIC(z) and present LSIC(w). For H_2^+ , a one-electron system, LSIC reproduces the identical behavior as PZSIC [Fig. 5 (a)]. For He_2^+ , a three-electron system, LSIC applies the correction to PZSIC only near-equilibrium regime [Fig. 5 (b)]. LSIC brings the equilibrium energy qualitatively closer to the CCSD(T) energy compared to PZSIC. The implication of Fig. 5 is that the scaling factor w performs well in differentiating the single-orbital-like regions and many-electron-like regions.

**Fig. 5** Dissociation curves of (a) H_2^+ and (b) He_2^+ using various methods. The CCSD(T) curve from Ref. [22] is plotted for comparison.

3.5 Water cluster binding energies: a case where LSIC(z) performs poorly

Sharkas *et al.*⁷⁸ recently studied binding energies of small water clusters using the PZSIC method in conjunction with FLOs to examine the effect of SIC on the binding energies of these systems. Water clusters are bonded by weaker hydrogen bonds and provide a different class of systems to test the performance of the LSIC method. Earlier studies using LSIC(z) on the water clusters of their polarizabilities and IPs have shown that LSIC(z) provides excellent descriptions of these properties when compared to CCSD(T) results^{67,141}. Here, we study the binding energies of the water clusters and find that the choice of iso-orbital indicator plays a critical role.

The structures considered in this work are $(\text{H}_2\text{O})_n$ ($n = 1 - 6$) whose geometries are from the WATER27 set¹⁴² optimized at the B3LYP/6-311++G(2d,2p) theory. The hexamer structure has a few known isomers, and we considered the book (b), cage (c), prism (p), and ring (r) isomers. The results are compared against the CCSD(T)-F12b values from Ref. 143 in Table 8. We obtained the MAEs of 118.9, 172.1, and 46.9 meV/H₂O for PZSIC, LSIC(z), and LSIC(w), respectively. Thus, LSIC(z) underestimates binding energies of water clusters by a roughly similar magnitude as LSDA (MAE, 183.4 meV/H₂O). This is one of few cases where LSIC(z) does not improve over PZSIC. A simple explanation for this behavior of LSIC(z) is that although z_σ can detect the weak bond regions, it cannot differentiate the slow-varying density regions from weak bond regions. $z_\sigma \rightarrow 0$ in both situations, causing the weak bond regions to be improperly treated. Fig. 6 (a) shows z_σ for the water dimer where both slow-varying density and weak interaction regions are detected but not differentiated. As a result, the total energies of the complex shift too much in comparison to the individual water molecules. Thus, the underestimation of water cluster binding energies is due to the choice of z and not necessarily of the LSIC method. Indeed, by choosing w as a scaling parameter, the binding energies are much improved. Fig. 6 (b) shows that there is no discontinuity of w between the two water molecules (w_i 's of two FLOs along the hydrogen bond are plotted together in the figure). Hence, unlike in LSIC(z), the weakly interacting region is not improperly scaled-down with LSIC(w). LSIC(w) shows an MAE of 46.9 meV/H₂O, comparable to SCAN (MAE, 35.2 meV/H₂O). This result is interesting as SCAN uses a function that can identify weak bond interaction. So LSIC(w)-LSDA may be behaving qualitatively similar to the detection function in SCAN in weak bond regions. The study of water cluster binding energies is one of few cases where the original LSIC(z) performed poorly. A recent study by Bhattarai *et al.*¹⁴⁴ examined the performance of LSIC(z) for interaction energies of non-covalently bonded complexes from the S22 data set and found that LSIC(z) performed poorly for the interaction energies of the weakly bonded complexes. A full understanding of how well LSIC performs for the binding energies of such complexes would require full self-consistency. The LSIC method may be improved by using a different iso-orbital indicator. These cases serve as motivation for identifying an appropriate iso-orbital indicator that would work for all bonding regions in LSIC.

Table 8 The binding energy of water clusters (in meV/H₂O).

n	PZSIC	LSIC(z)	LSIC(w)	CCSD(T) ^a
2	-153.7	-34.9	-82.7	-108.6
3	-321.6	-73.9	-183.0	-228.4
4	-425.2	-125.0	-248.6	-297.0
5	-446.9	-142.7	-264.8	-311.4
6b	-467.1	-133.6	-275.0	-327.3
6c	-466.8	-113.9	-274.8	-330.5
6p	-467.7	-104.8	-276.2	-332.4
6r	-458.1	-150.5	-275.5	-320.1
MAE	118.9	172.1	46.9	

^aReference [143]

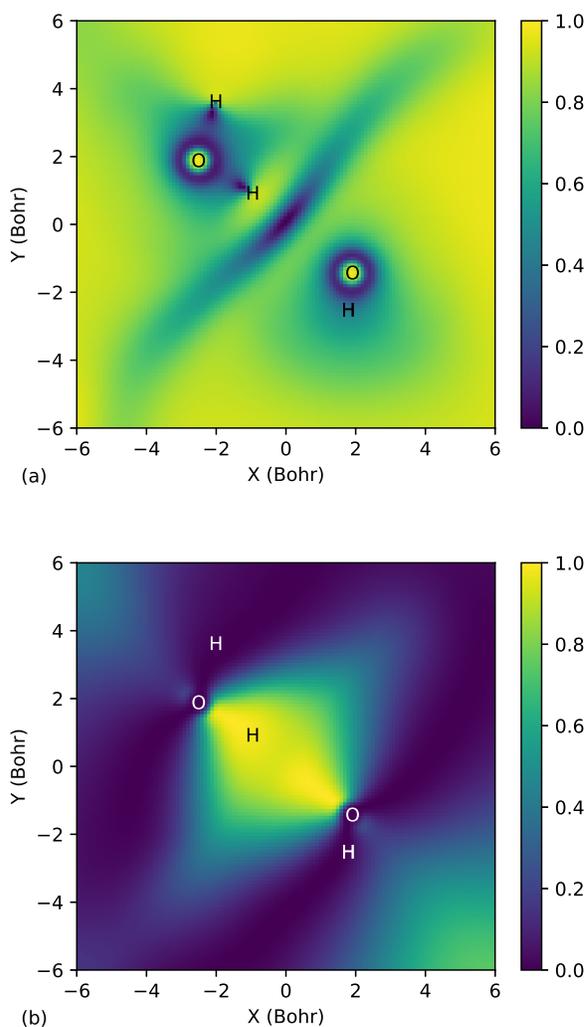


Fig. 6 Cross sectional plots of the iso-orbital indicators for water cluster dimer: (a) τ^w/τ and (b) ρ_i/ρ 's from the two FLOs along the hydrogen bond.

We now provide a qualitative explanation of why LSIC(w) gives improved results over PZSIC. This reasoning is along the same line as offered by Zope *et al.*¹⁷. As mentioned in Sec. 1, when SIE is removed using PZSIC, an improved description of barrier heights which involve stretched bonds is obtained, but the equilibrium properties like total energies, atomization energies, etc. are usually deteriorated compared to the uncorrected functional. This is especially so for the functionals that go beyond the simple LSDA. Typically this is because of the overcorrecting tendency of PZSIC. The non-empirical semilocal DFAs are designed to be exact in the uniform electron gas limit, this exact condition is no longer satisfied when PZSIC is applied to the functionals¹⁴⁵. This can be seen from the exchange energies of noble gas atoms and the extrapolation using the large- Z expansion of E_X as shown in Fig. 7. Following Ref. 145, we plot atomic exchange energies as a function of $Z^{-1/3}$. Thus, the region near the origin corresponds to the uniform gas limit. The plot was obtained by fitting the exchange energies (E_X) of Ne, Ar, Kr, and Xe atoms (within LSIC(w)-LSDA, LSIC(z)-LSDA, and LSDA) to the curve using the following fitting function¹⁴⁵.

$$\frac{E_X^{\text{approx}} - E_X^{\text{exact}}}{E_X^{\text{exact}}} \times 100\% = a + bx^2 + cx^3, \quad (12)$$

where $x = Z^{-1/3}$ and a , b , and c are the fitting parameters. The LSDA is exact in the uniform gas limit. So also is LSIC(z) since the scaling factor z_σ approaches zero as the gradient of electron density vanishes while the kinetic energy density in the denominator remains finite. The small deviation from zero seen near the origin (in Fig. 7) for LSIC(z) is due to the fitting error (due to limited data point). This error is -0.62% for LSIC(z). Thus, correcting LSDA using PZSIC introduces a large error in the uniform gas limit. The scaling factor w used here identifies single-electron regions since the density ratio approaches one in this limit. Fig. 7 shows that LSIC(w) also recovers the lost uniform gas limit. This partly explains the success of LSIC(w). Though the performance of LSIC(w) is substantially better than PZSIC-LSDA, it falls short of LSIC(z). On the other hand, unlike LSIC(z) it provides a good description of weak hydrogen bonds highlighting the need to identify suitable iso-orbital indicators or scaling factor(s) to apply pointwise SIC using the LSIC method. One possible choice may be scaling factors that are functions of α used in the construction of SCAN meta-GGA and the recently proposed¹⁴⁶ β parameter. A scaling factor containing β recently used by Yamamoto and coworkers with the OSIC scheme showed improved results¹¹². Future work would involve designing suitable scaling factors involving β for use in the LSIC method.

4 Conclusions

To recapitulate, we investigated the performance of LSIC with a simple scaling factor, w , that depends only on orbital and spin densities. Performance assessment has been carried out on atomic energies, atomization energies, ionization potentials, electron affinities, barrier heights, dissociation energies, and binding energies on standard data sets of molecules. The results show that LSIC(w) performs better than PZSIC for all properties with

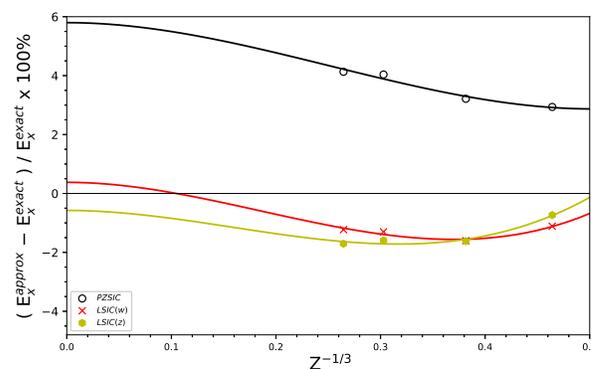


Fig. 7 Plot of percentage error of the approximated exchange energy compared to the exact exchange energy as a function of $Z^{-1/3}$.

exception of electron affinity and a SIE4x4 subset of dissociation energies. We also compared the performance of LSIC(w) against the OSIC method of Vydrov *et al.* Results indicate that although OSIC overall performs better than PZSIC, the improvement over PZSIC is somewhat limited. On the other hand, LSIC(w) is consistently better than OSIC(w). We have also studied the binding energies of small water clusters which are bonded by weak hydrogen bonds. Here, LSIC(w) performs very well compared to both PZSIC and LSIC(z), with performance comparable to SCAN. The present work shows the promise of the LSIC method and also demonstrates its limitation in describing weak hydrogen bonds if used with kinetic energy ratio, z_σ as an iso-orbital indicator. This limitation is due to the inability of z_σ to distinguish weak bonding regions from slowly varying density regions. The scaling factor w works differently than the scaling factor z , hence LSIC(w) provides a good description of weak hydrogen bonds in water clusters. The work thus highlights the importance of designing suitable iso-orbital indicators for use with LSIC that can detect weak bonding regions.

Data Availability Statement

The data that supports the findings of this study are available within the article and the supplementary information.

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgement

Authors acknowledge Drs. Luis Basurto, Carlos Diaz, Jorge Vargas, and Po-Hao Chang for discussions and technical supports. Authors also acknowledge Mr. Zachary Buschmann for proof-reading the manuscript. This work was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, as part of the Computational Chemical Sciences Program under Award No. DE-SC0018331. Support for the computational time at the Texas Advanced Computing Center through NSF Grant No. TG-DMR090071, and at NERSC is gratefully acknowledged.

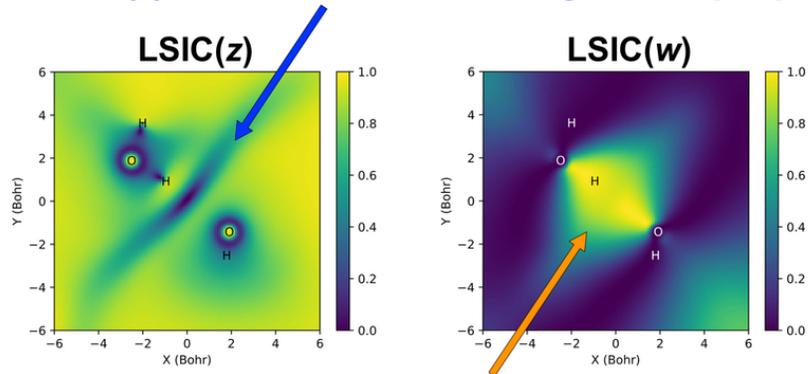
References

- 1 W. Kohn and L. Sham, *Phys. Rev.*, 1965, **140**, A1133–A1138.
- 2 R. O. Jones, *Rev. Mod. Phys.*, 2015, **87**, 897.
- 3 R. O. Jones and O. Gunnarsson, *Rev. Mod. Phys.*, 1989, **61**, 689–746.
- 4 J. P. Perdew and A. Zunger, *Phys. Rev. B*, 1981, **23**, 5048–5079.
- 5 J. P. Perdew and K. Schmidt, *AIP Conference Proceedings*, 2001, **577**, 1–20.
- 6 E. R. Johnson, P. Mori-Sánchez, A. J. Cohen and W. Yang, *The Journal of Chemical Physics*, 2008, **129**, 204112.
- 7 E. R. Johnson, A. Otero-de-la Roza and S. G. Dale, *The Journal of Chemical Physics*, 2013, **139**, 184116.
- 8 J. Gräfenstein, E. Kraka and D. Cremer, *J. Chem. Phys.*, 2004, **120**, 524–539.
- 9 I. Lindgren, *jqc*, 1971, **5**, 411–420.
- 10 M. S. Gopinathan, *Phys. Rev. A*, 1977, **15**, 2135–2142.
- 11 J. P. Perdew, R. G. Parr, M. Levy and J. L. Balduz Jr, *Physical Review Letters*, 1982, **49**, 1691.
- 12 U. Lundin and O. Eriksson, *jqc*, 2001, **81**, 247–252.
- 13 P. Mori-Sánchez, A. J. Cohen and W. Yang, *The Journal of Chemical Physics*, 2006, **125**, 201102.
- 14 N. I. Gidopoulos and N. N. Lathiotakis, *The Journal of Chemical Physics*, 2012, **136**, 224109.
- 15 T. Tsuneda, M. Kamiya and K. Hirao, *J. Comput. Chem.*, 2003, **24**, 1592–1598.
- 16 O. A. Vydrov and G. E. Scuseria, *J. Chem. Phys.*, 2006, **124**, 191101.
- 17 R. R. Zope, Y. Yamamoto, C. M. Diaz, T. Baruah, J. E. Peralta, K. A. Jackson, B. Santra and J. P. Perdew, *J. Chem. Phys.*, 2019, **151**, 214108.
- 18 D.-K. Seo, *Phys. Rev. B*, 2007, **76**, 033102.
- 19 G. Borghi, C.-H. Park, N. L. Nguyen, A. Ferretti and N. Marzari, *Physical Review B*, 2015, **91**, 155112.
- 20 T. Tsuneda and K. Hirao, *The Journal of Chemical Physics*, 2014, **140**, 18A513.
- 21 J. C. Slater, *Physical review*, 1951, **81**, 385.
- 22 A. Ruzsinszky, J. P. Perdew, G. I. Csonka, O. A. Vydrov and G. E. Scuseria, *jcp*, 2007, **126**, 104102.
- 23 M. R. Pederson, R. A. Heaton and C. C. Lin, *J. Chem. Phys.*, 1985, **82**, 2688–2699.
- 24 J. G. Harrison, R. A. Heaton and C. C. Lin, *Journal of Physics B: Atomic and Molecular Physics*, 1983, **16**, 2079–2091.
- 25 M. R. Pederson, R. A. Heaton and C. C. Lin, *J. Chem. Phys.*, 1984, **80**, 1972–1975.
- 26 J. Garza, J. A. Nichols and D. A. Dixon, *J. Chem. Phys.*, 2000, **112**, 7880–7890.
- 27 J. Garza, R. Vargas, J. A. Nichols and D. A. Dixon, *J. Chem. Phys.*, 2001, **114**, 639–651.
- 28 S. Patchkovskii, J. Autschbach and T. Ziegler, *J. Chem. Phys.*, 2001, **115**, 26–42.
- 29 M. K. Harbola, *Solid State Commun.*, 1996, **98**, 629–632.
- 30 S. Patchkovskii and T. Ziegler, *J. Chem. Phys.*, 2002, **116**, 7806–7813.
- 31 S. Patchkovskii and T. Ziegler, *J. Phys. Chem. A*, 2002, **106**, 1088–1099.
- 32 S. Goedecker and C. J. Umrigar, *Phys. Rev. A*, 1997, **55**, 1765–1771.
- 33 V. Polo, E. Kraka and D. Cremer, *Molecular Physics*, 2002, **100**, 1771–1790.
- 34 V. Polo, J. Gräfenstein, E. Kraka and D. Cremer, *Theor. Chem. Acc.*, 2003, **109**, 22–35.
- 35 J. Gräfenstein, E. Kraka and D. Cremer, *Phys. Chem. Chem. Phys.*, 2004, **6**, 1096–1112.
- 36 O. A. Vydrov and G. E. Scuseria, *J. Chem. Phys.*, 2004, **121**, 8187–8193.
- 37 O. A. Vydrov and G. E. Scuseria, *J. Chem. Phys.*, 2005, **122**, 184107.
- 38 O. A. Vydrov, G. E. Scuseria, J. P. Perdew, A. Ruzsinszky and G. I. Csonka, *J. Chem. Phys.*, 2006, **124**, 094108.
- 39 J. Chen, J. B. Krieger, Y. Li and G. J. Iafrate, *Phys. Rev. A*, 1996, **54**, 3939–3947.
- 40 R. R. Zope, M. K. Harbola and R. K. Pathak, *The European Physical Journal D-Atomic, Molecular, Optical and Plasma Physics*, 1999, **7**, 151–155.
- 41 R. R. Zope, *Physical Review A*, 2000, **62**, 064501.
- 42 E. S. Fois, J. I. Penman and P. A. Madden, *The Journal of chemical physics*, 1993, **98**, 6352–6360.
- 43 J. B. Krieger, Y. Li and G. J. Iafrate, *Phys. Rev. A*, 1992, **45**, 101–126.
- 44 J. B. Krieger, Y. Li and G. J. Iafrate, *Phys. Rev. A*, 1992, **46**, 5453–5458.
- 45 Y. Li, J. B. Krieger and G. J. Iafrate, *Phys. Rev. A*, 1993, **47**, 165–181.
- 46 S. Lehtola, M. Head-Gordon and H. Jónsson, *J. Chem. Theory Comput.*, 2016, **12**, 3195–3207.
- 47 G. I. Csonka and B. G. Johnson, *tca*, 1998, **99**, 158–165.
- 48 L. Petit, A. Svane, M. Lüders, Z. Szotek, G. Vaitheeswaran, V. Kanchana and W. Temmerman, *jpcm*, 2014, **26**, 274213.
- 49 S. Kümmel and L. Kronik, *Rev. Mod. Phys.*, 2008, **80**, 3.
- 50 T. Schmidt, E. Kraisler, L. Kronik and S. Kümmel, *pccp*, 2014, **16**, 14357–14367.
- 51 D.-y. Kao, M. Pederson, T. Hahn, T. Baruah, S. Liebing and J. Kortus, *Magnetochemistry*, 2017, **3**, 31.
- 52 S. Schwalbe, T. Hahn, S. Liebing, K. Trepte and J. Kortus, *J. Comput. Chem.*, 2018, **39**, 2463–2471.
- 53 H. Jónsson, K. Tsemekhman and E. J. Bylaska, Abstracts of Papers of the American Chemical Society, 2007, pp. 120–120.
- 54 M. M. Rieger and P. Vogl, *Phys. Rev. B*, 1995, **52**, 16567–16574.
- 55 W. Temmerman, A. Svane, Z. Szotek, H. Winter and S. Beiden, *Electronic Structure and Physical Properties of Solids*, Springer, 1999, pp. 286–312.
- 56 M. Daene, M. Lueders, A. Ernst, D. Ködderitzsch, W. M. Temmerman, Z. Szotek and W. Hergert, *jpcm*, 2009, **21**, 045604.

- 57 Z. Szotek, W. Temmerman and H. Winter, *Physica B: Condensed Matter*, 1991, **172**, 19–25.
- 58 J. Messud, P. M. Dinh, P.-G. Reinhard and E. Suraud, *prl*, 2008, **101**, 096404.
- 59 J. Messud, P. M. Dinh, P.-G. Reinhard and E. Suraud, *cpl*, 2008, **461**, 316–320.
- 60 M. Lundberg and P. E. M. Siegbahn, *jcp*, 2005, **122**, 224103.
- 61 C. D. Pemmaraju, T. Archer, D. Sánchez-Portal and S. Sanvito, *Phys. Rev. B*, 2007, **75**, 045101.
- 62 T. Körzdörfer, M. Mundt and S. Kümmel, *prl*, 2008, **100**, 133004.
- 63 T. Körzdörfer, S. Kümmel and M. Mundt, *jcp*, 2008, **129**, 014110.
- 64 I. Ciofini, C. Adamo and H. Chermette, *cp*, 2005, **309**, 67–76.
- 65 T. Baruah, R. R. Zope, A. Kshirsagar and R. K. Pathak, *Phys. Rev. A*, 1994, **50**, 2191–2196.
- 66 A. I. Johnson, K. P. K. Withanage, K. Sharkas, Y. Yamamoto, T. Baruah, R. R. Zope, J. E. Peralta and K. A. Jackson, *J. Chem. Phys.*, 2019, **151**, 174106.
- 67 J. Vargas, P. Ufondu, T. Baruah, Y. Yamamoto, K. A. Jackson and R. R. Zope, *Phys. Chem. Chem. Phys.*, 2020, **22**, 3789–3799.
- 68 N. Poilvert, G. Borghi, N. L. Nguyen, N. D. Keilbart, K. Wang and I. Dabo, *Advances in Atomic, Molecular, and Optical Physics*, Elsevier, 2015, vol. 64, pp. 105–127.
- 69 K. Treppe, S. Schwalbe, T. Hahn, J. Kortus, D.-Y. Kao, Y. Yamamoto, T. Baruah, R. R. Zope, K. P. K. Withanage, J. E. Peralta and K. A. Jackson, *J. Comput. Chem.*, 2019, **40**, 820–825.
- 70 K. P. K. Withanage, K. Treppe, J. E. Peralta, T. Baruah, R. Zope and K. A. Jackson, *J. Chem. Theory Comput.*, 2018, **14**, 4122–4128.
- 71 S. Lehtola and H. Jónsson, *J. Chem. Theory Comput.*, 2014, **10**, 5324–5337.
- 72 M. R. Pederson, T. Baruah, D.-y. Kao and L. Basurto, *J. Chem. Phys.*, 2016, **144**, 164117.
- 73 S. Schwalbe, L. Fiedler, T. Hahn, K. Treppe, J. Kraus and J. Kortus, *PyFLOSIC - Python based Fermi-Löwdin orbital self-interaction correction*, 2019.
- 74 D.-y. Kao, K. Withanage, T. Hahn, J. Batool, J. Kortus and K. Jackson, *J. Chem. Phys.*, 2017, **147**, 164107.
- 75 R. P. Joshi, K. Treppe, K. P. K. Withanage, K. Sharkas, Y. Yamamoto, L. Basurto, R. R. Zope, T. Baruah, K. A. Jackson and J. E. Peralta, *J. Chem. Phys.*, 2018, **149**, 164101.
- 76 K. Sharkas, L. Li, K. Treppe, K. P. K. Withanage, R. P. Joshi, R. R. Zope, T. Baruah, J. K. Johnson, K. A. Jackson and J. E. Peralta, *J. Phys. Chem. A*, 2018, **122**, 9307–9315.
- 77 K. A. Jackson, J. E. Peralta, R. P. Joshi, K. P. Withanage, K. Treppe, K. Sharkas and A. I. Johnson, *J. Phys. Conf. Ser.*, 2019, **1290**, 012002.
- 78 K. Sharkas, K. Wagle, B. Santra, S. Akter, R. R. Zope, T. Baruah, K. A. Jackson, J. P. Perdew and J. E. Peralta, *Proceedings of the National Academy of Sciences*, 2020, **117**, 11283–11288.
- 79 H. Gudmundsdóttir, E. O. Jónsson and H. Jónsson, *New Journal of Physics*, 2015, **17**, 083006.
- 80 R. A. Heaton, J. G. Harrison and C. C. Lin, *Phys. Rev. B*, 1983, **28**, 5992–6007.
- 81 R. A. Heaton and C. C. Lin, *Journal of Physics C: Solid State Physics*, 1984, **17**, 1853–1866.
- 82 A. Svane and O. Gunnarsson, *Europhysics Letters (EPL)*, 1988, **7**, 171–175.
- 83 A. Svane and O. Gunnarsson, *Phys. Rev. B*, 1988, **37**, 9919–9922.
- 84 S. C. Erwin and C. C. Lin, *Journal of Physics C: Solid State Physics*, 1988, **21**, 4285–4309.
- 85 Z. Szotek, W. Temmerman and H. Winter, *Solid State Communications*, 1990, **74**, 1031–1033.
- 86 Z. Szotek, W. Temmerman and H. Winter, *Physica B: Condensed Matter*, 1990, **165-166**, 275–276.
- 87 A. Svane and O. Gunnarsson, *Phys. Rev. Lett.*, 1990, **65**, 1148–1151.
- 88 A. Svane, *Phys. Rev. Lett.*, 1992, **68**, 1900–1903.
- 89 D. Vogel, P. Krüger and J. Pollmann, *Phys. Rev. B*, 1995, **52**, R14316–R14319.
- 90 M. Arai and T. Fujiwara, *Phys. Rev. B*, 1995, **51**, 1477–1489.
- 91 D. Vogel, P. Krüger and J. Pollmann, *Phys. Rev. B*, 1996, **54**, 5495–5511.
- 92 A. Svane, *Phys. Rev. B*, 1996, **53**, 4275–4286.
- 93 A. Svane, W. M. Temmerman, Z. Szotek, J. Lægsgaard and H. Winter, *International Journal of Quantum Chemistry*, 2000, **77**, 799–813.
- 94 A. Filippetti and N. A. Spaldin, *Phys. Rev. B*, 2003, **67**, 125109.
- 95 E. Bylaska, K. Tsemekhman and H. Jonsson, APS March Meeting Abstracts, 2004.
- 96 M. Lüders, A. Ernst, M. Däne, Z. Szotek, A. Svane, D. Ködderitzsch, W. Hergert, B. L. Györfy and W. M. Temmerman, *Phys. Rev. B*, 2005, **71**, 205109.
- 97 E. J. Bylaska, K. Tsemekhman and F. Gao, *Physica Scripta*, 2006, **T124**, 86–90.
- 98 B. Hourahine, S. Sanna, B. Aradi, C. Köhler, T. Niehaus and T. Frauenheim, *The Journal of Physical Chemistry A*, 2007, **111**, 5671–5677.
- 99 M. Stengel and N. A. Spaldin, *Phys. Rev. B*, 2008, **77**, 155106.
- 100 M. Däne, M. Lüders, A. Ernst, D. Ködderitzsch, W. M. Temmerman, Z. Szotek and W. Hergert, *Journal of Physics: Condensed Matter*, 2009, **21**, 045604.
- 101 N. L. Nguyen, N. Colonna, A. Ferretti and N. Marzari, *Phys. Rev. X*, 2018, **8**, 021051.
- 102 W. L. Luken and D. N. Beratan, *Theor. Chem. Acc.*, 1982, **61**, 265–281.
- 103 W. L. Luken and J. C. Culberson, *Theor. Chem. Acc.*, 1984, **66**, 279–293.
- 104 M. R. Pederson, A. Ruzsinszky and J. P. Perdew, *J. Chem.*

- Phys.*, 2014, **140**, 121103.
- 105 M. R. Pederson and T. Baruah, *Advances In Atomic, Molecular, and Optical Physics*, Academic Press, 2015, vol. 64, pp. 153–180.
- 106 R. P. Joshi, J. J. Phillips and J. E. Peralta, *Journal of Chemical Theory and Computation*, 2016, **12**, 1728–1734.
- 107 R. R. Zope, T. Baruah and K. A. Jackson, *FLOSIC 0.2*, <https://http://flosic.org/>, based on the NRLMOL code of M. R. Pederson.
- 108 K. P. K. Withanage, S. Akter, C. Shahi, R. P. Joshi, C. Diaz, Y. Yamamoto, R. Zope, T. Baruah, J. P. Perdew, J. E. Peralta and K. A. Jackson, *Phys. Rev. A*, 2019, **100**, 012505.
- 109 C. Shahi, P. Bhattarai, K. Wagle, B. Santra, S. Schwalbe, T. Hahn, J. Kortus, K. A. Jackson, J. E. Peralta, K. Treppe, S. Lehtola, N. K. Nepal, H. Myneni, B. Neupane, S. Adhikari, A. Ruzsinszky, Y. Yamamoto, T. Baruah, R. R. Zope and J. P. Perdew, *J. Chem. Phys.*, 2019, **150**, 174102.
- 110 Y. Yamamoto, C. M. Diaz, L. Basurto, K. A. Jackson, T. Baruah and R. R. Zope, *J. Chem. Phys.*, 2019, **151**, 154105.
- 111 M. R. Pederson and T. Baruah, unpublished.
- 112 Y. Yamamoto, S. Romero, T. Baruah and R. R. Zope, *The Journal of Chemical Physics*, 2020, **152**, 174112.
- 113 S. Schwalbe, L. Fiedler, J. Kraus, J. Kortus, K. Treppe and S. Lehtola, *The Journal of Chemical Physics*, 2020, **153**, 084104.
- 114 S. Klüpfel, P. Klüpfel and H. Jónsson, *J. Chem. Phys.*, 2012, **137**, 124102.
- 115 S. Klüpfel, P. Klüpfel and H. Jónsson, *Phys. Rev. A*, 2011, **84**, 050501.
- 116 E. Ö. Jónsson, S. Lehtola and H. Jónsson, *Procedia Computer Science*, 2015, **51**, 1858–1864.
- 117 S. Lehtola, E. Ö. Jónsson and H. Jónsson, *J. Chem. Theory Comput.*, 2016, **12**, 4296–4302.
- 118 J. P. Perdew, A. Ruzsinszky, J. Sun and M. R. Pederson, *Advances In Atomic, Molecular, and Optical Physics*, Academic Press, 2015, vol. 64, pp. 1–14.
- 119 A. Ruzsinszky, J. P. Perdew, G. I. Csonka, O. A. Vydrov and G. E. Scuseria, *J. Chem. Phys.*, 2006, **125**, 194112.
- 120 J. Jaramillo, G. E. Scuseria and M. Ernzerhof, *jcp*, 2003, **118**, 1068–1073.
- 121 T. Schmidt, E. Kraisler, A. Makmal, L. Kronik and S. Kummel, *The Journal of Chemical Physics*, 2014, **140**, 18A510.
- 122 J. C. Slater and J. H. Wood, *International Journal of Quantum Chemistry*, 1970, **5**, 3–34.
- 123 P. Bhattarai, K. Wagle, C. Shahi, Y. Yamamoto, S. Romero, B. Santra, R. R. Zope, J. E. Peralta, K. A. Jackson and J. P. Perdew, *The Journal of Chemical Physics*, 2020, **152**, 214109.
- 124 Y. Yamamoto, L. Basurto, C. M. Diaz, R. R. Zope and T. Baruah, *Self-interaction correction to density functional approximations using Fermi-Löwdin orbitals: methodology and parallelization*, Unpublished.
- 125 D. Porezag and M. R. Pederson, *Phys. Rev. A*, 1999, **60**, 2840–2847.
- 126 M. R. Pederson and K. A. Jackson, *Phys. Rev. B*, 1990, **41**, 7453–7461.
- 127 J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh and C. Fiolhais, *prb*, 1992, **46**, 6671–6687.
- 128 M. R. Pederson, *J. Chem. Phys.*, 2015, **142**, 064112.
- 129 Z.-h. Yang, M. R. Pederson and J. P. Perdew, *Phys. Rev. A*, 2017, **95**, 052505.
- 130 M.-C. Kim, E. Sim and K. Burke, *Phys. Rev. Lett.*, 2013, **111**, 073003.
- 131 S. J. Chakravorty, S. R. Gwaltney, E. R. Davidson, F. A. Parpia and C. F. Fischer, *Phys. Rev. A*, 1993, **47**, 3649–3670.
- 132 A. Kramida, Yu. Ralchenko, J. Reader and NIST ASD Team, NIST Atomic Spectra Database (ver. 5.6.1), [Online]. Available: <https://physics.nist.gov/asd> [2018, July 25]. National Institute of Standards and Technology, Gaithersburg, MD., 2018.
- 133 National Institute of Standards and Technology, NIST Computational Chemistry Comparison and Benchmark Database NIST Standard Reference Database Number 101 Release 19, April 2018, Editor: Russell D. Johnson III <http://cccbdb.nist.gov/> DOI:10.18434/T47C7Z.
- 134 L. A. Curtiss, K. Raghavachari, G. W. Trucks and J. A. Pople, *J. Chem. Phys.*, 1991, **94**, 7221–7230.
- 135 B. J. Lynch and D. G. Truhlar, *J. Phys. Chem. A*, 2003, **107**, 8996–8999.
- 136 R. Peverati and D. G. Truhlar, *J. Chem. Phys.*, 2011, **135**, 191102.
- 137 J. Zheng, Y. Zhao and D. G. Truhlar, *jctc*, 2007, **3**, 569–582.
- 138 L. Goerigk, A. Hansen, C. Bauer, S. Ehrlich, A. Najibi and S. Grimme, *Phys. Chem. Chem. Phys.*, 2017, **19**, 32184–32215.
- 139 L. Goerigk and S. Grimme, *J. Chem. Theory Comput.*, 2010, **6**, 107–126.
- 140 A. Ruzsinszky, J. P. Perdew and G. I. Csonka, *The Journal of Physical Chemistry A*, 2005, **109**, 11006–11014.
- 141 S. Akter, Y. Yamamoto, C. M. Diaz, K. A. Jackson, R. R. Zope and T. Baruah, *The Journal of Chemical Physics*, 2020, **153**, 164304.
- 142 V. S. Bryantsev, M. S. Diallo, A. C. T. van Duin and W. A. Goddard, *Journal of Chemical Theory and Computation*, 2009, **5**, 1016–1026.
- 143 D. Manna, M. K. Kesharwani, N. Sylvetsky and J. M. L. Martin, *J. Chem. Theory Comput.*, 2017, **13**, 3136–3152.
- 144 P. Bhattarai, B. Santra, K. Wagle, Y. Yamamoto, R. R. Zope, K. A. Jackson and J. P. Perdew, unpublished.
- 145 B. Santra and J. P. Perdew, *J. Chem. Phys.*, 2019, **150**, 174106.
- 146 J. W. Furness and J. Sun, *Phys. Rev. B*, 2019, **99**, 041119.

LSIC(z) treats weak bound regions improperly



LSIC(w) scales down correctly

79x39mm (300 x 300 DPI)