Self-interaction corrected SCAN functional for molecules and solids in the numeric atom-center orbital framework

Dissertation
zur Erlangung des akademischen Grades
doctor rerum naturalium
(Dr. rer. nat.)
im Fach: Physik
Spezialisierung: Theoretische Physik
eingereicht an der Mathematisch-Naturwissenschaftlichen Fakultät
der Humboldt Universität zu Berlin

von
Sheng Bi

Präsidentin der Humboldt Universität zu Berlin:

Prof. Dr. Julia von Blumenthal

Dekanin der Mathematisch-Naturwissenschaftlichen Fakultät:

Prof. Dr. Caren Tischendorf

Gutachter/innen:
1. Prof. Dr. Matthias Scheffler
2. Prof. Dr. Claudia Draxl
3. Prof. Dr. Hong Jiang

Tag der mündlichen Prüfung: 21.02.2023

Abstract

The state-of-the-art "Strongly Constrained and Appropriately Normed" (SCAN) functional pertains to the family of meta-generalized-gradient approximation (meta-GGA) exchange-correlation functionals. Compared to lower-order approximations or other meta-GGA functionals, it provides a remarkable accuracy in reproducing and predicting a manifold of chemical and physical properties, both for molecules and solid materials. Nonetheless, also SCAN suffers from some well-documented deficiencies. From a practical point of view, SCAN calculations often require a huge number of iterations to reach self-consistency due to considerable numerical instabilities in the evaluation of the functional itself. From a more fundamental point of view, SCAN suffers from the notorious self-interaction error (SIE) known from semi-local exchange-correlation functionals, which limits its accuracy in several cases, e.g., for the description of charge transfer and for the assessment of electronic band gaps.

In the first part of this thesis, I revisited the known numerical instability problems of the SCAN functional in the context of the numerical, real-space integration framework used in the FHI-aims code. This analysis revealed that applying standard density-mixing algorithms to the kinetic energy density attenuates and largely cures these numerical issues. By this means, SCAN calculations converge towards the self-consistent solution as fast and as efficiently as lower-order GGA calculations.

In the second part of the thesis, I investigated strategies to alleviate the SIE in SCAN calculations by using the self-interaction correction algorithm proposed by Perdew and Zunger (PZ-SIC). Besides implementing and benchmarking the PZ-SIC algorithm in *FHI-aims*, fundamental advancements with promising features are proposed for the PZ-SIC algorithm itself. Typically, the PZ-SIC method suffers under the fact that the associated equations allow for multiple solutions (several local minima in addition to the wanted lowest-energy global minimum). In practice, this results into a pronounced starting-point dependence, i.e., the outcome of the correction varies considerably with the choice of orbitals viz. states used to initialize the description of the localized representation. For instance, the uncertainty of the SIC orbitals can differ the cohesive energy of the molecule H-F by about 1 eV and change the band-gap of the LiH crystal by more than 4 eV. Inspired by the original arguments in PZ-SIC and other localized methods, I introduced a mathematical constraint, i.e., the orbital density constraint, that forces the orbitals to retain their localization throughout the self-consistency cycle. In turn, this alleviates the multiple-solutions problem and facilitates the convergence towards the correct, lowest-energy solution both for *complex* and *real* SIC orbitals. This is demonstrated by exploring the performance of SIC-SCAN calculations for a wide range of chemical and physical properties for molecules and solids, including reaction barriers, ionization potentials, cohesive energies, and fundamental band-gaps for different kinds of materials.

The developments and investigations performed in this thesis pave the road towards a more wide-spread use of SIC-SCAN calculations in the future, allowing more accurate predictions within only moderate increases of computational cost.

Zusammenfassung

Das “Strongly Constrained and Appropriately Normed” (SCAN) Austausch-Korrelations-Funktional gehört zur Familie der meta-GGA (generalized gradient approximation) Funktionale. Im Vergleich zu anderen Funktionalen dieser Klasse besitzt SCAN eine bemerkenswerte Genauigkeit bei der Beschreibung vieler chemischen und physikalischen Eigenschaften von Molekülen und Festkörpern. Es gibt aber auch Nachteile Zum einen leiden SCAN Rechnungen oft unter numerischen Instabilitäten, wodurch sehr viele Iteration zum Erreichen von Selbst-Konsistenz benötigt werden. Zum anderen leidet SCAN unter dem von GGA Methoden bekannten Selbstwechselwirkung-Fehler.

Im ersten Teil der Arbeit habe ich die numerischen Stabilitätsprobleme in SCAN Rechnungen im Rahmen der numerischen Realraum-Integrationsroutinen im Code *FHI-aims* untersucht. Diese Analyse zeigt, dass die genannte Probleme durch Anwendung von standardisierten Dichte-Mischalgorithmen für die kinetische Energiedichte abgemildert werden können. Dadurch wird auch in SCAN-Rechnungen eine schnelle und stabile Konvergenz zur selbstkonsistenten Lösung ermöglicht.

Im zweiten Teil der Arbeit habe ich untersucht, in welchem Rahmen sich der Selbstwechselwirkung-Fehler in SCAN mittels des von Perdew und Zunger vorgeschlagenen Selbstinteraktionskorrekturalgorithmus (PZ-SIC) verringern lässt. Dazu wurde zum einen der PZ-SIC-Algorithmus in *FHI-aims* implementiert. Es wurden aber auch Optimierungen für die PZ-SIC Methode entwickelt. Bekannterweise leidet die PZ-SIC Methode unter einer Anfangswertabhängigkeit, dass heißt die Resultate der Methode können eine deutliche Abhängigkeit davon zeigen, welche Orbitale anfangs zur Initialisierung der lokalisierten Zustände gewählt werden. Beispielsweise kann die Unsicherheit der SIC-Orbitale die Atomisierungsenergie des Moleküls H-F um etwa 1 eV verändern und die Bandlücke des LiH-Kristalls um mehr als 4 eV. Inspiriert von den ursprünglichen Argumenten in der PZ-SIC-Methode und anderen lokalisierten Methoden, wird in dieser Arbeit eine neuartige Randbedingung (orbital density constraint) vorgeschlagen, die sicherstellt, dass die PZ-SIC Orbitale während des Selbstkonsistenzzyklus lokalisiert bleiben. Dies mildert die Anfangswertabhängigkeit deutlich ab und hilft dabei, in die korrekte selbst-konsistente Lösung mit minimaler Energie zu konvergieren, unabhängig davon ob reelle oder komplexe SIC Orbitale verwendet werden. Dies ist in dieser Arbeit an Hand einer breiten Palette chemischer und physikalischer Eigenschaften von Molekülen und Festkörpern gezeigt, einschließlich der Reaktionsbarrieren, Ionisierungspotentiale, Kohäsionsenergien und der fundamentalen Bandlücke für verschiedene Arten von Materialien.

Die in dieser Arbeit getägtigen Entwicklungen und Untersuchungen sind Wegbereiter dafür, in Zukunft mit SIC-SCAN Rechnungen deutlich genauere *ab initio* Rechnungen mit nur gering höherem Rechenaufwand durchführen zu können.

Acknowledgement

Looking back to the period of pursuing my Ph.D., it is definitely an unforgettable experience in my life. First and foremost, I would like to give my heartfelt thanks to my supervisor Prof. Dr. Matthias Scheffler, who brought me into the DFT world and provided me with this SIC-SCAN project. It is indeed a difficult project. It is not possible to finalize this project and write this thesis without his insightful guidance and generous support. I'd also want to express my gratitude to Prof. Dr. Claudia Draxl, my second supervisor at HU Berlin, for her unwavering support. To finish my thesis under their supervision will be the greatest honor of my life.

My gratitude especially goes out to my group leaders, Igor Ying Zhang and Christian Carbogno, who are always willing to help me and generously share their knowledge. I started empty-handed, and I learned a lot from them. I'd want to show my thankfulness to Zhenkun Yuan, Yuanyuan Zhou, and Maria Dragoumi for their discussions and encouragement. Many thanks to all my colleagues, especially Nathaniel Raimbault, Xiaojuan Hu, Runhai Ouyang, Haiyuan Wang, Dmitrii Maksimov, Florian Knoop, Nikita Rybin, Julia Pach, for fostering a positive work environment.

I am deeply grateful to my parents and my girlfriend because they stand behind me in all that I do. I will keep making my great efforts to live up to their expectations.

Table of Contents

Abbreviation

AE	absolute error.
BH76	benchmark database of barrier heights for heavy atom transfer, nucleophilic substitution, association, and unimolecular reactions.
CC	conventional cell.
CCSD(T)	single, double, and perturbative triple excitations.
DFA	density-functional approximation.
DFT	density-functional theory.
EA	electron affinity.
ELPA	eigenvalue solvers for petaflop application.
ERKALE	a flexible program package for X-ray properties of atoms and molecules.
FA	Fermi and Amaldi.
FHI-aims	Fritz Haber Institute *ab initio* molecular simulations.
FLO	Fermi orbital with Löwdin orthogonalization (Fermi-Löwdin orbital).
G2-1	Gaussian-2 theoretical procedure, based on *ab initio* molecular orbital theory, for calculation of molecular energies of compounds containing first- and second-row atoms.
Gaussian	computational chemistry software package initially released by John Pople and his research group at Carnegie Mellon University and named as Gaussian.
GE2	second-order gradient correction (part in density functional) approximation.
GE4	gradient expansion (of the exchange part in density functional approximation) up to the fourth-order.
GGA	generalized gradient approximation.
gKS	generalized Kohn-Sham.
GPAW	DFT and beyond within the projector-augmented wave method.
GSC	global scaling correction.

GTO	Gaussian-type orbital.
HD	hexagonal diamond.
HF	Hartree-Fork.
HHL	Heaton, Harrison, and Lin.
HK	Hohenberg and Kohn.
HOMO	highest occupied (molecular) orbital.
IP	ionization potential.
KS	Kohn-Sham.
LDA	local-density approximation.
LOSC	local orbital scaling correction.
LUMO	lowest unoccupied (molecular) orbital.
MAE	mean absolute error.
MAPE	mean absolute percentage error.
ME	mean error.
MLWO	maximally localized Wannier orbital.
NAO	numeric atom-centered orbital.
NRLMOL	naval research laboratory molecular orbital library.
PBC	periodic boundary condition.
PBE	Perdew-Burke-Ernzerhof.
PPLB	Perdew, Parr, Levy, and Balduz.
PyFLOSIC	python-based Fermi-Löwdin orbital self-interaction correction soft.
PZ-SIC	self-interaction correlation algorithm proposed by Perdew and Zunger.
SCAN	strongly constrained and appropriately normed.
SCF	self-consistent field.
sd-SIC	scaled-down self-interaction correlation.
SIC	self-interaction correction.
SIC-SCAN(real)	SIC-SCAN using *real* SIC orbitals.
SIC-SCAN(complex)	SIC-SCAN using *complex* SIC orbitals.

SIE	self-interaction error.
slSIC	self-consistently localized self-interaction correlation.
SSH	Su-Schrieffer-Heeger.
TPA	crystalline *trans*-polyacetylene.
TPSS	Tao-Perdew-Staroverov-Scuseria.
VASP	Vienna *ab initio* software package.
vdW	van der Waals.
W1	Weizmann-1.
WFLO	Wannier function-derived Fermi orbital with Löwdin orthogonalization (Wannier Fermi-Löwdin orbital).
XC	exchange-correlation.

1 Introduction

Kohn-Sham (KS) density-functional theory (DFT) is currently the only practical method of choice for studying the electronic structure of complex materials [1–3]. This fact is due to the good balance between computational efficiency and accuracy provided by the existing semi-local density-functional approximations (DFAs) on the exchange-correlation energy $E_{\mathrm{xc}}[n]$. The non-empirical generalized gradient approximation (GGA) of Perdew, Burke, and Ernzerhof, namely PBE [4], has been the most widely used semi-local DFA with great success in computational materials science for many decades. However, subjecting to the limitation of the GGA formula, the PBE functional is often not sufficiently accurate to describe some important physical and chemical properties [5, 6].

In 2015, Perdew and co-workers proposed a new non-empirical semi-local DFA in the context of the meta-GGA framework, which is **S**trongly **C**onstrained by obeying 17 known exact constraints and **A**ppropriately **N**ormed to the accurate exchange-correlation energies of the rare-gas atoms and the binding-energy curves of the Ar_2 dimer [7] (SCAN). The resulting SCAN functional provides a notable and consistent improvement over the PBE functional for various kinds of chemical interactions for both molecules and solids [8–10]. Despite being a promising candidate to supplant the PBE functional, there are growing shreds of evidence showing that the SCAN functional inherits the notorious self-interaction error (SIE) from PBE [6, 11, 12], meanwhile suffers from a heavier numerical instability than PBE, which undermines to some extent the potential of the SCAN functional to be the next-generation working horse in computational materials science.

The numerical instability problem of the SCAN functional can be partly traced back to the use of the kinetic energy density in the meta-GGA functional construction. It was first reported by Johnson and co-workers in 2004 [13]. Compared with the standard semi-local DFAs that depend on the density and density gradient only, a much denser integration grid is often needed to converge the total energy of dispersion-driven systems for the meta-GGAs of TPSS, VSXC, and BB95 [14], which later turned out to be a common problem of meta-GGAs, including the Minnesota family of empirical meta-GGAs [15, 16].

Unfortunately, the aforementioned numerical instability problem is more serious for the SCAN functional and often leads to the failure of SCF convergence [17, 18]. Increasing the integration grid only cannot solve this problem completely. It was found that the use of a new kind of dimensionless kinetic energy variable α gives rise to ill-behaved exchange-correlation potential values with heavy numerical noise in certain potential regions [17, 19]. In consequence, recent efforts resort to revising the functional formula. The resulting SCAN variants notably enhance the numerical robustness and steady the SCF convergence, which, however, is paid for by some changes in the performance.

My project was carried out with the Fritz Haber Institute *ab initio* molecular simulations (*FHI-aims*) code, which is an all-electron full-potential electronic-structure code in the numerical integration framework [20]. Taking advantage of the locality of the numeric atom-centered orbitals (NAOs), the formal scaling of the numerical integration implemented in *FHI-aims* is nearly linear for large systems. It thus allows us to employ very dense grids at an affordable cost for routine use. In the first part of my project, I revisited the grid sensitivity of semi-local DFAs and confirmed that the SCAN functional tends to produce much rougher XC potentials than those of PBE. On the one hand, I found that the *tight* grid setting in *FHI-aims* is sufficient to capture the SCAN XC potential. On the other hand, such an ill-behaved potential induces larger changes (or oscillation) in the electron density and kinetic energy density during the iteration. It thus makes the SCF procedure much harder to get converged. In this consideration, I generalized the standard electron density mixing algorithm to the kinetic energy density variant. Further investigations confirm that including the kinetic energy density in the mixing algorithm is the key to achieving a fast and stable convergence towards the standard SCAN self-consistent solution.

In addition to the numerical instability problem, the SCAN functional suffers from the SIE. The self-interaction originates from the Hartree approximation and, in principle, should be completely canceled out by the XC functional. Unfortunately, all existing semi-local DFAs, including SCAN, cannot meet the challenge. SIE takes the blame for the poor performance of semi-local DFAs on the fractional-charge molecular systems, the charge-transfer driven properties, and the band-gaps of insulators/semiconductors [6, 21–23]. For one-electron systems, the SIE to the semi-local DFAs is well defined, which has motivated the self-interaction correction (SIC) strategy proposed by Perdew and Zunger in the early 1980s (PZ-SIC) [24]. The PZ-SIC strategy, despite the one-electron nature, is able to accurately correct the one-/two-electron SIE and has shown to be an effectual remedy against the many-electron SIE [25, 26].

Although the concept of the PZ-SIC strategy is simple, the PZ-SIC equations are challenging to solve, hard to converge, and often yield multiple solutions (several local minima in addition to the wanted lowest-energy global minimum). Without further constraints, it is difficult to pick out the correct localized solution, which, according to the original PZ-SIC definition, should produce the lowest SIC total energy for a given semi-local DFA. This multiple-solution problem was observed initially in molecules and, subsequently, much more severely for solids. For example, by applying the PZ-SIC correction to the local-density approximation (LDA), Heaton, Harrison and Lin reported the band-gap of the Ar crystal as 13.9 eV, which is about 0.7 eV different from the value of Stengel and Spaldin [27, 28].

The past decade has witnessed a revival of the PZ-SIC strategy. A fundamental technical advance enables the solution of the PZ-SIC equations in the *complex* domain. For finite molecules, it appeared that the *complex* SIC orbitals often provide lower SIC energies than those of *real* SIC orbitals [12, 29, 30]. For solids,

to solve the PZ-SIC equations in the *complex* domain should lead to the *complex* Wannier SIC orbitals, which is expected to deliver lower SIC total energies for solids as well. However, no relevant research has yet been carried out under periodic boundary conditions.

The second part of my project realized the PZ-SIC implementation in the *FHI-aims* package. My SIC implementation solves the PZ-SIC equations in the *complex* domain and enables the treatment of both molecule and solids on an equal footing. In the spirit of Edmiston-Ruedenberg localized orbitals [31], I proposed a new SIC constraint, namely *orbital density constraint*, which can be used for all semi-local DFAs, including LDA, PBE, and SCAN. This new constraint tends to minimize the SIC total energy and maximize the spatial localization of the SIC orbitals simultaneously. I proved that this orbital density constraint does not break the already known SIC constraints and is useful for pinpointing the correct SIC orbitals for both molecules and solids. Moreover, my research demonstrates the necessity of constructing the SIC orbitals in the *complex* domain. It provides consistent and notable lower PZ-SIC total energies than the *real* SIC orbitals, which is valid for both molecules and solids.

This thesis is organized as follows: Chapter 2 presents an overview of density-functional theory, including the density-functional approximations, self-interaction errors and self-interaction corrections. In Chapter 3, we demonstrate that the calculation of SCAN functional implemented in *FHI-aims* achieves a fast and stable convergence, using the *tight* grid setting at an affordable cost and including the kinetic energy density in the mixing algorithm. The self-consistent PZ-SIC method, containing the *orbital density constraint* proposed by us, is presented in detail in Chapter 4. The importance of using *complex* SIC orbitals for both molecules and solids is discussed in this chapter. Moreover, this chapter shows the performance of our PZ-SIC method using the calculations of formation energies, the highest occupied orbital (HOMO) energies, band-gaps, reactions barriers and symmetry prediction of a group of typical molecules and solids. Finally, a brief summary concludes the whole work in Chapter 5.

2 Density-Functional Theory

2.1 Kohn-Sham Framework

For the many-body systems with N_{nuc} nuclei and N_{e} electrons, the quantum-mechanical behavior obeys the many-particle Schrödinger's equation. However, this equation can only be solved for a limited number of simple systems, such as hydrogen atoms, hydrogen-like ions, and so forth. When the motion of electrons and nuclei is uncoupled by the Born-Oppenheimer approximation [32], the ground-state energy can be written as the sum of the nuclei-nuclei Coulomb energy V_{n} and the observation value of the electron Hamiltonian operator $\hat{H}_{\text{e}}$ on the ground-state electron wave-function $\Psi_{\text{gs}}(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_{N_{\text{e}}})$:

$$E_{\text{gs}} = V_{\text{n}} + \langle \Psi_{\text{gs}} | \hat{H}_{\text{e}} | \Psi_{\text{gs}} \rangle, \tag{2.1}$$

where V_{n} comes from the nuclei-nuclei Coulomb interaction:

$$V_{\text{n}} = \frac{1}{2} \sum_{I=1}^{N_{\text{nuc}}} \sum_{\substack{J=1 \\ J \neq I}}^{N_{\text{nuc}}} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \quad . \tag{2.2}$$

$\hat{H}_{\text{e}}$ includes the electron kinetic operator $\hat{T}_{\text{e}}$, the nuclei-electron Coulomb interaction operator $\hat{V}_{\text{ne}}$, and the electron-electron Coulomb interaction operator $\hat{V}_{\text{ee}}$:

$$\begin{aligned} \hat{H}_{\text{e}} &= \hat{T}_{\text{e}} + \hat{V}_{\text{ne}} + \hat{V}_{\text{ee}} \\ \hat{T}_{\text{e}} &= \sum_{i}^{N_{\text{e}}} -\frac{1}{2} \nabla_i^2 \\ \hat{V}_{\text{ne}} &= -\sum_{i}^{N_{\text{e}}} \sum_{J}^{N_{\text{nuc}}} \frac{Z_J}{|\mathbf{r}_i - \mathbf{R}_J|} \\ \hat{V}_{\text{ee}} &= \frac{1}{2} \sum_{i=1}^{N_{\text{e}}} \sum_{\substack{j=1 \\ j \neq i}}^{N_{\text{e}}} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \quad . \end{aligned} \tag{2.3}$$

Under the Born-Oppenheimer approximation, the nuclei coordinates $\{\mathbf{R}_I\}$ and their charges $\{Z_I\}$ are treated as parameters in the Hamiltonian. The aforementioned multi-dimensional problem is reduced to $3N_{\text{e}}$ degrees of freedom, which, however, is still too complicated to be tractable in practice.

Alternatively, the ground-state electron density $n(\mathbf{r})$ is much easier to handle since it has only 3 degrees of freedom instead of $3N_{\text{e}}$ as the ground-state wave-function mentioned above. The ground-state density can be evaluated from the wave function by

$$n(\mathbf{r}) = N_{\text{e}} \int d^3\mathbf{r}_2 \cdots d^3\mathbf{r}_{N_{\text{e}}} |\Psi_{gs}(\mathbf{r}, \mathbf{r}_2, \cdots, \mathbf{r}_{N_{\text{e}}})|^2 \quad . \tag{2.4}$$

The attempt of using the ground-state electron density as the basic input can be traced back to the 1920s [33, 34]. Modern density-functional theory (DFT) was established in the 1960s started from the important works of Hohenberg and Kohn (HK) [35], which provide rigorous proof of two theorems:

- HK Existence Theorem: The external potential $\hat{v}_{\text{ext}}(\mathbf{r})$, and hence the ground-state total energy E, is a unique functional of the electron density $n(\mathbf{r})$.
- HK Variational Theorem: The trial density that minimizes the ground-state total energy is the exact ground-state density.

To be specific, the HK existence theorem suggests the existence of the density functional of the ground-state energy $E^{\text{HK}}[n(\mathbf{r})]$:

$$E^{\text{HK}}[n(\mathbf{r})] = E_{\text{ext}}[n(\mathbf{r})] + F[n(\mathbf{r})] \quad , \tag{2.5}$$

which can be further divided into two terms. The first term $E_{\text{ext}}[n(\mathbf{r})]$ is an explicit density functional, which is system-dependent, containing the energy contributions from the nuclei-nuclei/electron Coulomb interaction $\hat{v}_{\text{nuc}}(\mathbf{r})$ and any other external field $\hat{v}_{\text{other}}$.

$$\begin{aligned} E_{\text{ext}} &= \int d^3\mathbf{r}\hat{v}_{\text{ext}}(\mathbf{r})n(\mathbf{r}) \\ \hat{v}_{\text{ext}}(\mathbf{r}) &= \hat{v}_{\text{nuc}}(\mathbf{r}) + \hat{v}_{\text{other}}(\mathbf{r}) \\ \hat{v}_{\text{nuc}}(\mathbf{r}) &= \frac{1}{N_{\text{e}}}V_{\text{n}} - \sum_{J}^{N_{\text{nuc}}} \frac{Z_J}{|\mathbf{r}-\mathbf{R}_J|} \quad . \end{aligned} \tag{2.6}$$

The second term $F[n]$ is called the universal functional, which comprises the kinetic and the electron-electron interaction energies $\langle\Psi_{\text{gs}}|\hat{T}+\hat{V}_{\text{ee}}|\Psi_{\text{gs}}\rangle$. Although the HK existence theorem suggests $F[n]$ as a functional of the ground-state density, the explicit formula is still challenging to derive and remains unknown [1]. The Kohn-Sham (KS) framework proposed in 1965 smartly invents an artificial non-interacting system to conquer this problem. The KS non-interacting system shares the same ground-state density as the fully interacting real system [36]. However, it is easier to solve, and its wave function is nothing but a single Slater determinant, i.e., an expression of the multiple-electron system represented by a combination of the lowest N_{e} occupied single-electron orbitals $\{\psi_i(r)\}$.

$$\begin{aligned} \Psi_{\text{gs}}^{\text{non}} &\equiv |\psi_1\psi_2\cdots\psi_{N_{\text{e}}}\rangle \equiv \frac{1}{\sqrt{N_{\text{e}}!}}\det||\psi_i(\mathbf{r}_n)|| \\ &= \frac{1}{\sqrt{N_{\text{e}}!}}\begin{vmatrix} \psi_1(\mathbf{r}_1) & \cdots & \psi_1(\mathbf{r}_{N_{\text{e}}}) \\ \vdots & \ddots & \vdots \\ \psi_{N_{\text{e}}}(\mathbf{r}_1) & \cdots & \psi_{N_{\text{e}}}(\mathbf{r}_{N_{\text{e}}}) \end{vmatrix} \end{aligned} \tag{2.7}$$

The ground state density $n(\mathbf{r})$ of the KS non-interacting system is evaluated via

$$n(\mathbf{r}) = N_{\text{e}}\int d\mathbf{r}_2\cdots d\mathbf{r}_{N_{\text{e}}}|\Psi_{\text{gs}}^{\text{non}}(\mathbf{r},\mathbf{r}_2,\cdots,\mathbf{r}_{N_{\text{e}}})|^2 \quad . \tag{2.8}$$

The density at position ($\mathbf{r} \equiv \mathbf{r}_1$) is the sum of occupied orbital electron densities $\{|\psi_i(\mathbf{r})|^2\}$ as obtained from Eq. (2.7) :

$$\begin{aligned} n(\mathbf{r}) &= N_\mathrm{e} \frac{1}{N_\mathrm{e}!} \int d^3\mathbf{r}_2 \cdots d^3\mathbf{r}_{N_\mathrm{e}} \left(\sum_{n=1}^{N_\mathrm{e}!} (-1)^{n-1} \prod_{i=1}^{N_\mathrm{e}} \psi_i^*(\mathbf{r}_{P_n(i)}) \right) \times \\ &\left(\sum_{m=1}^{N_\mathrm{e}!} (-1)^{m-1} \prod_{j=1}^{N} \psi_j(\mathbf{r}_{P_m(j)}) \right) \\ &= \frac{1}{(N_\mathrm{e}-1)!} \int d^3\mathbf{r}_2 \cdots d^3\mathbf{r}_{N_\mathrm{e}} \sum_{n=1}^{N_\mathrm{e}!} \prod_{i=1}^{N_\mathrm{e}} |\psi_i(\mathbf{r}_{P_n(i)})|^2 \\ &= \frac{1}{(N_\mathrm{e}-1)!} \sum_{n=1}^{N_\mathrm{e}!} |\psi_{P_n^{-1}(1)}(\mathbf{r})|^2 = \frac{1}{(N_\mathrm{e}-1)!} \sum_{i}^{N_\mathrm{e}!} |\psi_i(\mathbf{r})|^2 \\ &= \sum_{i}^{N_\mathrm{e}} |\psi_i(\mathbf{r})|^2 \quad . \end{aligned} \tag{2.9}$$

Here $P_n(i)$ is the permutation operator labeling a different ordered set of numbers ranged from 1 to N_e on each index n out of $N_\mathrm{e}!$ indexes.

In the KS framework, the ground-state density functional of the real system can be reformulated as

$$E^\mathrm{KS}[n] = T_\mathrm{s}[n] + E_\mathrm{es}[n] + E_\mathrm{xc}[n] + E_\mathrm{ext}[n] \quad , \tag{2.10}$$

where $T_\mathrm{s}[n]$ is the kinetic energy of the KS non-interacting system and captures most part of the kinetic energy of the real system.

$$T_\mathrm{s}[n] = \langle \Psi_\mathrm{gs}^\mathrm{non} | \hat{T} | \Psi_\mathrm{gs}^\mathrm{non} \rangle = -\frac{1}{2} \sum_{i}^{N_\mathrm{e}} \langle \psi_i | \nabla_i^2 | \psi_i \rangle \quad . \tag{2.11}$$

For the same reason, the Hartree energy $E_\mathrm{es}[n]$ is introduced to cover the most part of the electron-electron Coulomb interaction

$$\begin{aligned} E_\mathrm{es} &= \frac{1}{2} \sum_{i=1}^{N_\mathrm{e}} \sum_{j=1}^{N_\mathrm{e}} \langle \psi_i(r)\psi_j(r') || \mathbf{r} - \mathbf{r}' |^{-1} | \psi_j(r')\psi_i(r) \rangle \\ &= \frac{1}{2} \int \int \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r} d^3\mathbf{r}' \quad . \end{aligned} \tag{2.12}$$

The remaining unknown contributions are combined and define the so-called exchange-correlation (XC) energy in KS-DFT. It is usually split into the exchange $E_\mathrm{x}[n]$ and correlation $E_\mathrm{c}[n]$ functionals defined as

$$\begin{aligned} E_\mathrm{xc} &= E_\mathrm{x} + E_\mathrm{c} \\ E_\mathrm{x} &= \langle \Psi_\mathrm{gs}^\mathrm{non} | \hat{V}_\mathrm{ee} | \Psi_\mathrm{gs}^\mathrm{non} \rangle - E_\mathrm{es} \\ E_\mathrm{c} &= \langle \Psi_\mathrm{gs} | \hat{T}_\mathrm{e} + \hat{V}_\mathrm{ee} | \Psi_\mathrm{gs} \rangle - \langle \Psi_\mathrm{gs}^\mathrm{non} | \hat{T}_\mathrm{e} + \hat{V}_\mathrm{ee} | \Psi_\mathrm{gs}^\mathrm{non} \rangle \quad . \end{aligned} \tag{2.13}$$

The XC energy can also be defined in terms of the Fermi-Coulomb XC hole $n_{\mathrm{xc}}(\lambda, \mathbf{r}, \mathbf{r}')$ with the adiabatic connection parameter λ:

$$E_{\mathrm{xc}}[n] = \int_0^1 d\lambda (\int d^3\mathbf{r} n(\mathbf{r}) (0.5 \int \frac{n_{\mathrm{xc}}(\lambda, \mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r})) \quad , \tag{2.14}$$

where the XC hole defines the difference between the pair-correlation density $g(\lambda, \mathbf{r}, \mathbf{r}')$ and the ground-state density:

$$\begin{aligned} n_{\mathrm{xc}}(\lambda, \mathbf{r}, \mathbf{r}') &= g(\lambda, \mathbf{r}, \mathbf{r}')/n(\mathbf{r}) - n(\mathbf{r}') \\ g(\lambda, \mathbf{r}, \mathbf{r}') &= N_{\mathrm{e}}(N_{\mathrm{e}} - 1) \int d^3\mathbf{r}_3 \cdots d^3\mathbf{r}_{N_{\mathrm{e}}} |\Psi^{\lambda}_{\mathrm{gs}}(\mathbf{r}, \mathbf{r}', \mathbf{r}_3, \cdots, \mathbf{r}_{N_{\mathrm{e}}})|^2 \quad . \end{aligned} \tag{2.15}$$

More details about the XC energy derivation can be found in Ref. [37].

According to the HK variational theorem, the true ground-state density minimizes the total energy $E^{\mathrm{KS}}[n]$. Enforcing the conservation of total electrons number N_{e} and the orthonormalization of the orbitals $\{\psi_i(r)\}$ results in the KS Euler equations

$$\delta \left[E^{\mathrm{KS}} - \mu (\int n(\mathbf{r}) d^3\mathbf{r} - N_{\mathrm{e}}) - \sum_{ij}^{N_{\mathrm{e}}} \gamma_{ij} (\langle \psi_i | \psi_j \rangle - \delta_{ij}) \right] = 0 \quad , \tag{2.16}$$

with the Lagrange parameter γ_{ij} and μ. The variation introduces the well-known KS one-electron equations

$$\hat{h}^{\mathrm{KS}} \psi_l = \epsilon_l \psi_l \quad , \tag{2.17}$$

with the eigenvalues ϵ_l. The KS one-electron Hamiltonian $\hat{h}^{\mathrm{KS}}$ includes the kinetic energy operator $\hat{t}_{\mathrm{s}}$, the external potential $\hat{v}_{\mathrm{ext}}$ (defined in Eq. (2.6)), the Hartree potential $\hat{v}_{\mathrm{es}}$, and the XC potential $\hat{v}_{\mathrm{xc}}$.

$$\begin{aligned} \hat{h}^{\mathrm{KS}} &= \hat{t}_{\mathrm{s}} + \hat{v}_{\mathrm{ext}} + \hat{v}_{\mathrm{es}} + \hat{v}_{\mathrm{xc}} \\ \hat{t}_{\mathrm{s}} &= \frac{\delta T_{\mathrm{s}}}{\delta n} = -\frac{1}{2}\nabla^2 \ , \ \hat{v}_{\mathrm{xc}} = \frac{\delta E_{\mathrm{xc}}}{\delta n} \\ \hat{v}_{\mathrm{es}} &= \frac{\delta E_{\mathrm{es}}}{\delta n} = \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r}' \\ \hat{v}_{\mathrm{ext}} &= \frac{\delta E_{\mathrm{ext}}}{\delta n} = \hat{v}_{\mathrm{nuc}} + \hat{v}_{\mathrm{other}} = -\sum_{J}^{N_{\mathrm{nuc}}} \frac{Z_J}{|\mathbf{r} - \mathbf{R}_J|} + \hat{v}_{\mathrm{other}} \end{aligned} \tag{2.18}$$

If the XC energy E_{xc} and its potential $\hat{v}_{\mathrm{xc}}$ are known, the system can be solved in a self-consistent manner. The calculation flow is shown in the following:

1 Guess an initial density or a set of KS orbitals $\{\psi_l\}$;

2 Solve the KS one-electron Eq. (2.17) and produce a new set of the KS orbitals $\{\psi'_l\}$;

3 Update the density in Eq. (2.9) and the total energy;

4 If the density and the total energy are not converged according to a preset threshold, go back to Step 2. Otherwise, stop the self-consistent cycle and print out the required electronic-structure information, like the total energy, the coefficient matrix for the KS orbitals, the ground-state density, and so forth.

2.2 Density-Functional Approximations

Under the KS framework of density-functional theory (KS-DFT), the remaining unknown functional is associated with the XC energy only. The central task of KS-DFT is to find more and more accurate approximations of the XC functional. In fact, the huge success of DFT achieved in the past decades owns much to the existence of many excellent density-functional approximations (DFAs) for $E_{\mathrm{xc}}[n(\mathbf{r})]$. For a given DFA, the approximated one-electron equations in the KS framework are written as

$$\hat{h}^{\mathrm{DFA}}\psi_l = \epsilon_l \psi_l \quad , \tag{2.19}$$

with the approximated one-electron Hamiltonian

$$\hat{h}^{\mathrm{DFA}} = \hat{t}_{\mathrm{s}} + \hat{v}_{\mathrm{ext}} + \hat{v}_{\mathrm{es}} + \hat{v}_{\mathrm{xc}}^{\mathrm{DFA}} \quad . \tag{2.20}$$

According to the complexity of variables used in the DFA construction, Perdew proposed the famous Jacob's Ladder in DFT to classify different levels of DFAs (See Fig. 2.1) [38].

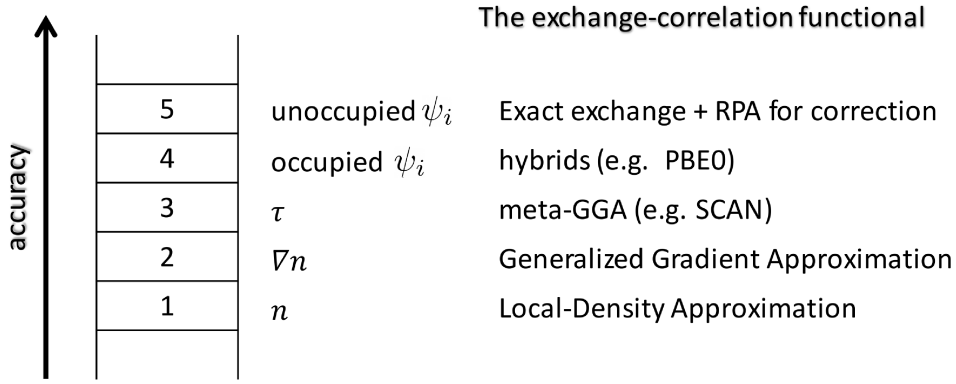

Figure 2.1: The Jacob's Ladder in DFT to classify different levels of DFAs [38].

2.2.1 Local-Density Approximation

The first rung of Jacob's Ladder is the local-density approximation (LDA):

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[n(n_\uparrow, n_\downarrow)] \equiv E_{\mathrm{xc}}^{\mathrm{LDA}}[n_\uparrow, n_\downarrow] = \int d^3\mathbf{r} n(\mathbf{r}) \epsilon_{\mathrm{xc}}^{\mathrm{unif}}[n_\uparrow, n_\downarrow] \quad , \tag{2.21}$$

where $\epsilon_{\mathrm{xc}}^{\mathrm{unif}}[n_\uparrow, n_\downarrow]$ is the XC energy density of the uniform electron gas for spin distinct densities $n_\uparrow$ and $n_\downarrow$ [39–41]. LDA satisfies many important physical

constraints that an exact XC functional must fulfill. For example, the exchange energy is less than zero, and the correlation energy is less than or equal to zero. Moreover, LDA is size-consistent, which means that for a set of well-separated systems with the densities of $\{n_1, \cdots, n_N\}$, the LDA total energy $E^{\mathrm{LDA}}[n_{\mathrm{tot}}]$ by taking the whole density $n_{\mathrm{tot}} = \sum_i^N n_i$ into account equals the sum of the separated energy of each [42]

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[n_{\mathrm{tot}}] = E_{\mathrm{xc}}^{\mathrm{LDA}}[n_1] + \cdots + E_{\mathrm{xc}}^{\mathrm{LDA}}[n_N] \quad . \tag{2.22}$$

Furthermore, LDA satisfies the following DFA constraints:

- the exact spin-scaling relation

$$E_{\mathrm{x}}[n_\uparrow, n_\downarrow] = \frac{1}{2}E_{\mathrm{x}}[n_\uparrow, n_\uparrow] + \frac{1}{2}E_{\mathrm{x}}[n_\downarrow, n_\downarrow] \quad ; \tag{2.23}$$

- the exact uniform-density-scaling relation with scaling factor γ

$$\begin{aligned} &n_\gamma(\mathbf{r}) = \gamma^3 n(\gamma\mathbf{r}) \\ &E_{\mathrm{x}}[n_\gamma(\mathbf{r})] = \gamma E_{\mathrm{x}}[n(\mathbf{r})] \\ &\lim_{\gamma\to 0} E_{\mathrm{c}}[n_\gamma(\mathbf{r})] = \gamma E_{\mathrm{c}}[n(\mathbf{r})] \quad ; \end{aligned} \tag{2.24}$$

- the Lieb-Oxford lower bound relation with the exchange energy for the uniform electron gas $\epsilon_{\mathrm{x}}^{\mathrm{unif}}$

$$E_{\mathrm{x}} \geq E_{\mathrm{xc}} \geq 1.679 \int d^3\mathbf{r} n(\mathbf{r}) \epsilon_{\mathrm{x}}^{\mathrm{unif}}[n] \quad ; \tag{2.25}$$

- a good description of the linear response of the density of a uniform electron gas to a weak static perturbation potential.

Because of these promising qualities, LDA performs well for slowly varying densities, metallic surfaces, and the predictions of lattice constants [43]. However, due to the intrinsic limitation of the approximation, LDA is not accurate enough for many cases, i.e., atomic energies are predicted too low [24]; the transfer energies between s and p/d orbitals are poorly described [44]; the bandwidths of alkali metals are noticeably overestimated in comparison to photoemission measurements [45].

2.2.2 Generalized Gradient Approximation

The second rung of Jacob's Ladder is the generalized gradient approximation (GGA), which introduces, in addition to the density itself, the density gradient ∇n in the XC energy functional [5]:

$$E_{\mathrm{xc}}^{\mathrm{GGA}}[n_\uparrow, n_\downarrow] = \int d^3\mathbf{r} n(\mathbf{r}) \epsilon_{\mathrm{xc}}^{\mathrm{GGA}}[n_\uparrow, n_\downarrow, \nabla n_\uparrow, \nabla n_\downarrow] \quad . \tag{2.26}$$

The first approximation attempting to include the information of the density gradient is the second-order gradient correction approximation (GE2) [35, 36]

$$\begin{aligned} E_{\mathrm{x}}^{\mathrm{GE2}}[n] &= \int d^3\mathbf{r}[n\epsilon_{\mathrm{x}}^{\mathrm{unif}}[n] + C_{\mathrm{x}}\frac{|\nabla n|^2}{n^{4/3}}] \\ E_{\mathrm{c}}^{\mathrm{GE2}}[n_\uparrow, n_\downarrow] &= \int d^3\mathbf{r}[n\epsilon_{\mathrm{x}}^{\mathrm{unif}} + C_{\mathrm{c}}[n]((1+\zeta)^{2/3} + (1-\zeta)^{2/3})\frac{|\nabla n|^2}{2n^{4/3}}] \quad , \end{aligned} \tag{2.27}$$

where $C_{\mathrm{c}}[n]$ is a weak function of the density tending to $C_{\mathrm{c}}[\infty] \approx 4.235\times10^{-3}$ (in Hartree atomic units) in the high-density limit, ζ is defined as $\zeta = (n_\uparrow - n_\downarrow)/n$, and C_{x}, as a negative constant, can be obtained by fitting against the known exact constraints. However, the GE2 approximation was demonstrated to be not successful since it leads to a divergence in the expansion of the XC hole. The GGAs were proposed to overcome this deficiency [5, 46]. Moreover, the high-density limit for uniform scaling is satisfied in the GGA framework:

$$\lim_{\gamma\to\infty} E_{\mathrm{c}}^{\mathrm{GGA}}[\gamma^3 n(\gamma\mathbf{r})] = \text{constant} \quad . \tag{2.28}$$

Furthermore, the non-uniform density scaling is also satisfied for GGAs

$$\begin{aligned} \lim_{s\to\infty} \epsilon_{\mathrm{x}}[s] &= \epsilon_{\mathrm{x}}^{\mathrm{unif}} O(s^{-1/2}) \\ s[n] &= |\nabla n|/[2(3\pi^2)]^{1/3} n^{4/3} \quad , \end{aligned} \tag{2.29}$$

where s is a dimensionless density gradient, and $O(s^{-1/2})$ describes the behavior of a function acting as $s^{-1/2}$ when s tends towards infinity.

The most popular GGA functional in computational materials science is the non-empirical Perdew-Burke-Ernzerhof (PBE) functional. PBE exhibits a notable and consistent improvement over LDA for the description of many different chemical and/or physical properties [4]. However, PBE is not satisfactory enough for accuracy. For example, PBE tends to overestimate lattice parameters but significantly underestimate reaction barrier heights by more than 7 kcal/mol [5, 6].

2.2.3 Meta Generalized Gradient Approximation

The third rung of Jacob's Ladder is the meta generalized gradient approximation (meta-GGA). Compared to LDA and GGA, the new ingredient introduced in the meta-GGA functional construction is the kinetic energy density:

$$\tau_\sigma = \frac{1}{2}\sum_i^{\mathrm{occ}} |\nabla\psi_{i\sigma}|^2 \quad . \tag{2.30}$$

The meta-GGA XC energy functional is written as [47–49]

$$E_{\mathrm{xc}}^{\mathrm{mGGA}}[n_\uparrow, n_\downarrow] = \int d^3\mathbf{r} n(\mathbf{r}) \epsilon_{\mathrm{xc}}^{\mathrm{mGGA}}[n_\uparrow, n_\downarrow, \nabla n_\uparrow, \nabla n_\downarrow, \tau_\uparrow, \tau_\downarrow] \quad . \tag{2.31}$$

The development of the exchange part of meta-GGA starts from the gradient expansion up to the fourth-order (GE4)

$$E_{\mathrm{x}}^{\mathrm{GE4}}[n] = \int d^3\mathbf{r}[n\epsilon_{\mathrm{x}}^{\mathrm{unif}}(n) + C_{\mathrm{x}}\frac{|\nabla n|^2}{n^{4/3}} + \alpha_{\mathrm{x}}\frac{|\nabla^2 n|}{n^2} + \beta_{\mathrm{x}}\frac{|\nabla n|^2\nabla^2 n}{n^3} + \gamma_{\mathrm{x}}\frac{|\nabla n|^4}{n^4}], \tag{2.32}$$

where $C_{\mathrm{x}} = -2.382 \times 10^{-3}, \alpha_{\mathrm{x}} = -3.633 \times 10^{-5}, \beta_{\mathrm{x}} = 9.083 \times 10^{-5}$ in Hartree atomic units. The unknown coefficient γ_{x} introduces extra freedom to the meta-GGA construction. The meta-GGA provides much larger flexibility than the GGA for the functional constructions. It is possible to satisfy more exact constraints, for example, the zero correlation condition for any one-electron spin-polarized density:

$$\begin{aligned} E_{\mathrm{c}}^{\mathrm{mGGA}}[n, 0] = 0, \\ \int d^3\mathbf{r} n(\mathbf{r}) = 1 \quad . \end{aligned} \tag{2.33}$$

To solve the KS one-electron equations (See Eqs. (2.18) and (2.20)), we need the XC potential of a given meta-GGA. The XC potential of meta-GGAs has the form of

$$v_{\mathrm{xc}}^{\mathrm{mGGA}}(\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}^{\mathrm{mGGA}}}{\delta n(\mathbf{r})} = v_{\mathrm{xc}}^{\mathrm{GGA}}(\mathbf{r}) + v_{\mathrm{xc}}^{\tau\text{-dep}}(\mathbf{r}) \quad , \tag{2.34}$$

where the GGA-like part and the τ-dependent part are defined as:

$$\begin{aligned} v_{\mathrm{xc}}^{\mathrm{GGA}}(\mathbf{r}) &= \frac{\partial \epsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial n}(\mathbf{r}) - \nabla \cdot [\frac{\partial \epsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial \nabla n}(\mathbf{r})] \quad , \\ v_{\mathrm{xc}}^{\tau\text{-dep}}(\mathbf{r}) &= \int d^3\mathbf{r}' \frac{\partial \epsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial \tau(\mathbf{r}')} \frac{\delta \tau(\mathbf{r}')}{\delta n(\mathbf{r})} \quad , \end{aligned} \tag{2.35}$$

respectively. Unlike the GGA-like part, the τ-dependent part is an implicit functional of density. Therefore, the so-called optimized effective potential technique is needed to calculate the density functional potential in the standard KS framework [50]. In practice, this requires going beyond the standard KS scheme and establishing the potential in the generalized Kohn-Sham (gKS) framework [51], which is defined as follows,

$$v_{\mathrm{xc}}^{\tau\text{-dep}}(\mathbf{r})\psi_i(\mathbf{r}) = \hat{v}_{\mathrm{xc}}^{\mathrm{mGGA}} \frac{\delta n(\mathbf{r})}{\delta \psi_i^*(\mathbf{r}')} = -\frac{1}{2}\nabla \cdot [\frac{\partial \epsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial \tau}(\mathbf{r})\nabla\psi_i(\mathbf{r})] \quad . \tag{2.36}$$

In consequence, the meta-GGA potential operator in the gKS framework [52, 53] is written as

$$\hat{v}_{\mathrm{xc}}^{\mathrm{mGGA}}\psi_i(\mathbf{r}) = v_{\mathrm{xc}}^{\mathrm{GGA}}(r)\psi_i(\mathbf{r}) + v_{\mathrm{xc}}^{\tau\text{-dep}}(\mathbf{r})\psi_i(\mathbf{r}) \quad . \tag{2.37}$$

Here, the orbital-dependent effective potential defined in Eq. (2.35) is treated as a semi-local potential of the reference systems.

The nonempirical Tao-Perdew-Staroverov-Scuseria (TPSS) meta-GGA was constructed under the aforementioned exact constraints and normed by the exchange-correlation energy of the H atom [54]. The exchange energy in TPSS is defined as the exchange energy for the uniform electron gas $\epsilon_{\mathrm{x}}^{\mathrm{unif}}$ multiplied by an enhancement factor $F_{\mathrm{x}}^{\mathrm{TPSS}}[p,z]$

$$E_{\mathrm{x}}^{\mathrm{TPSS}}[n] = \int d^3\mathbf{r} n \epsilon_{\mathrm{x}}^{\mathrm{unif}}[n] F_{\mathrm{x}}^{\mathrm{TPSS}}[p,z] \quad . \tag{2.38}$$

Here $p = s^2$ is the square of the dimensionless density gradient s defined in Eq. (2.29), and $z = \tau^{\mathrm{W}}/\tau$ is an inhomogeneity parameter related to the kinetic energy density τ and the Weizsäcker kinetic energy density $\tau^{\mathrm{W}} = |\nabla n|^2/8n$. TPSS shows an improvement over PBE for atomization energies of molecules, lattice constants of solids, and surface energies [55]. However, a limit anomaly exists for TPSS in the single-bond orbital area and makes it less accurate than PBE for structural phase transitions under the pressure of solids, such as Si, SiO_2, and Zr crystals [55]. This has been partly traced back to the inhomogeneity parameter z, which can be either 1 or 0 when approaching the limit from different directions $p \to 0$, especially in systems containing single-bond orbital [56]. A dimensionless kinetic energy variable introduced by Perdew *et al.* overcomes the order-of-limits problem [57]

$$\alpha = (\tau - \tau^{\mathrm{W}})/\tau^{\mathrm{unif}} \quad , \tag{2.39}$$

where τ^{unif} is the kinetic-energy density of the uniform electron gas. For instance, the two different limits of the very slowly varying densities ($\alpha \sim 1$) and single-bond orbitals ($\alpha \sim 0$) show no anomalies [58, 59]. The so-called "strongly constrained and appropriately normed" (SCAN) non-empirical meta-GGA was then proposed in 2015. It satisfies all the 17 known exact constraints mentioned above [7]. For convenience, I summarize the names of the 17 known exact constraints in Appendix 7.1.

Similar to the TPSS method, the exchange energy in SCAN has the form of

$$E_{\mathrm{x}}^{\mathrm{SCAN}}[n] = \int d^3\mathbf{r} n \epsilon_{\mathrm{x}}^{\mathrm{unif}}[n] F_{\mathrm{x}}^{\mathrm{SCAN}}[s,\alpha] \quad . \tag{2.40}$$

But the enhancement factor in SCAN ($F_{\mathrm{x}}^{\mathrm{SCAN}}$) depends on the dimensionless density gradient s and the dimensionless kinetic energy variable α. In the limit of $\alpha \approx 1$, the SCAN functional energy reduces to a PBE-like expression, capturing the correct behavior for slowly varying densities. It is worth noticing that a correct description of one-/two-electron systems cannot be achieved in PBE and other GGAs [5]. Perdew and co-workers argued that this problem is associated with the limit $\alpha = 0$ and demonstrated that the SCAN exchange functional solves this issue. The SCAN behavior in the limit of $\alpha = \infty$ is extrapolated from the limits of $\alpha \approx 1$ and $\alpha = 0$. Moreover, SCAN satisfies the lower Lieb-Oxford bound. In particular, for two-electron densities, we have:

$$E_{\mathrm{x}}^{\mathrm{SCAN}}[n] \geq 1.174 \int d^3\mathbf{r} n(\mathbf{r}) \epsilon_{\mathrm{x}}^{\mathrm{unif}}[n] \quad , \tag{2.41}$$

a condition not satisfied in the GGA PBE and even the meta-GGA TPSS [7]. SCAN has been demonstrated to yield consistent and notable improvements in describing many other chemical and physical properties of both molecules and solids compared to earlier semi-local DFAs, including LDAs, GGAs, and even meta-GGAs. Examples include the descriptions of ferroelectric materials, structural properties, electronic polarizability [8], and intermediate-range van der Waals (vdW) interaction [9].

2.3 Self-interaction Error (SIE)

As discussed above, the Hartree energy E_{es} (See Eq. (2.12)) introduced in the KS scheme captures most part of the electron-electron Coulomb contribution $\left\langle \Psi_{\text{gs}} \middle| \hat{V}_{\text{ee}} \middle| \Psi_{\text{gs}} \right\rangle$. The remaining part of the electron-electron Coulomb contribution, i.e., $\left\langle \Psi_{\text{gs}} \middle| \hat{V}_{\text{ee}} \middle| \Psi_{\text{gs}} \right\rangle - E_{\text{es}}$, is expected to be relatively small in most cases. Generally, this is true, for which the strongest evidence is the success of KS-DFT in the past decades. However, this formulation also introduces a notorious error, namely the self-interaction error (SIE), to the KS-DFT [24, 60]. SIE is easy to understand in one-electron systems, where there is no electron-electron Coulomb interaction. However, in the Hartree approximation (Eq. (2.12)), E_{es} is nonzero in such a one-electron system due to the spurious Coulomb interaction of the single electron with itself. Because of this error, LDA, PBE, and SCAN underestimate the (single-particle) highest occupied orbital (HOMO) energy for hydrogen atoms by about 30% - 40% [11, 29]. In the exact KS XC functional, the SIE would cancel out exactly with the exchange functional, as it appears in HF theory (See Eq. (2.13)). Unfortunately, none of the existing semi-local exchange DFAs can meet this requirement, therefore they fail in the description of one-electron systems, like the dissociation of H-H$^+$ [25, 26, 61].

For many-electron systems, it has been found that the SIE associates several worrisome failures of semi-local DFAs, including LDA, PBE, and SCAN. Examples are the overestimation of the magnetic coupling [62, 63], wrong description of charge-transfer systems [64, 65], and poor performance for band-gaps [27, 61, 66]. The accuracy of predicting band-gap is crucial because it is highly relevant to the simulation accuracy of many properties of chemical and/or physical interest, like optical properties [67–69]. In the KS-DFT scheme where the XC energy is an explicit functional of electron density, the negative of the (single-particle) HOMO energy is equal to the ionization potential (IP), and the negative of the (single-particle) lowest unoccupied orbital (LUMO) energy is expected to the electron affinity (EA) [61, 65]:

$$\begin{aligned} \text{IP} &= -\epsilon_{\text{HOMO}} \\ \text{AE} &= -\epsilon_{\text{LUMO}} \quad . \end{aligned} \tag{2.42}$$

The eigenenergy $\epsilon_{\text{LUMO/HOMO}}$ is spontaneously obtained after self-consistently solving the KS equations (Eq. (2.17)). The IP and EA can be obtained from

photoemission and inverse photoemission spectroscopy, respectively [69]. Consequently, the KS-DFT predicted band-gap between LUMO and HOMO should be equal to the difference between IP and EA, which is experimentally observable and is defined as the fundamental band-gap [70, 71]. Unfortunately, the optical processes, such as photon absorption, need also be modeled in the band-gap calculations [72, 73]. Moreover, due to the deficiencies of widely used semi-local DFAs, particularly the notorious SIEs, the KS-DFT band-gaps calculated by these DFAs are often smaller than the exact values [36, 61]. Here the KS-DFT band-gap is defined as

$$\text{KS gap} = \epsilon_{\text{LUMO}} - \epsilon_{\text{HOMO}} \quad . \tag{2.43}$$

Furthermore, only slight improvements in the band-gap performances have been found in the calculations of meta-GGAs under the gKS scheme compared with GGAs under the KS scheme [17]. This is because that gKS regards the orbital contribution as part of the reference systems rather than being part of the local effective potential. As a result, the band-gap errors about DFAs in the KS scheme carry over to the meta-GGAs in the gKS scheme. For instance, the band-gap for the diamond-C is about 5.48 eV observed by the experiment [74] but is about 4.17 eV and 4.57 eV calculated by PBE and SCAN, respectively. Furthermore, it was argued that the SIE is responsible for the failure of semi-local DFAs in the description of the transition-metal d bands and oxygen p bands in Mott insulators, like MnO, CoO, FeO, and NiO [24]. It has been indicated that using exact-exchange DFT can help to describe the d-electron hybridization compared to the semi-local DFAs [69]. Consequently, it is worth detecting the impact of SIEs and correcting the errors for DFAs in many-electron systems, especially the meta-GGAs like SCAN functional.

However, the SIE for many-electron systems is much harder to quantify than for one-electron systems [24]. In 1982, Perdew, Parr, Levy, and Balduz (PPLB) proved that, for a system with a fractional electron $N_{\text{e}} = N_{\text{e0}} + q$ [75], where N_{e0} is an integer and $0 \leq q \leq 1$, the total energy must obey

$$E(N_{\text{e}}) = (1-q)E(N_{\text{e0}}) + qE(N_{\text{e0}}+1) \quad . \tag{2.44}$$

It linearly connects the total energies of integer-charge systems [75]. This observation imposes a fundamental condition of piecewise linearity with respect to electron number for the total energy. As the electrons under semi-local DFAs always feel self-interaction and have lower energy than fact, the many-electron SIE has been linked to the wrong convex behavior of the total energy of fractional-charge systems calculated with semi-local DFAs [76]:

$$E^{\text{DFA}}(N_{\text{e0}}+q) < (1-q)E(N_{\text{e0}}) + qE(N_{\text{e0}}+1) \quad . \tag{2.45}$$

Figure 2.2 shows the energy ΔE for several semi-local DFAs in the fractionally charged Carbon atom. We can see that the SIE leads to a nonphysical stabilization of fractional-charge states, resulting in a wrong convex curve. Therefore, the SIE in semi-local DFAs is also called delocalization error [66].

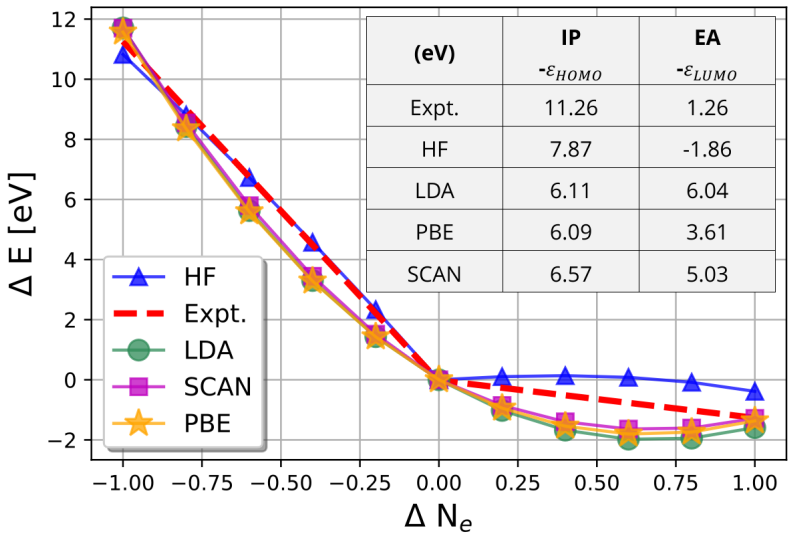

Figure 2.2: Total energy (ΔE) of the C atom as a function of the electron number ($N_\mathrm{e} = 6 + \Delta N_\mathrm{e}$) calculated by Hartree-Fork (HF), LDA, PBE, and SCAN in *FHI-aims* using *tier*-1 NAO basis sets. The energy axis is set as zero at $N_\mathrm{e} = 6$. The red dash line is plotted by straightly connecting the point $(0, 0)$ with the points $(1, -E_\mathrm{EA})$ and $(-1, E_\mathrm{IP})$, where E_IP and E_EA are the experimental ionization potential (IP) and the electron affinity (EA), respectively [77]. (insert) The calculated negative HOMO ($-\epsilon_\mathrm{HOMO}$) and LUMO ($-\epsilon_\mathrm{LUMO}$) energies of C atom ($N_\mathrm{e} = 6$) compared with experimental IP and EA.

In the following, I introduce several model systems for which the SIE plays a major role. These model systems compose two well-separated subsystems and can be categorized into two types. Let us begin with the first type. Here, we investigate two equal neutral subsystems with an integer number of electrons N_e. For each subsystem, we use E and E^DFA to denote the exact energy and the approximated energy calculated by a given DFA, respectively. Assuming one electron is removed from the whole system, both subsystems become fractionally charged with the number of electrons $N_\mathrm{e} - q$ and $N_\mathrm{e} - 1 + q$, respectively. Here $(q,\ 1-q)$ with $0 \leq q \leq 1$ denotes this fractional-charge state. According to the PPLB condition (Eq. (2.44)), the total energy of the systems with different charge distributions obeys the relationship of

$$\begin{aligned} &E(N_\mathrm{e} - q) + E(N_\mathrm{e} - 1 + q) \\ &= (1 - (1-q))E(N_\mathrm{e} - 1) - (1-q)E(N_\mathrm{e}) \\ &+ (1-q)E(N_\mathrm{e} - 1) + qE(N_\mathrm{e}) \\ &= E(N_\mathrm{e} - 1) + E(N_\mathrm{e}). \end{aligned} \tag{2.46}$$

which means that any fractional-charge states $(q,\ 1-q)$ should lead to the same energy as that of the integer-charge state $(1,\ 0)$ or $(0,\ 1)$. Unfortunately, as

shown in Eq. (2.45) and Fig. 2.2, the SIE (or the delocalization error) in semi-local DFAs breaks the PPLB condition and has a partiality for the fractional-charge states

$$\begin{aligned} &E^{\mathrm{DFA}}(N_{\mathrm{e}}-0.5)+E^{\mathrm{DFA}}(N_{\mathrm{e}}-0.5) \\ &\leq E^{\mathrm{DFA}}(N_{\mathrm{e}}-q)+E^{\mathrm{DFA}}(N_{\mathrm{e}}-1+q) \\ &< E^{\mathrm{DFA}}(N_{\mathrm{e}}-1)+E^{\mathrm{DFA}}(N_{\mathrm{e}}), \end{aligned} \tag{2.47}$$

resulting in the symmetrically distributed state (0.5, 0.5) as the most stable state for any semi-local DFAs.

The dissociation of one-electron system H-H^+ is the simplest model system belonging to this type of problem. As discussed above, the electron-electron Coulomb interaction should be zero for any one-electron system. In KS-DFT, it means that

$$E_{\mathrm{es}} = -E_{\mathrm{xc}}. \tag{2.48}$$

However, it cannot be satisfied by any semi-local DFAs, indicating a strong one-electron SIE in the approximation. Fig. 2.3-top shows the performance of different semi-local DFAs on the H-H^+ dissociation curve. The well-separated integer-charge state (1, 0) is taken as the reference. In the equilibrium region, all the semi-local DFAs perform well and result in an error smaller than 0.1 eV. However, in the dissociation limit, the two hydrogen atoms are well separated with one electron removed. It is not surprising that in the dissociation limit, all semi-local DFAs, including LDA, PBE, and SCAN, predict the well-separated fractional-charge state (0.5, 0.5) to be the most stable one, yielding dissociation energy too low with an error larger than 1.5 eV.

It is worth noting that the condition of Eq. (2.48) is not satisfied by the semi-local DFAs for the whole dissociation curve. Therefore, from the absolute energy point of view, the semi-local DFAs suffer from the heavy one-electron SIE, whether in the equilibrium region or the dissociation limit. The good performance in the equilibrium region is due to the fact that the SIE in this region is on par with the SIE in the reference state, i.e., the well-separated integer-charged state (1, 0). However, this error cancellation breaks in the dissociation limit. The violation of the PPLB condition (Eqs. (2.44) and (2.45)) suggests that the one-electron SIE in the fractional-charge state is much larger and reaches its maximum value at the symmetrically distributed state (0.5, 0.5).

Another first-type model system investigated here is the Ne-Ne^+ dissociation curve. Fig. 2.3-bottom shows the performance of different semi-local DFAs on this system. Same as for the H-H^+ dissociation, the well-separated integer-charge state (1, 0) is taken as the reference. For comparison, the single, double, and perturbative triple excitations (CCSD(T)) curve is provided in Fig. 2.3-bottom as well. CCSD(T) is a high-level wave-function method that is often considered the "gold standard" in quantum chemistry [78–80]. CCSD(T) can degenerate to the Hartree-Fork (HF) method in one-electron systems.

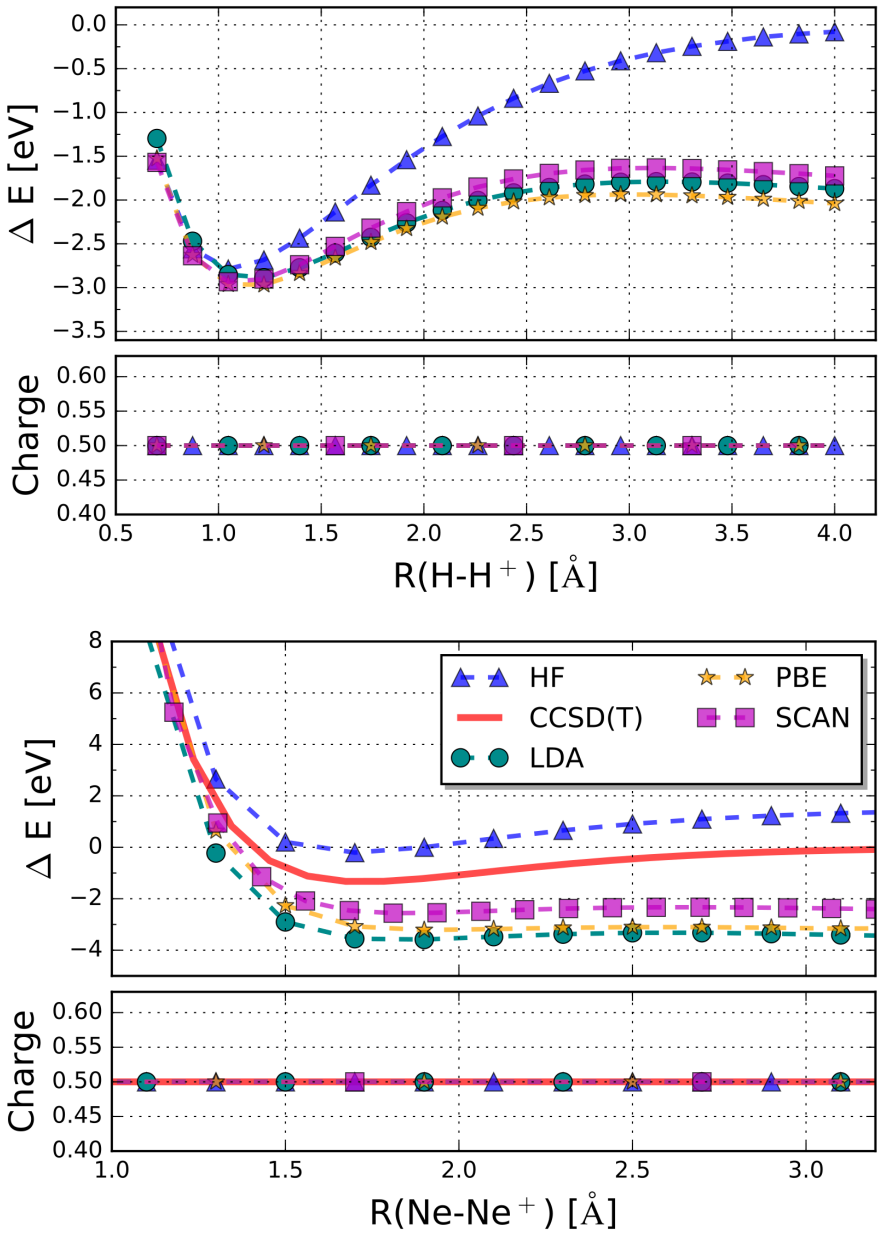

Figure 2.3: Dissociation curves of H-H$^+$ (top) and Ne-Ne$^+$ (bottom) with the zero energy level set to the total energy of isolated atoms/ions. The Mulliken charge analyses in the dissociation H-H$^+$ and Ne-Ne$^+$ are studied, for which the charge transfer to the H atom and the Ne atom are provided. Hartree-Fork (HF), LDA, PBE, SCAN are calculated in *FHI-aims* using *tight* basis sets. CCSD(T) is calculated in *FHI-aims* using cc-pVTZ Gaussian-type basis sets. The HF method is exact for one-electron systems and therefore is treated as a reference in the curve of H-H$^+$.

For many-electron systems, the electron-electron Coulomb interaction is non-zero. In consequence, the condition of Eq. (2.48) does not hold anymore and cannot be used to diagnose the many-electron SIE in the semi-local DFAs. Fortunately, the PPLB condition is still valid. The dissociation curves of LDA and PBE deviate significantly from the CCSD(T) values. The maximum deviation (larger than 3.0 eV) appears in the dissociation limit. It originates from the many-electron SIE of semi-local DFAs, which mistakenly stables the fractional-charge states. The SCAN dissociation curve displays a notable improvement over LDA and PBE, but it is not yet satisfactory.

The second type of model system comprises two different well-separated subsystems, denoted S_a and S_b. This type of system is relative to novel interface materials, like luminescence diodes [81–83]. Due to the loss of the mirror symmetry, the charge transfer between S_a and S_b is allowed, reflecting the different electron affinity of the subsystems. The discussion above reveals that the SIE of semi-local DFA tends to overemphasize the fractional-charge (or delocalized) state and can lead to a non-physical charge transfer. While the first-type model system uncovers the influence of the SIE on the total (or relative) energy, the second-type model system can be used to directly study the influence of SIE on the density and even the electronic structure. The dissociation of the H-F molecule is a typical example. Fig. 2.4 shows the performance of SCAN on this

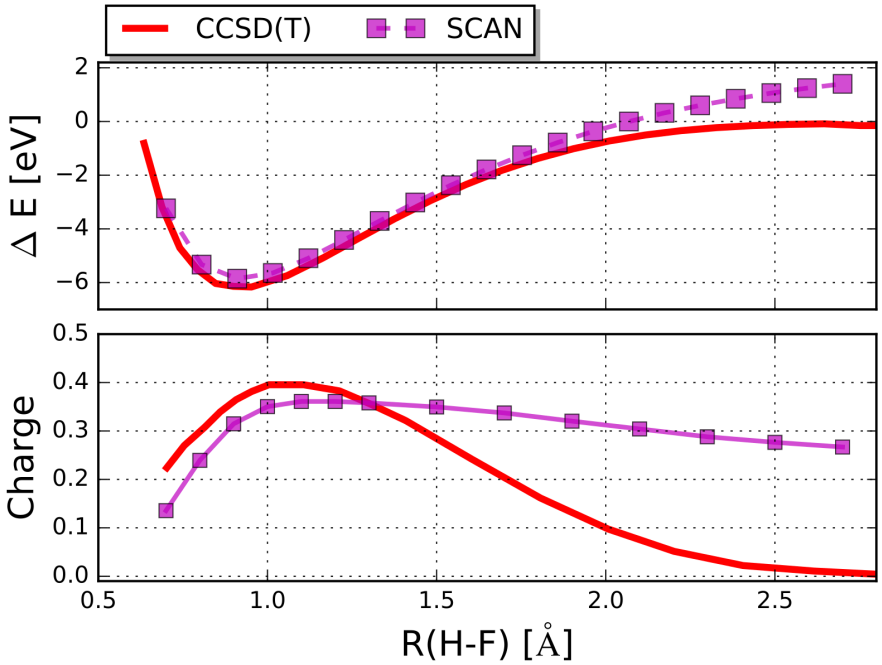

Figure 2.4: Dissociation curves of H-F with the zero energy level set to the total energy of isolated atoms/ions. The Mulliken charge analysis in this dissociation is studied, for which the charge transfer to the H atom is provided. SCAN and CCSD(T) are calculated in *FHI-aims* using *tight* basis sets and cc-pVTZ Gaussian-type basis sets respectively.

system, and a CCSD(T) curve is provided as a reference. Near the equilibrium region (0.8-1.2 Å), SCAN reproduces the CCSD(T) and results in an energy error smaller than 0.2 eV. However, in the dissociation limit, the H and F atoms are well separated and do not have any charge transfer with each other. SCAN, as a semi-local functional, predicts a fractional-charge (or delocalized) state, having about $0.26e$ charge transfer from the F to the H atom. The maximum deviation (larger than 1.5 eV) appears in the dissociation limit.

2.4 Self-interaction Correction (SIC)

Because the notorious SIEs impeded the development of DFT when it was born, the attempt of correcting SIEs was acknowledged a long time ago in the 1930s [60]. With a long period of development, the modern self-interaction correction (SIC) methods for directly removing SIEs in semi-local DFAs are classified into two popular options: the scaling correction method [61] and the Perdew-Zunger SIC method [24].

2.4.1 Scaling Correction

Imposing the PPLB condition (Eq. (2.44)) on any given fractional-charge system gives the global scaling corrected (GSC) DFA method developed by Xiao *et al.* in 2011 [61]. The GSC energy of a system with fractional charges $N_\mathrm{e} + q$ is calculated by interpolating the energy of the system with integer charges N_e and $N_\mathrm{e} + 1$:

$$E^\mathrm{GSC}(N_\mathrm{e} + q) = (1 - q)E^\mathrm{DFA}(N_\mathrm{e}) + qE^\mathrm{DFA}(N_\mathrm{e} + 1) - E^\mathrm{DFA}(N_\mathrm{e} + q) \quad . \tag{2.49}$$

It has been discussed that the GSC spontaneously restores the energy linearity behavior at any fractional charge q for the LDAs and GGAs [61]. However, GSC does not offer any contribution to the integer charged systems according to its definition, which implies that the GSC-DFAs are not size-consistent between the calculations for the integer and fractional-charge systems [84].

Since 2015, a series of local orbital scaling corrections (LOSCs) [84–86] on DFAs have been proposed to enforce the PPLB condition and process integer/neutral-charged systems. These LOSCs are developed from an inner perspective of the fractional-charge orbital densities instead of the total density (namely LOSC orbital densities). As the extrapolation of the GSC-DFA, the LOSC-DFA retains the ability to describe the PPLB condition (Eq. (2.44)). Furthermore, it has been demonstrated that the LOSC-DFA effectively discards the SIEs in the integer/neutral charged systems. For example, LOSC-DFA can well describe the dissociation limits of molecules (like ionic species, single bonds, and multiple bonds) [86]. Although LOSC-DFA has led to significant improvement in the HOMO/LUMO energy prediction for molecules (like polyacene and pentacene) [84], the screening of feasible LOSC orbitals is intense. The post-self-consistent

field approach (post-LOSC) has been well investigated at the beginning of the development, in which the converged electron density from the parent DFA is directly used to evaluate the energy correction of LOSC [87, 88]. The Boys localization procedure is used to define the LOSC orbitals [89]:

$$\min \sum_i [\langle \phi_i | \mathbf{r}^2 | \phi_i \rangle - \langle \phi_i | \mathbf{r} | \phi_i \rangle^2] \quad . \tag{2.50}$$

Using the Boys localization, the spectrum of LOSC orbitals is identical to the spectrum of auxiliary canonical orbitals and therefore gives effective correction to the HOMO/LUMO energy for molecules [84]. However, applying post-LOSC in these cases is insufficient because it improves only the energy for the parent DFA but leaves the significant error in electron density unchanged. In 2020, a reliable self-consistent field LOSC (SCF-LOSC) method for molecular systems has been proposed [90]. The electron density is updated and iterated by solving the LOSC-DFA Hamiltonian until a converged electron density is obtained. The SCF-LOSC method is demonstrated to correctly describe the electron densities, total energies, and energy-level alignments for the molecular dissociation process, like Li-H, Li-F, and benzenediamine-tetracyanoethylene. LOSC is now a promising method for studying problems for correct electron densities and energy-level alignments in molecular systems [90, 91].

2.4.2 Perdew-Zunger SIC

Another SIC option generalizes the correction for one-electron systems to the correction for many-electron systems. As the Hartree energy introduces the SIEs, a prototypical version of the SIC method aims to remove the Hartree energy in one-electron systems, as Fermi and Amaldi (FA) proposed in the 1930s [60]. In FA-SIC, the Hartree energy is directly shifted to zero for one-electron systems:

$$E_{\text{es}}^{\text{FA-SIC}} = E_{\text{es}}[n] - N_{\text{e}} E_{\text{es}}[\frac{n}{N_{\text{e}}}] \quad . \tag{2.51}$$

Although the FA-SIC Hartree energy vanishes in one-electron systems, the FA-SIC method can not cancel the remaining DFA XC energy. While the FA-SIC method is not a particularly good model for the DFA XC energy, the FA-SIC-DFA calculations can result in spectra of alkali metals semi-quantitatively matching the experiment [92].

As there is no electron-electron interaction by definition for one-electron systems, the SIE in one-electron systems can be consequently evaluated by Eq. (2.48) exactly. This analysis has motivated the celebrated self-interaction correction strategy for the semi-local XC DFAs proposed by Perdew and Zunger in the early 1980s (PZ-SIC) [24]:

$$E^{\text{PZ-SIC}}[\{n_i\}] = -\sum_i^{N_{\text{e}}} \left(E_{\text{xc}}^{\text{DFA}}[n_i] + E_{\text{es}}[n_i] \right) \quad . \tag{2.52}$$

Here $\{n_i(\mathbf{r}) = |\phi_i(\mathbf{r})|^2\}$ are the densities of the N_e single-electron orbitals. $\{\phi_i(\mathbf{r})\}$ are called SIC orbitals in the following. Accordingly, $E_\mathrm{xc}^\mathrm{DFA}[n_i]$ is the self-exchange-correlation energy evaluated for the orbital density $n_i(\mathbf{r})$ for a given DFA, while $E_\mathrm{es}[n_i]$ is the self-electrostatic energy of $n_i(\mathbf{r})$, evaluated via the Hartree equation (2.12). The SIC orbitals must fulfill the *total density constraint*, i.e., the SIC orbitals must be unitary transformations of occupied KS orbitals (See Section 4.1 for details). Different from the LOSC method [84], the correction $E^\text{PZ-SIC}$ is exact for one-electron systems, in which the only KS orbital defines the SIC orbitals uniquely [26]. For many-electron systems, the total-density constraint is, however, not sufficient to uniquely define the SIC orbitals. This has significantly hindered the routine use of PZ-SIC approaches since it has been demonstrated that PZ-SIC is only capable of correcting for large portions of the SIE if an appropriate choice of SIC orbitals is employed [27, 76].

In aid of filtering out the solutions of the PZ-SIC equations so to get the lowest total energy of the ground state of SIC-DFAs, Lin *et al.* introduced an orbital potential constraint to the SIC orbitals [93], in which the SIC orbitals make the variation of the total energy of SIC-DFAs equal to zero. (A detailed description of the orbital potential constraint is given in Sec. 4.1.) In 1988, the orbital potential constraint was used in SIC-LDA calculations for atomic systems, resulting in the experimental ionization potential even better than HF results [94]. However, in the case of benzene, erroneous geometry with alternating bond lengths [30] was observed. Moreover, many localization methods, including the Boys localization used in LOSC, were tested to determine that different localization procedures can violently impact the SIC performance [30]. This was traced back to the fact that the orbital potential constraint is insufficient to determine the correct SIC orbitals. In 2014, Pederson and co-workers suggested that the problem can be circumvented by using the Fermi orbitals with Löwdin orthogonalization (FLOs) to construct the SIC orbitals under SIC potential constraint [95, 96]. Although the FLO-SIC method requires a careful self-consistency check [97], the resulting FLO-SIC scheme makes it easier to find the correct SIC solutions than the standard PZ-SIC equations, particularly for molecules with π bonds [98].

In fact, the problem of multiple solutions is even more serious for solids. Due to the localized nature of the SIC orbitals, Heaton, Harrison, and Lin (HHL) suggested that it is more convenient to expand the SIC orbitals by using Wannier orbitals [27] instead of expanding the SIC orbitals in **k**-space. In early 2021, Shinde *et al.* proposed a tentative implementation of Wannier function-derived FLOs (WFLOs) for correcting the SIE in periodic systems [99]. Calculations on several (17 in total) prototypical molecular solids, semiconductors, and wide-band-gap materials show that the WFLO-SIC approach gives better band-gaps and bulk moduli compared to the semi-local PBE functional, largely due to the partial removal of SIEs. Although the Wannier orbitals $\{\Psi_{l,\mathbf{L}}(\mathbf{r})\}$ are localized molecular orbitals in solids, the same problem of solids as molecules is that the

orbital potential constraint does not ensure a unique set of SIC orbitals. The uncertainty of SIC orbitals passes to the SIC-DFA Hamiltonian and eventually impacts the orbital energies. For example, the SIC-LDA band-gap of the Ar crystal reported by HHL is different from that of Stengel and Spaldin by a deviation of 0.7 eV [27, 28]. The SIC orbitals in those results were both filtered by the orbital potential constraint but used different localized representations for the Wannier orbitals.

Another challenge for the original PZ-SIC method is that the corrections on GGAs or meta-GGAs are not accurate enough, especially for describing the equilibrium properties, like atomization energies, enthalpies, and equilibrium bond lengths [100, 101]. The accuracies are not improved by its application on GGAs or meta-GGAs compared with its application on LDAs, while the accuracy is improved by the nonempirical semi-local functional climbing the DFT Jacob's ladder (e.g., from LDA to GGA PBE and meta-GGA TPSS). For example, degressive deviations from the experiment of formation energies for the G2-1 test set [102, 103] are expected to be -36.0, -6.6, and -3.6 (a.u.) for LDA, PBE, and TPSS, but SIC-LDA, SIC-PBE and SIC-TPSS show deviations valued as -21.5, 10.0 and 15.0 (a.u.) [101]. One explanation is that the physical constraints or/and experimental data are reproduced by the DFA functionals with self-interaction terms included [101]. A way to systematically improve the PZ-SIC for (meta-)GGAs is to introduce the *complex* SIC orbitals to systems. It has been found that *complex* SIC orbitals generate lower total energies of atoms H to Ar and make the HOMO energies closer to experimental ionization energies [29]. In 2019, Shahi *et al.* [12] documented that the *complex* SIC orbitals can help SIC-PBE and SIC-SCAN to predict more accurate reaction barriers than the *real* SIC orbitals for a small representative BH6 set [104]. Unfortunately, only the *real* Wannier SIC orbitals were used in the WFLO-SIC work of Shinde *et al.* and other early SIC works for solids. To my best knowledge, the effect of *complex* SIC Wannier orbitals for solids is not invested and presented.

3 SCAN in FHI-aims

The meta-GGA SCAN method provides a consistent improvement over PBE for different kinds of chemical and physical properties, holding the promise to be the next-generation working horse in computational materials science. However, it suffers from a heavy numerical instability problem. It was found that meta-GGAs are more sensitive than the GGAs to the grids used in the numerical integration for finite and periodic systems, such as Ge atom [105] and bulk Si [17]. By employing the standard grid setting that is dense enough to converge the GGA calculations, the meta-GGA calculations often show spurious oscillations in the dissociation curve of the argon dimer [13, 14], and give rise to a numerical error of more than 6 kcal/mol in the reaction energies [15, 16]. These problems are partly due to including a more complicated ingredient in the functional construction, i.e., the kinetic energy density τ.

In addition to the grid sensitivity, several studies suggested that the meta-GGAs converge much slower than the standard GGAs or even diverge during the self-consistent field (SCF) procedure [14, 17, 106, 107]. It is documented that the τ-dependent dimensionless variable α (See Eq. (2.39)) defined in those meta-GGAs may cause numerical instability to the XC potential [108]. Specifically, ill-conditioned oscillations may arise in the XC potential when the α approaches 1 (i.e., in an area having slowly varying densities). This observation has motivated proposing several α variants to the SCAN functional, i.e., rSCAN [19] and r^{2}SCAN [105], both of which efficaciously reduce the grid sensitivity and accelerate the SCF convergence by smoothing/damping the sharp oscillations in the XC potential. However, rSCAN and r^{2}SCAN no longer fulfill all the 17 exact constraints as the standard SCAN, thus showing less accuracy for closed-shell complexes and lattice constant prediction [109]. (The detailed cross-comparison for the constraints satisfied by SCAN, rSCAN, and r^{2}SCAN is presented in Appendix 7.1.)

The research in this chapter was carried out with the *FHI-aims* package, which seeks to examine and circumvent the aforementioned instability problem of SCAN in the fully numerical integration framework using numeric atom-center orbitals.

3.1 Basis Set: Numeric Atom-Centered Orbitals

The Fritz Haber Institute *ab initio* molecular simulations (*FHI-aims*) code is an all-electron, full-potential electronic-structure package in the numerical integration framework [20, 110]. The numeric atom-centered orbitals (NAOs) used in *FHI-aims* have the form

$$\varphi_i(\mathbf{r}) = \frac{u_i(r)}{r} Y_m^l(\Omega) \quad , \tag{3.1}$$

where $\{Y_m^l\}$ are the spherical harmonics with the implicitly i-dependent index functions $l(i)$ and $m(i)$. The radial shape $u_i(r)$ typically is the numerical solutions of the Schrödinger-like radial equation:

$$[-\frac{1}{2}\frac{d^2}{dr^2} + \frac{l(l+1)}{r^2} + v_i(r) + v_{\text{cut}}]u_i(r) = \epsilon_i u_i(r). \tag{3.2}$$

The cut-off potential v_{cut} is a steeply increasing confining potential, which ensures a smooth decay of each radial function to zero outside a limit radius. The potential $v_i(r)$ mainly controls the shape of the radial function. The NAO basis sets in *FHI-aims* take the self-consistent free-atom radial potential to generate the radial functions $\{u_i(r)\}$ used on the *minimal basis*. In order to approach the complete basis-set limit, the so-called *tier-n* basis sets (with $n = 1, 2, 3, \cdots$) introduce more NAO basis functions on top of the minimal basis tier by tier. Compared with Dunning's cc-pVnZ Gaussian-type basis sets, the *tier*-1 basis set has a similar size as the cc-pVDZ, *tier*-2 as cc-pVTZ, and *tier*-3 as cc-pVQZ [110].

The KS orbitals are expanded using a given set of basis functions, $\{\varphi_i(\mathbf{r})\}$,

$$\psi_l = \sum_{i=1}^{N_{\text{b}}} c_{il}\varphi_i(\mathbf{r}) \quad , \tag{3.3}$$

and the electron density using basis functions is represented as:

$$n(\mathbf{r}) = \sum_{ijl} \varphi_i^*(\mathbf{r})\varphi_j(\mathbf{r})c_{il}^* c_{jl} \quad . \tag{3.4}$$

The gradient density and kinetic energy density are written as

$$\begin{aligned} \nabla n(\mathbf{r}) &= \sum_{ijl} (\varphi_j(\mathbf{r})\nabla\varphi_i^*(\mathbf{r}) + \varphi_i^*(\mathbf{r})\nabla\varphi_j(\mathbf{r}))c_{il}^* c_{jl} \\ \tau(\mathbf{r}) &= \frac{1}{2}\sum_{ijl} (\nabla\varphi_i^*(\mathbf{r}) \cdot \nabla\varphi_j(\mathbf{r}))c_{il}^* c_{jl} \quad , \end{aligned} \tag{3.5}$$

which require the gradient basis functions $\nabla\varphi_i(\mathbf{r})$ in (meta-)GGAs. It is favorable to update the gradient density and kinetic energy density based on the basis functions because the gradient of the basis functions can be prepared once and stored in the program instead of calculating the gradient of updated KS orbitals in each SCF iteration.

The corresponding DFA one-electron KS equations (Eq. (2.19)) can then be written as

$$\sum_j h_{ij}^{\text{DFA}} c_{jl} = \epsilon_l \sum_j s_{ij} c_{jl} \quad . \tag{3.6}$$

Here the Hamilton and overlap matrix elements are obtained by the numerical integration

$$\begin{aligned} h_{ij}^{\text{DFA}} &= \int d^3\mathbf{r}[\varphi_i^*(\mathbf{r})\hat{h}^{\text{DFA}}\varphi_j(\mathbf{r})] \\ s_{ij} &= \int d^3\mathbf{r}[\varphi_i^*(\mathbf{r})\varphi_j(\mathbf{r})] \quad . \end{aligned} \tag{3.7}$$

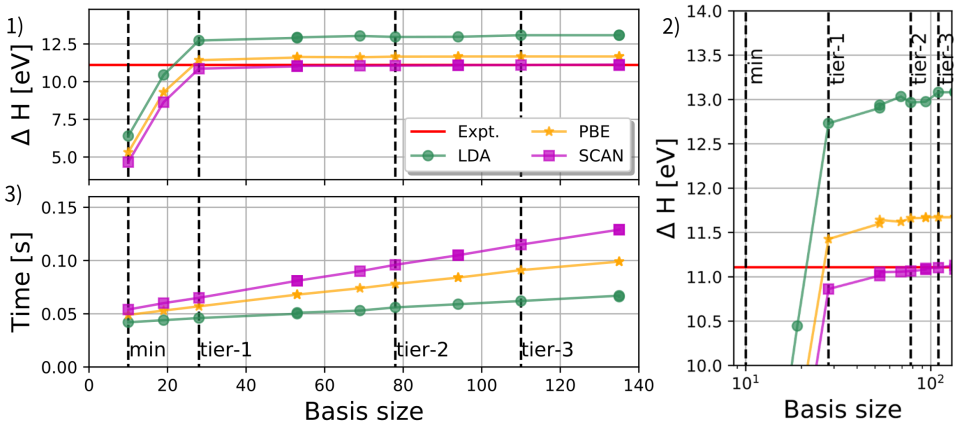

Figure 3.1: 1)-2) The convergence of the cohesive energy ΔH and 3) the time cost per iteration for the calculations of molecule CO in LDA, PBE and SCAN as a function of basis size. The black dash lines, marked with min and tier-1,2,3, denote the predefined *minimal* and *tier*-1,2,3 basis sets, respectively. The red lines present the experimental cohesive energies of CO [77].

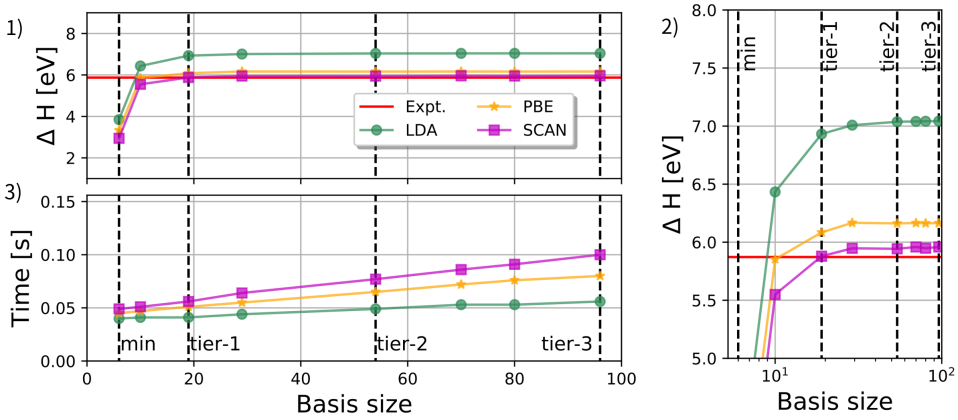

Figure 3.2: 1)-2) The convergence of the cohesive energy ΔH and 3) the time cost per iteration for the calculations of molecule HF in LDA, PBE and SCAN as a function of basis size. The black dash lines, marked with min and tier-1,2,3, denote the predefined *minimal* and *tier*-1,2,3 basis sets, respectively. The red lines present the experimental cohesive energies of HF [77].

Fig. 3.1 and 3.2 present the basis-set convergence of the SCAN functional in the calculations of cohesive energies. The small molecules CO and HF are considered. Our results support evidence from the previous work that *tier*-1 is good enough for LDAs and GGAs [3, 20]. Fig. 3.1 and 3.2 also suggests that the

SCAN functional shares the same basis-set convergence as LDA and PBE. The SCAN cohesive energies at the basis-set level of *tier*-1 have been well converged on the condition of speedy computations.

In the case of solids, the periodic boundary conditions (PBCs) are used in practical numerical calculations [20, 111]. The PBCs arise a translationally invariant area, which is characterized by the lattice vectors of unit cell $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$. Under the PBCs, the KS orbitals $\{\psi_{l,\mathbf{k}}(\mathbf{r})\}$ depend on $\mathbf{k}$-points and are the eigenstates of the translation operations $\hat{T}$,

$$\hat{T}^{\dagger}(\mathbf{R_L})\psi_{l,\mathbf{k}}(\mathbf{r}) = \psi_{l,\mathbf{k}}(\mathbf{r}+\mathbf{R_L}) = e^{i\mathbf{k}\cdot\mathbf{R_L}}\psi_{l,\mathbf{k}}(\mathbf{r}) \quad . \tag{3.8}$$

Here $\mathbf{R_L} = \sum_i L_i \mathbf{a}_i$ are the lattice vectors, with the subscript $\mathbf{L} = [L_1, L_2, L_3]$ being the index of a given lattice vector. Because the eigenvalues are valued as phases $|e^{i\mathbf{k}\cdot\mathbf{R_L}}| = 1$, KS orbitals have the invariant value in each cell $|\psi_{l,\mathbf{k}}(\mathbf{r}+\mathbf{R_L})| = |\psi_{l,\mathbf{k}}(\mathbf{r})|$. Consequently, the total density is translationally invariant in the PBC area

$$\begin{aligned} n(\mathbf{r}) &= \sum_{l,\mathbf{k}} |\psi_{l,\mathbf{k}}(\mathbf{r})|^2 \\ \hat{T}^{\dagger}(\mathbf{R_L})n(\mathbf{r}) &= n(\mathbf{r}+\mathbf{R_L}) = n(\mathbf{r}) \quad . \end{aligned} \tag{3.9}$$

For simulating the extended systems under the PBCs, a set of Bloch-like basis functions $\chi_{i,\mathbf{k}}(\mathbf{r})$ for different $\mathbf{k}$-points in the first Brillouin zone have to be prepared from the *real*-space NAO basis functions $\varphi_i(\mathbf{r})$

$$\chi_{i,\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{L}} e^{i\mathbf{k}\cdot\mathbf{R_L}} \cdot \varphi_i(\mathbf{r}-\mathbf{R_L}). \tag{3.10}$$

The KS orbitals in a given $\mathbf{k}$-point are expanded using the Bloch-like basis functions at that $\mathbf{k}$-point

$$\psi_{l,\mathbf{k}} = \sum_{i=1}^{N_\mathrm{b}} c_{il,\mathbf{k}}\chi_{i,\mathbf{k}}(\mathbf{r}). \tag{3.11}$$

In consequence, the corresponding KS equations become

$$\sum_j h^{\mathrm{DFA}}_{ij,\mathbf{k}} c_{jl,\mathbf{k}} = \epsilon_{l,\mathbf{k}} \sum_j s_{ij,\mathbf{k}} c_{jl,\mathbf{k}} \quad , \tag{3.12}$$

which can be solved independently for each $\mathbf{k}$ point. The corresponding Hamilton and overlap matrix elements are defined as

$$\begin{aligned} h^{\mathrm{DFA}}_{ij,\mathbf{k}} &= \langle\chi_{i,\mathbf{k}}|\hat{h}^{\mathrm{DFA}}|\chi_{j,\mathbf{k}}\rangle = \sum_{\mathbf{I},\mathbf{J}} e^{i\mathbf{k}\cdot[\mathbf{R_I}-\mathbf{R_J}]}\langle\varphi_{i,\mathbf{I}}|\hat{h}^{\mathrm{DFA}}|\varphi_{i,\mathbf{J}}\rangle \\ s_{ij,\mathbf{k}} &= \langle\chi_{i,\mathbf{k}}|\chi_{j,\mathbf{k}}\rangle = \sum_{\mathbf{I},\mathbf{J}} e^{i\mathbf{k}\cdot[\mathbf{R_I}-\mathbf{R_J}]}\langle\varphi_{i,\mathbf{I}}|\varphi_{i,\mathbf{J}}\rangle \quad . \end{aligned} \tag{3.13}$$

Fig. 3.3 presents the basis-set convergence of LDA, PBE, and SCAN to calculate band-gaps and cohesive energies of the fcc Si crystal. Same as the case for

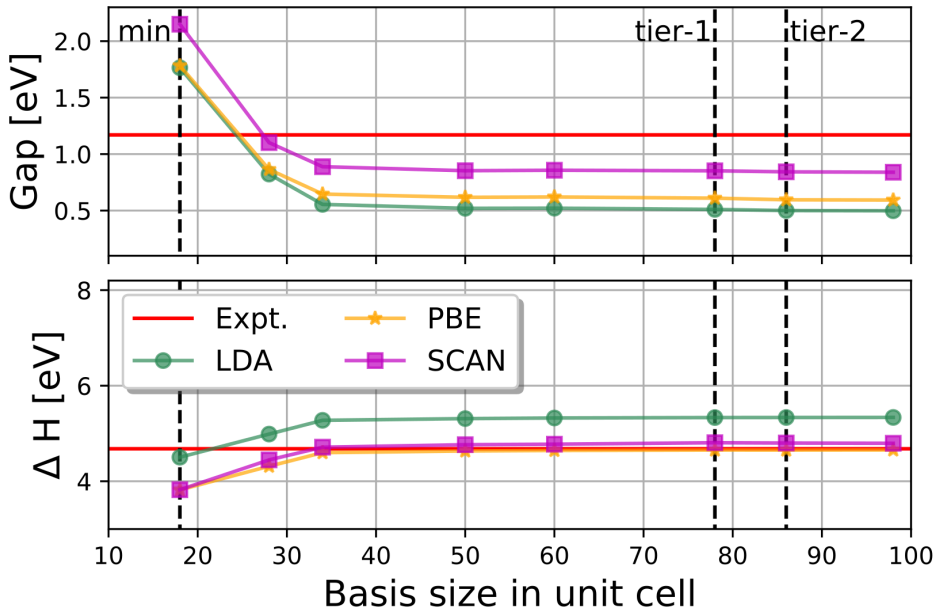

Figure 3.3: The convergence of the band-gap energy and the cohesive energy ΔH of fcc Si as a function of basis set size. LDA, PBE and SCAN are calculated in *FHI-aims* using $10\times10\times10$ **k**-grids and *tight* numerical settings. The red line presents the experimental band-gap (top) and cohesive energy (bottom) [112]. The black dash lines, marked with min and tier-1,2, denote the predefined *minimal* and *tier*-1,2 basis sets, respectively.

molecules mentioned above, SCAN has the same basis-set convergence as LDA and PBE for solids. The *tier*-1 has been good enough to reach the complete basis-set limit for calculating both band-gaps and cohesive energies. Fig. 3.4 shows that the time costs per iteration of LDA, PBE, and SCAN in the periodic systems have the same scaling factor $O(N_\mathrm{b}^{2.29})$. The scaling factor in the periodic systems is different from the linear scaling behavior in the finite systems because the Hartree potential is caused by summing over all unit cells in periodic boundaries. However, it has been determined that the computational cost of solving the KS one-electron equations scales as $O(N_\mathrm{b}^3)$ for finite and periodic systems, which stands out among other calculation steps to be the major cost for large systems (i.e., containing more than 1000 atoms) [113, 114]. It has been proven that this heavy computational step can be speeded up to very large core counts [115] by the Eigenvalue Solvers for Petaflop Application (*ELPA*) [114]. So far, *ELPA* has been used in many electronic structure communities, including *FHI-aims*.

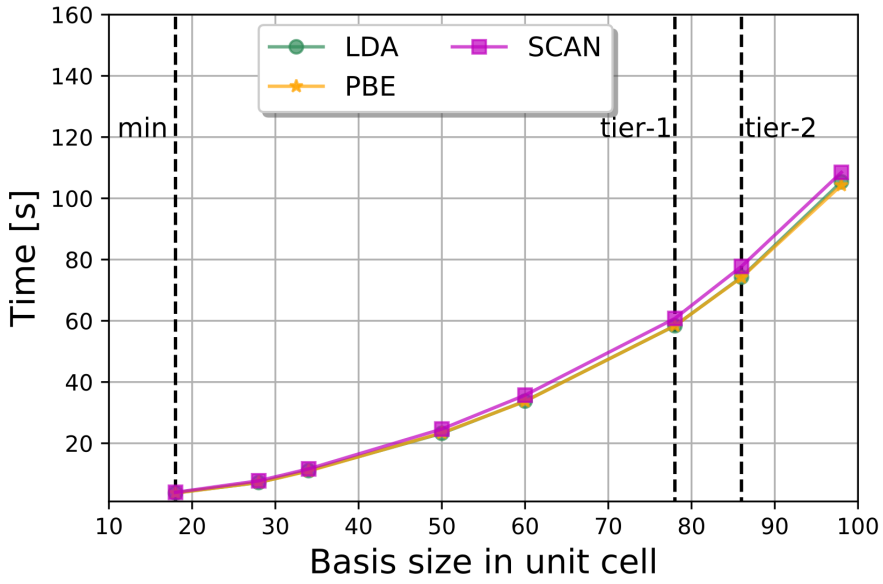

Figure 3.4: The time cost per iteration for the LDA, PBE and SCAN calculations for fcc Si as a function of basis size in unit cell. LDA, PBE and SCAN are calculated in *FHI-aims* using $10 \times 10 \times 10$ **k**-grids and *tight* basis sets. The black dash lines, marked with min and tier-1,2, denote the predefined *minimal* and *tier*-1,2 basis sets, respectively. The time scales as $O(N_{\mathrm{b}}^{2.29})$ for LDA, PBE and SCAN calculations in a similar manner.

3.2 Integration Grid

In practice, *FHI-aims* numerically solves the integrations in Eqs. (3.7) and (3.11) on partitioned grids. It has been proved that the numerical integration implementation achieves a nearly linear-scaling performance for large systems [20]. This advantage can be traced to the Stratmann atom-center partition algorithm [116] for LDA and GGAs realized in *FHI-aims* for the sparse grid operation [117, 118]. Moreover, it was demonstrated that this partition algorithm needs fewer grid points for the NAO basis sets than analytical Gaussian-type orbital (GTO) basis sets to converge electronic total-energy calculations in *FHI-aims*. One explanation is that NAOs do better than GTOs describe the area very close to the nucleus [110]. Therefore I employed the Stratmann grid scheme and NAO basis keeping the same treatment used for LDA and GGAs to study the grid sensitivity of the SCAN functional in *FHI-aims*.

The standard basis-set pool of *FHI-aims* provides a hierarchy of predefined settings for integration grids, namely, *light*, *tight*, and *really-tight*. The accuracy of basis sets is simultaneously increased by the following aspects in the program [20]: 1) number of shells along radial direction; 2) position of the outermost shell; 3) Hartree potential expansion; and 4) the divisions in one shell used for

three-dimensional integrations. (The control technics behind these aspects are described in Ref. [20].) It has been reported that the SCAN calculation for Ge

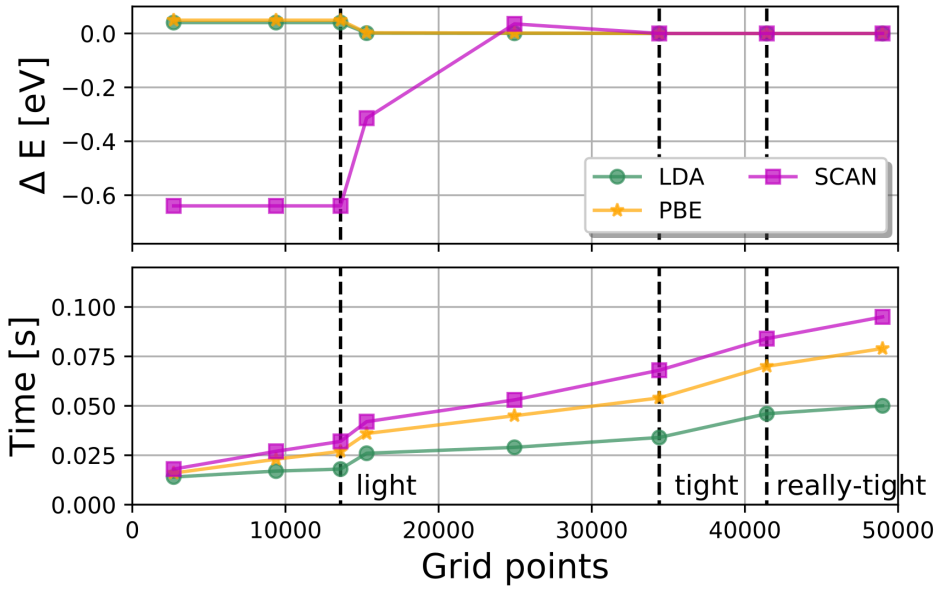

Figure 3.5: (top) The numerical convergence of the total energies of the Ge atom with respect to the predefined grid settings of *light*, *tight* and *really-tight*. (bottom) The performances of LDA, PBE, and SCAN methods are studied, for which the corresponding time costs per iteration are also provided. ΔE represents the total energy difference with the energy using "*dense*" grid settings, which denotes a grid setting 25 times denser than the standard *tight* setting.

atom is very sensitive to the choice of grids [108]. Fig. 3.5 presents the total energies of the Ge atom self-consistently calculated by LDA, PBE, and SCAN to study the numerical convergence with respect to different predefined grid settings. We can see that the *tight* predefined grid setting is enough to converge the Ge total energy for all three semi-local DFAs. However, the SCAN functional suffers from a notable slower convergence than LDA and PBE. Specifically, the error of SCAN is larger than 0.6 eV by using the *light* grid setting. For comparison, it is about 0.03 eV for LDA and PBE.

To trace the origin of the slow convergence behavior, I plotted the XC potentials along the radius distance on Fig. 3.6, which were evaluated based on the self-consistent KS orbitals of PBE and SCAN themselves, respectively. It is apparent from this figure that the GGA PBE method possesses a smooth XC potential, which can be properly described by using the *light* grids. For comparison, the SCAN XC potential is much more rugged, and a denser grid setting is required to sample this behavior so to converge the numerical integration over the potential. Although the grid sensitivity of SCAN can be ascribed to its rugged XC potential, closer inspection of Fig. 3.6 suggests that the SCAN

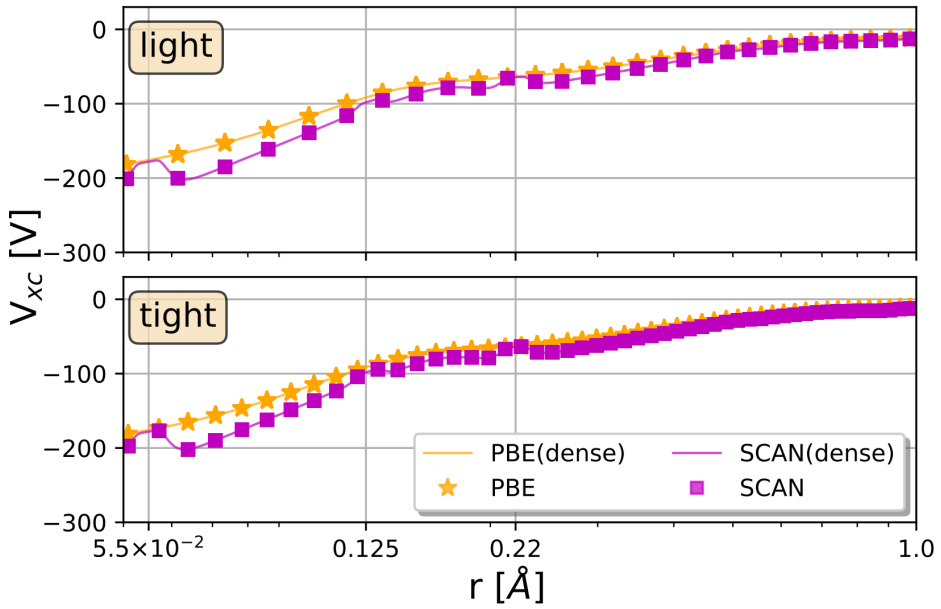

Figure 3.6: The XC potential along the radius distance evaluated based on self-consistent PBE and SCAN KS orbitals. "*dense*" denotes a grid setting 25 times denser than the standard *tight* setting.

XC potential calculated in my investigation is much less rugged than those reported in the previous research by Furness and Sun in 2019 [108]. In the next section, I presented the differences and proposed a meta-GGA-specific mixing algorithm, which was demonstrated to be efficacious to get the XC potential for the standard SCAN functional.

3.3 Kinetic Energy Density Mixing Algorithm

For deeply figuring out the problem detected in the last section, the SCAN XC potential of the Ge atom was re-calculated based on the self-consistent PBE orbitals (denoted as "SCAN@PBE"). It was plotted in Fig. 3.7 together with the SCAN and PBE potentials using their own self-consistent KS orbitals. The dimensionless kinetic energy α (defined in Eq. (2.39)) and density were also plotted in Fig. 3.8 and Fig. 3.9, respectively. α(SCAN@PBE) or α(SCAN) represents the α based on the self-consistent PBE or SCAN orbitals in the following.

As shown in Fig. 3.7, the SCAN@PBE XC potential brings about sharp oscillations, identical to the report of Furness and Sun [108]. The ill-behaved points on the SCAN@PBE XC potential curve are observed near the approximated orbital radius defined in the Bohr model. However, taking Fig. 3.8 and Fig. 3.7 together, we can find that the ill-behaved points exactly match the region of α(SCAN@PBE) approaching 1. Furthermore, the region of α(SCAN) ≈ 1 also

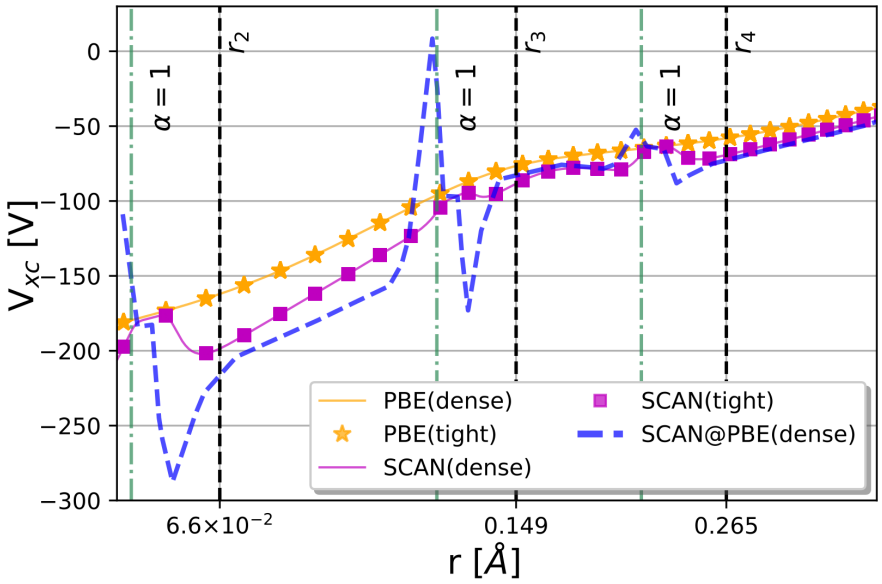

Figure 3.7: XC potentials of PBE and SCAN along the radius distance for the Ge atom. "*dense*" denotes a grid setting 25 times denser than the standard *tight* setting. The SCAN potential evaluated by the PBE self-consistent KS orbitals is labeled as "SCAN@PBE". The vertical dashed lines indicate the approximated orbital radius defined in the Bohr model ($r_i = \frac{0.53}{32} i^2$). The vertical dot-dashed lines indicate the positions where the dimensionless kinetic energy α of SCAN@PBE is equal to 1.

raises weak ruggedness in the SCAN XC potential based on the self-consistent SCAN orbitals, which explains the failure of SCAN in describing the energy using *light* basis in the case of the Ge atom (shown in Fig 3.5). The area of $\alpha \approx 1$ describes the asymptotic regions of the non-covalent density [57]. The recent work of Furness *et al.* in 2021 has documented that $\alpha(r)$ can have diverged scaling behaviors of r when $\alpha(r)$ closes 1. The improper scaling behavior destroys the overall smoothness of the functional and introduces oscillations into the XC potential. A series of works have been proposed to reform α [19, 105, 109, 119]. In this work, we argued that the self-consistent calculations of SCAN itself finely adjust α and can result in a less rugged XC potential. From Fig. 3.9 to Fig. 3.7, we see that the almost overlapping (kinetic energy) densities of PBE and SCAN can lead to the α(SCAN) and α(SCAN@PBE) with slight differences. The differences are further enlarged by the SCAN XC potential and are expressed in Fig. 3.7. Likewise, for any given set of trial KS orbitals, the use of τ-dependent dimensionless parameter α in the SCAN functional poses a high risk of yielding an ill-behaved and nonphysical XC potential. It was worth trying a meta-GGA-specific mixing algorithm to find the self-consistent SCAN orbitals and further converge the SCF iteration fast and stable for the standard

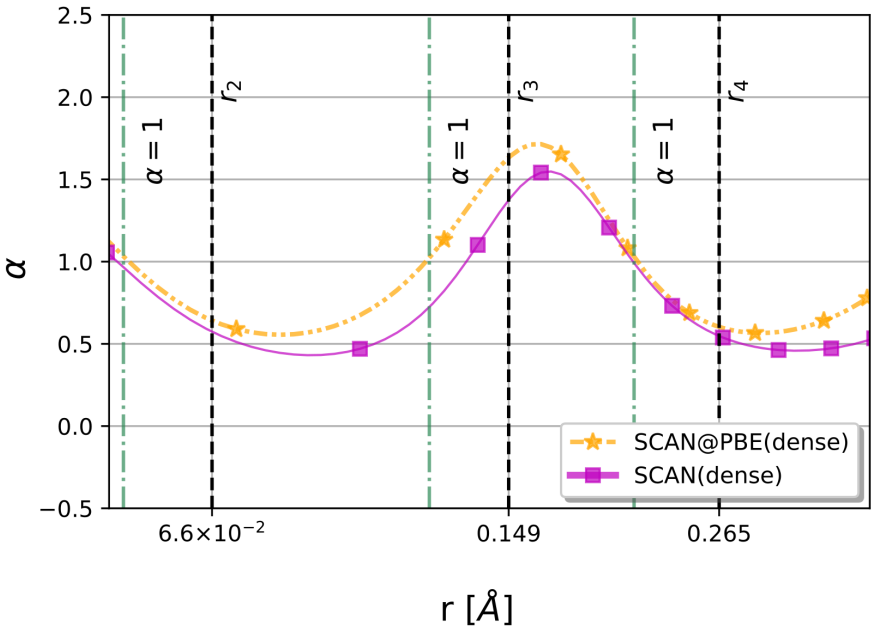

Figure 3.8: Dimensionless kinetic energy α (defined in Eq. (2.39)) of SCAN@PBE and SCAN along the radius distance for the Ge atom. "*dense*" denotes a grid setting 25 times denser than the standard *tight* setting. The α evaluated by the PBE self-consistent KS orbitals is labeled as "SCAN@PBE". Each marker on the curve of α is plotted every 1000 grid points along the radius distance. The vertical dashed lines indicate the approximated orbital radius defined in the Bohr model ($r_i = \frac{0.53}{32} i^2$). The vertical dot-dashed lines indicate the positions where the dimensionless kinetic energy α of SCAN@PBE is equal to 1.

SCAN functionals.

In *FHI-aims*, the fast and stable SCF convergence of the LDA and PBE calculations is achieved using Pulay's direct inversion of the iterative subspace scheme for the (gradient) density mixing [20, 120]. The example of the Ge atom has determined that inconspicuous changes in the kinetic energy density can lead the oscillations in the SCAN XC potential. In consequence, only counting (gradient) density in the mixing algorithm might not be enough to stabilize the SCF convergence for the SCAN functional. In order to make the kinetic energy density τ to be self-consistent with the τ-dependent meta-GGA XC functionals, I developed a new mixing algorithm of $\tau(\mathbf{r})$ in *FHI-aims* and applied it together with the standard density mixing algorithm during the SCF procedure for meta-GGAs.

Assuming that the kinetic energy density in the μth SCF iteration step is τ_{in}^{μ}, the Hamiltonian matrix can be established based on it together with the electron

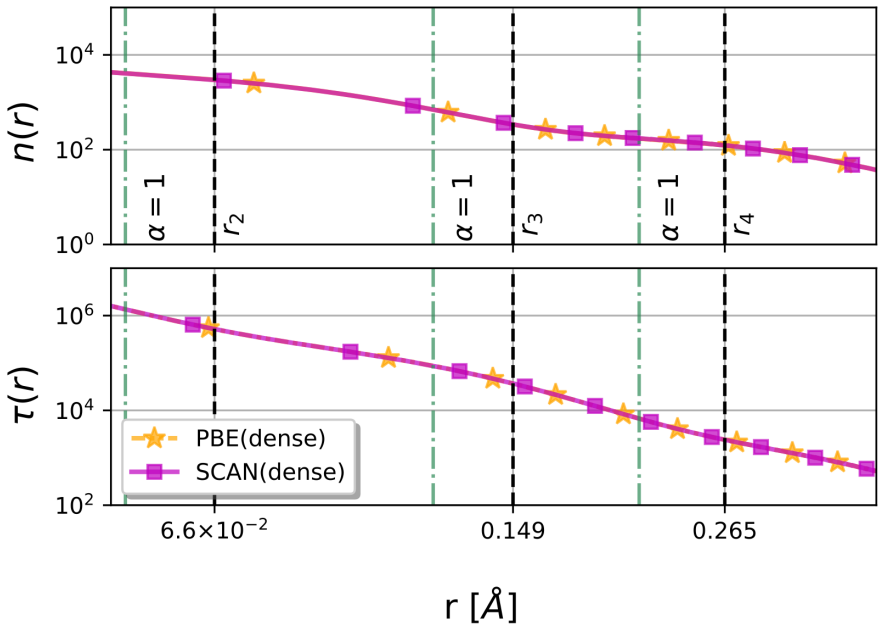

Figure 3.9: Density $n(r)$ and kinetic energy density $\tau(r)$ of the self-consistent PBE and SCAN orbitals along the radius distance for the Ge atom. "*dense*" denotes a grid setting 25 times denser than the standard *tight* setting. Each marker on the curve of (kinetic energy) density is plotted every 1000 grid points along the radius distance.

density and orbitals in the same iteration step (See Eqs. (2.19), (2.20), (2.37), and (3.7)). By solving the KS equations, we obtained the updated kinetic energy density τ^{μ}_{out} (See Eqs. (2.30) and (3.5)). The residual functional between the input and output kinetic energy density is defined as

$$R[\tau^{\mu}_{\text{in}}] = \tau^{\mu}_{\text{out}}[\tau^{\mu}_{\text{in}}] - \tau^{\mu}_{\text{in}} \quad . \tag{3.14}$$

The residual functional is zero, if input kinetic energy density τ^{μ}_{in} has been converged in SCF procedures. Pulay mixing scheme uses the converged condition [120], in which the residual functional depends linearly on a series of historical input kinetic energy densities

$$R[\sum_{\mu} \alpha_{\mu}\tau^{\mu}_{\text{in}}] = \sum_{\mu} \alpha_{\mu} R[\tau^{\mu}_{\text{in}}] \quad , \tag{3.15}$$

where $\{\alpha_{\mu}\}$ minimize the norm of the residual functional $\langle R[\sum_{\mu}\alpha_{\mu}\tau^{\mu}_{\text{in}}]|R[\sum_{\mu}\alpha_{\mu}\tau^{\mu}_{\text{in}}]\rangle$ under the normalization constraint of $\sum_{\mu}\alpha_{\mu} = 1$. Such kind of minimization determines $\{\alpha_{\mu}\}$ to be with the form of

$$\alpha_{\mu} = \frac{\sum_{v}\langle R[\tau^{v}_{\text{in}}]|R[\tau^{\mu}_{\text{in}}]\rangle^{-1}}{\sum_{vw}\langle R[\tau^{v}_{\text{in}}]|R[\tau^{w}_{\text{in}}]\rangle^{-1}} \quad . \tag{3.16}$$

The resulting optimal kinetic energy density $\tau^{\mu}_{\text{opt}} = \sum_{\mu}\alpha_{\mu}\tau^{\mu}_{\text{in}}$ was then utilized to generate the input kinetic energy density in the $(\mu+1)$-iteration via the

equation of

$$\tau_{\text{in}}^{\mu+1} = \tau_{\text{opt}}^{\mu} + \hat{G}^1 R[\tau_{\text{opt}}^{\mu}] \quad . \tag{3.17}$$

Here, the operator $\hat{G}^1$ is a special preconditioning matrix, e.g., the non-local matrix proposed by Kerker in 1981 [121].

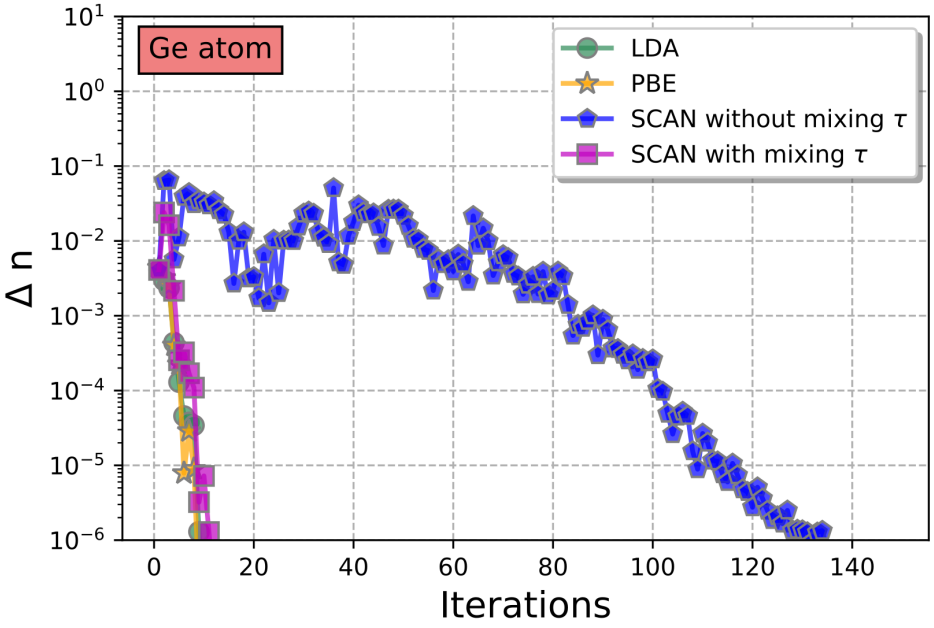

Figure 3.10: The density residual of Ge atom SCAN with and without kinetic energy density mixing as a function of the iteration steps in SCF calculations. The density residual is calculated by integrating the difference of the densities in the μth and $(\mu-1)$th SCF iteration step: $\Delta n = \int d^3\mathbf{r}(n^{\mu}(\mathbf{r}) - n^{\mu-1}(\mathbf{r}))$. The SCF procedure is stopped when the density residual drops to 10^{-6}. The results of LDA and PBE are given as references. The SCAN total energy with τ-mixing differs from the energy without τ-mixing less than 10^{-7} eV.

I have implemented the aforementioned mixing algorithm for the kinetic energy density in *FHI-aims*. To better understand the outcome of mixing $\tau(\mathbf{r})$ in *FHI-aims*, I firstly invested the SCF convergence behavior of LDA, PBE, and SCAN for Ge atom and also plotted the SCAN calculations with/without mixing $\tau(\mathbf{r})$ in Fig. 3.10. In this case, the SCAN calculations without mixing τ converged over 130 iterations, but LDA and PBE converged fast within 20 steps. The mixing algorithm of $\tau(\mathbf{r})$ speeds up the SCAN convergence (within 20 steps), with the same behavior as LDA and PBE. This result confirms that updating the kinetic energy density directly during SCF iterations indeed undermines the convergence behavior for SCAN calculations, and the mixing algorithm of $\tau(\mathbf{r})$ can obviously weaken this impact in this simple case.

Besides atoms, the calculations for the solid fcc Si has been reported to have

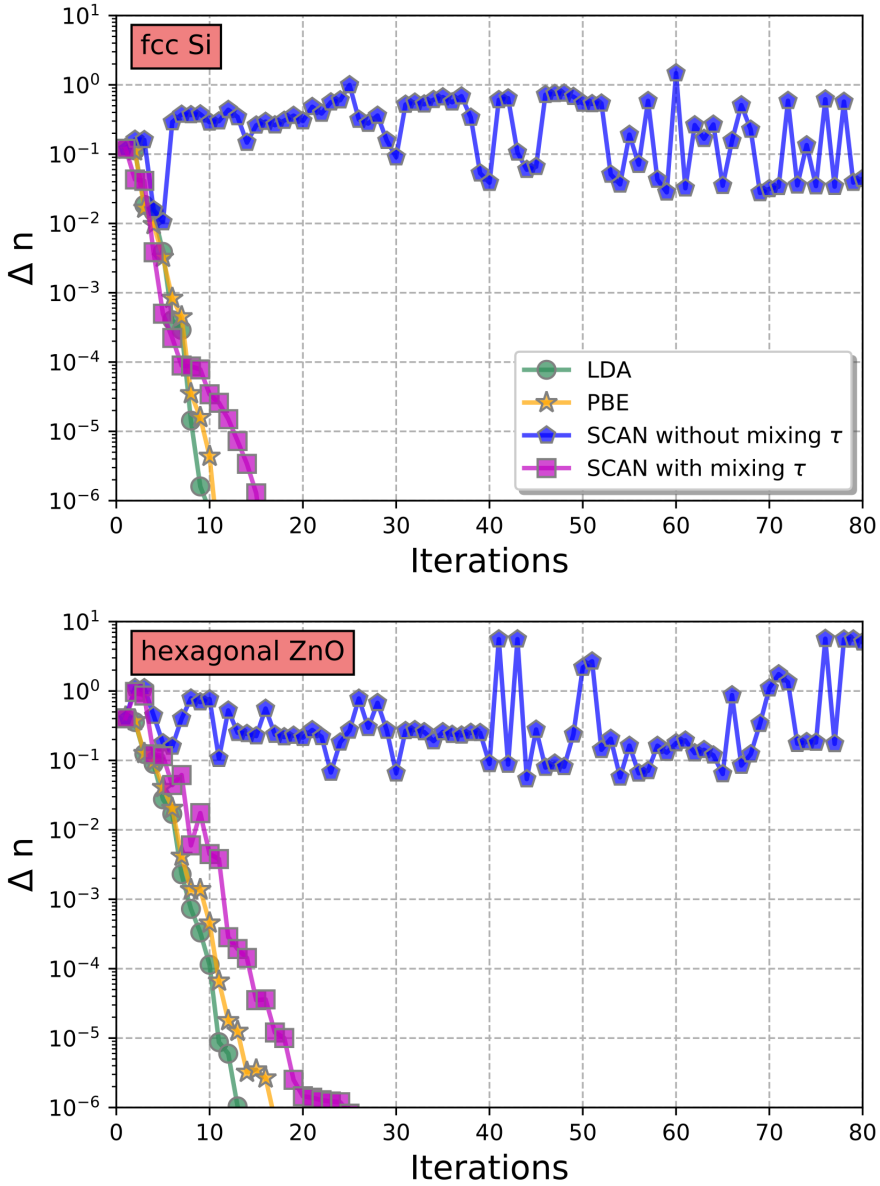

Figure 3.11: The density residual of the non-spin fcc Si (top) and hexagonal ZnO (bottom) unit cells using SCAN with and without kinetic energy density mixing as a function of the iteration steps in SCF calculations. The density residual is calculated by integrating the difference of the densities in the μth and $(\mu-1)$th SCF iteration step: $\Delta n = \int d^3\mathbf{r}(n^{\mu}(\mathbf{r}) - n^{\mu-1}(\mathbf{r}))$. The SCF procedure is stopped when the density residual drops to 10^{-6}. The results of LDA and PBE are given as references. The $3\times3\times3$ **k**-grids and *tight tier*-1 NAO basis sets are used for all the calculations.

a convergence problem in the early SCAN implementation [17]. In further investigations of solids, I found the SCAN functional encounters a widely serious SCF convergence problem. To better illustrate the SCF convergence problem,

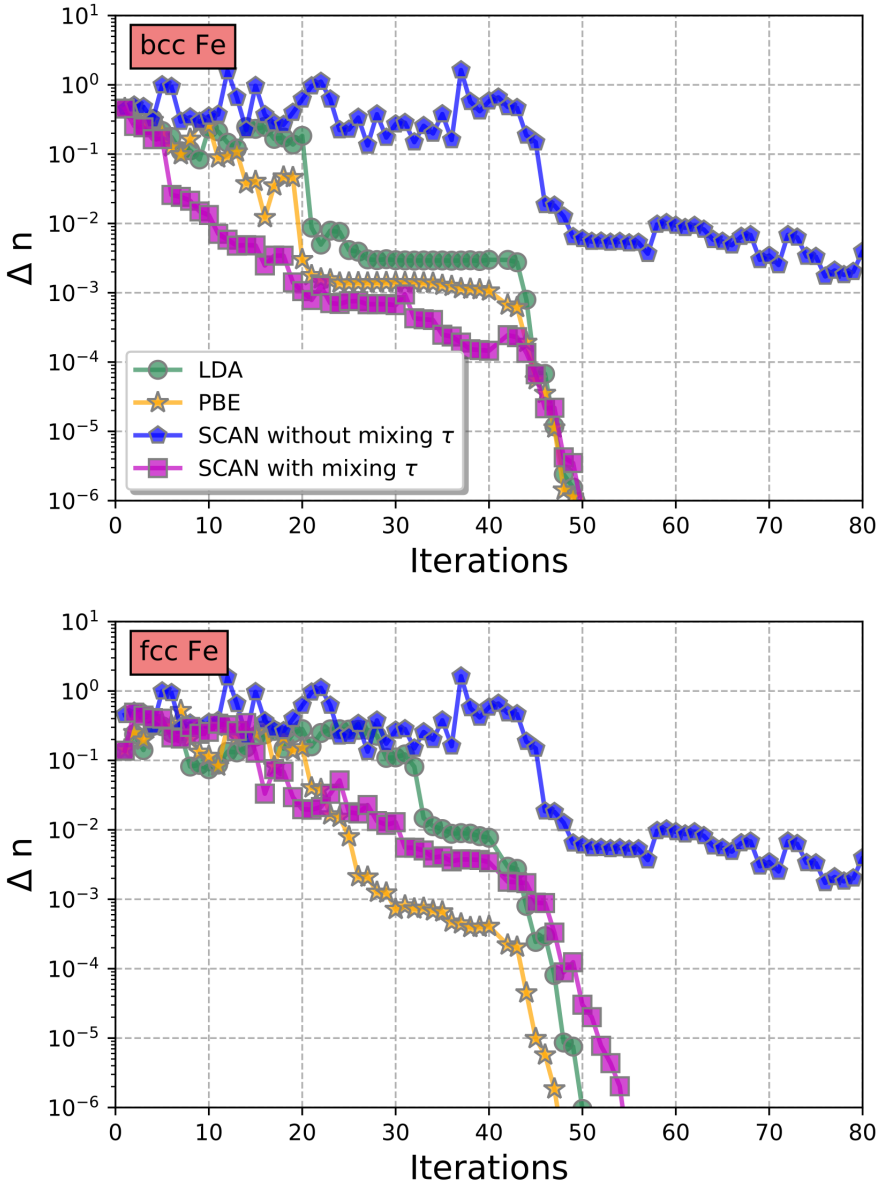

Figure 3.12: The density residual of the magnetic bcc Fe (top) and fcc Fe (bottom) unit cells using SCAN with and without kinetic energy density mixing as a function of the iteration steps in SCF calculations. The density residual is calculated by integrating the difference of the densities in the μth and $(\mu-1)$th SCF iteration step: $\Delta n = \int d^3\mathbf{r}(n^{\mu}(\mathbf{r}) - n^{\mu-1}(\mathbf{r}))$. The SCF procedure is stopped when the density residual drops to 10^{-6}. The results of LDA and PBE are given as references. The spin-polarized, $3 \times 3 \times 3$ **k**-grids and *tight tier*-1 NAO basis sets are used for all the calculations.

Figs. 3.11 and 3.12 present the SCF convergence behavior of LDA, PBE, and SCAN for the crystals of fcc Si, hexagonal ZnO, bcc Fe, and fcc Fe.

For all the four cases, the SCAN calculations without mixing τ cannot converge within 100 iterations. For comparison, 20 iterations are enough to converge the LDA and PBE calculations for the non-magnetic systems (fcc Si and hexagonal ZnO), as shown in Fig. 3.11. More than 50 iterations are needed to converge in LDA and PBE calculations for the magnetic systems (bcc Fe and fcc Fe). We can see that the SCAN functional with mixing τ now exhibits the same convergence behavior as LDA and PBE for all the cases, including the non-magnetic fcc Si and hexagonal ZnO, and the magnetic bcc Fe and fcc Fe.

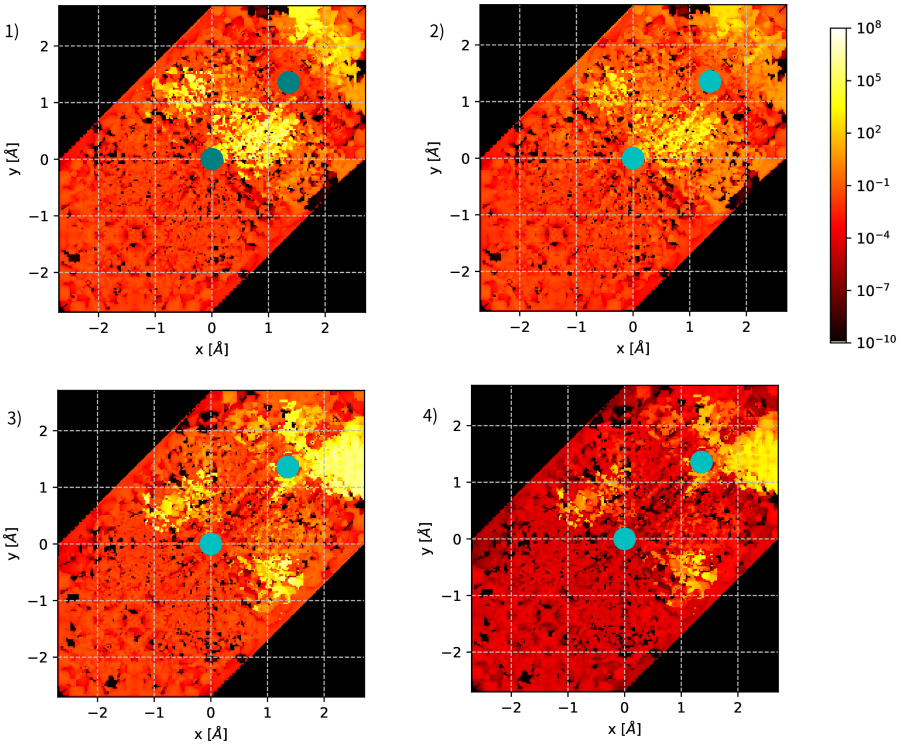

Figure 3.13: The evolution of kinetic energy density τ on *tight* grids in the SCAN SCF procedure for fcc Si. 1) The value of τ in the 1st SCF iteration, which is same for the procedures with and without mixing τ. 2) The value of τ evolved without mixing itself in the 40th SCF iteration. The value of τ evolved with mixing itself in the 3) 3rd and 4) 16th SCF iteration. All density maps are limited in the unit cell and are viewed from the z-axis. The blue circles present the positions of Si atoms. The calculation is already converged in the 16th SCF iteration with mixing τ (See Fig. 3.11). The $3 \times 3 \times 3$ **k**-grids and *tier*-1 NAO basis sets are used for all the calculations.

To trace the convergence failure for solids, I plotted the τ evolution in the SCAN SCF procedure of fcc Si in Fig. 3.13. Specifically, Fig. 3.13-2) determines that the value of τ in the 40th step still oscillates around the value in the first step if only the (gradient) density is mixed in the SCAN calculation. Figure 3.13-3) displays that the positive effect of mixing τ can emerge from the 3rd step, because the value of τ at this step has approached the final converged value plot-

ted in Figure 3.13-4). These results confirm that the SCF convergence problem reported in the previous work [108] is not inherent to the SCAN functional. I demonstrated that the SCAN implementation in *FHI-aims* can converge as fast and stable as LDA and PBE. Now the SCAN calculation for solids (like fcc Si, hexagonal ZnO, bcc Fe and fcc Fe) is expected as 1.26 times the total cost of PBE in *FHI-aims*, but was reported in 2020 as 4.94 times of the PBE cost in other implementations [122].

In the next section, I benchmarked my SCAN implementation in *FHI-aims* with mixing $\tau(\mathbf{r})$ by using 1) 24 HOMO energies of the atoms from H to Ar, 2) 76 reaction barriers from the BH76 test set [123, 124], and 3) 10 fundamental band-gaps from a subset of the reference [6] and reference [125]. The following calculations of LDA, PBE and SCAN in *FHI-aims* use the *tight* basis-sets with *tier*-1, which are analysed in the Sec. 3.1 and Sec. 3.2 to have well convergence.

3.4 Benchmark SCAN in *FHI-aims*

Since the birth of the SCAN functional in 2015, the SCAN functional has been used and analysed in many works [6–10]. Now SCAN functional is the promising meta-GGA functional for both molecules and solids. However, the SCAN can not be widely applied because of the convergence problem, especially for solids (See Sec. 3.3). I pick up three confirmed subsets, including HOMO energies of the atoms, reactions barriers and band-gaps. In this section, these subsets are used to determine the correctness of the SCAN implementation in *FHI-aims* rather than discuss the performance of SCAN functional.

We compare the LDA, PBE and SCAN HOMO energies of atoms from H to Ar against the experimental IPs [77] in Fig. 3.14. The figure shows that the calculated points of SCAN are closer than LDA and PBE to perfect line. But all calculations of DFAs (including LDA, PBE and SCAN) have large distance to the perfect line and hence heavily underestimate the IPs. Table 3.1 shows that the MAEs (e.g., mean absolute errors) of the orbital eigenvalues are 4.92, 4.90, and 4.51 eV for LDA, PBE, and SCAN, respectively. The corresponding results, taken from Ref. [11] and calculated by the naval research laboratory molecular orbital library (*NRLMOL*) [126–128], are also included in Table 3.1 for comparison. The *NRLMOL* results are obtained by the default *NRLMOL* basis sets, and a finer grid dense, typically containing 25000 grid points per atom. The MAEs between *FHI-aims* and *NRLMOL* for LDA, PBE, and SCAN calculations are all within a 0.03 eV difference.

In Fig. 3.15, we compare the barrier heights of BH76 for LDA, PBE, and SCAN calculations against the Weizmann-1 (W1) calculations. Here the BH76 test set comprises 76 barrier heights for chemical reactions [123, 124], and W1 is high-accuracy computational thermochemistry designed to extrapolate to the complete basis limit of a CCSD(T) [129, 130]. SCAN and PBE have significant improvements over LDA for the predicted barrier heights of the BH76 test

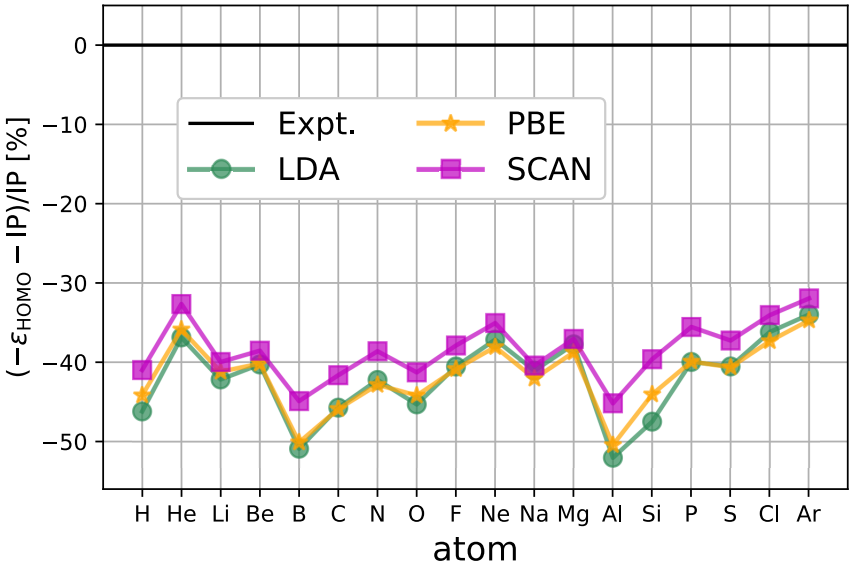

Figure 3.14: Comparison of IPs and the calculated $-\epsilon_{\mathrm{HOMO}}$ energies by LDA, PBE, and SCAN of H to Ar atoms. The relative errors of $-\epsilon_{\mathrm{HOMO}}$ energies to IPs of H to Ar atoms are given using LDA, PBE, and SCAN. The *tight tier*-1 NAO basis sets are used for all calculations. The black line presents the experimental IPs [77].

[eV]	FHI-aims		NRLMOL	
	ME	MAE	ME	MAE
LDA	-4.92	4.92	-4.89	4.89
PBE	-4.90	4.90	-4.88	4.88
SCAN	-4.51	4.51	-4.54	4.54

Table 3.1: Deviations of $-\epsilon_{\mathrm{HOMO}}$ from the corresponding experimental ionization potentials for H to Ar atoms calculated by LDA, PBE and SCAN in *FHI-aims*. Mean errors (MEs) and mean absolute errors (MAEs) are given in eV. The *FHI-aims* results use the *tight* basis-sets with *tier*-1. The *NRLMOL* results are given as a comparison and taken from Ref. [11], The default *NRLMOL* basis set and a finer grid dense, typically containing 25000 grid points per atom, are used in Ref. [11].

set. At the same time, the differences in the prediction accuracy for PBE and SCAN are less noticeable. In Table 3.2, the MAEs of LDA, PBE, and SCAN calculations are given as 0.66, 0.40 and 0.33 eV (15.2, 9.3, and 7.8 kcal/mol), respectively, which match the MAEs shown in the original SCAN work [7] within 0.005 eV (0.1 kcal/mol). The results shown in the original SCAN work are calculated using *Gaussian* code [131] with 6-311++G (3df, 3pd) basis set.

Figure 3.16 compares the computed band-gaps against the experimental values

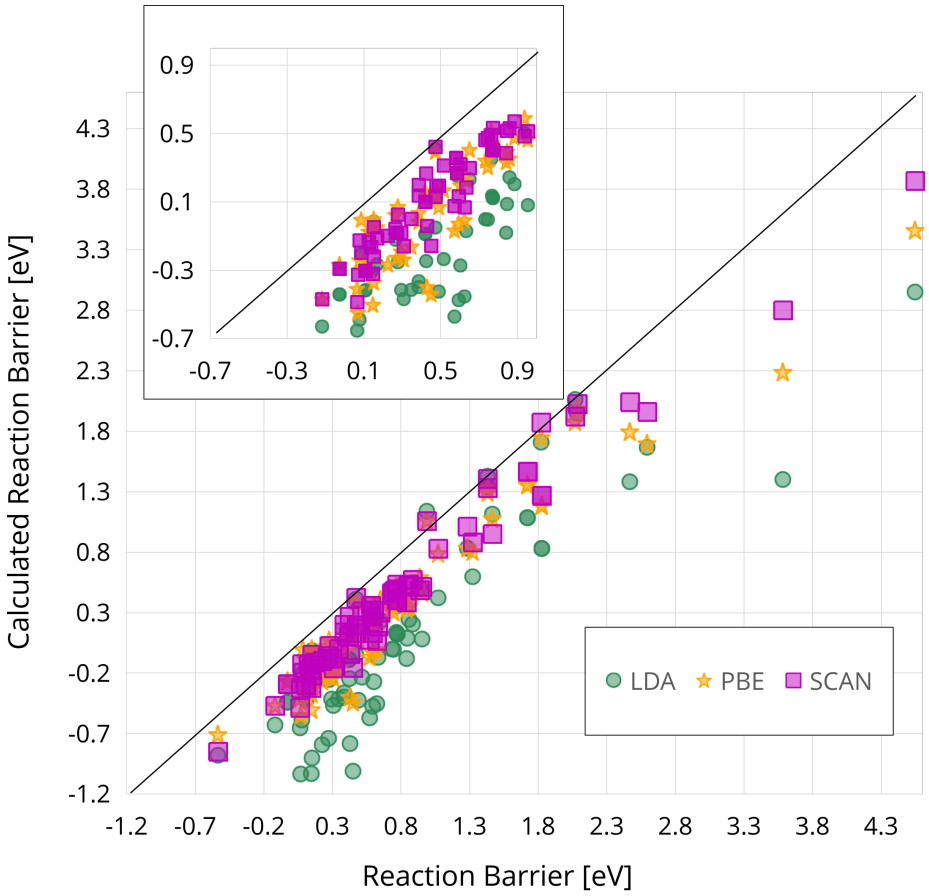

Figure 3.15: Comparison of the calculated results by LDA, PBE and SCAN and results by the Weizmann-1 (W1) method [129, 130] for the reactions included in BH76 [123, 124]. The black diagonal line corresponds to the calculated line of perfect agreement. *Tight tier*-1 NAO basis sets are used for LDA, PBE and SCAN calculations in *FHI-aims*. (1 kcal/mol = 0.0434 eV)

[eV]	FHI-aims		Gaussian	
	ME	MAE	ME	MAE
LDA	-0.655	0.660	-0.660	0.668
PBE	-0.399	0.404	-0.395	0.399
SCAN	-0.334	0.339	-0.334	0.334

Table 3.2: Mean error (ME) and mean absolute error (MAE) of the calculated chemical barrier heights given in eV for the BH76 test set [123, 124] calculated by LDA, PBE and SCAN in *FHI-aims*. The *FHI-aims* results use the *tight* basis-sets with *tier*-1. The *Gaussian* results with the 6-311++G (3df, 3pd) basis sets are given as a comparison and taken from Ref. [7]. (1 kcal/mol = 0.0434 eV)

of a test set with 10 fundamental solids. It is worth noting that the electron-phonon interaction can impact band-gaps at absolute zero temperature due to the zero-point vibrations [71], but the exact effect is difficult to access from experiments [132]; thus, we ignore it when calculating band-gaps in this thesis. More theoretical studies for this effect on band-gaps can be found in Ref. [133, 134]. The test set includes covalent crystals (such as Si, C), ionic crystals (such as NaCl, LiF), and also Mott insulators composed of transition-metal oxides (MnO). Nearly, all the calculated points by DFAs (including LDA, PBE and SCAN) lie below the dashed line of the perfect agreement. The band-gaps of SCAN calculations have a small MAE (2.29 eV), compared with the MAE of LDA (3.24 eV) and PBE (3.04 eV). These results echo the work of Isaacs and Wolverton, in which SCAN has moderate improvements (more than 20% decrease in MAE) compared to PBE [6]. For LiF (with a very large experimental gap of 14.20 eV) as an example, the SCAN band-gap is 0.98 eV larger than that of PBE using FHI-aims, and the enhancement of SCAN for PBE is 1.10 eV in the Isaacs and Wolverton work using the Vienna *ab initio* software package (*VASP*) [135–138] with a 600-eV plane-wave cutoff.

In this chapter, I verified that my implementation in *FHI-aims* with mixing $\tau(\mathbf{r})$ had overcome the worrisome numerical instability in SCAN functional and can correctly give calculated results. However, in the three benchmark cases, only slight improvements of SCAN exist from PBE comparing with the experimental or reference values. These extensive derivations of the HOMO energies, reaction barriers, and band-gaps for the semi-local DFAs, including SCAN, are caused by the notorious SIEs. In the next chapter, I developed and implemented a self-consistent SIC method in *FHI-aims* for solving this problem.

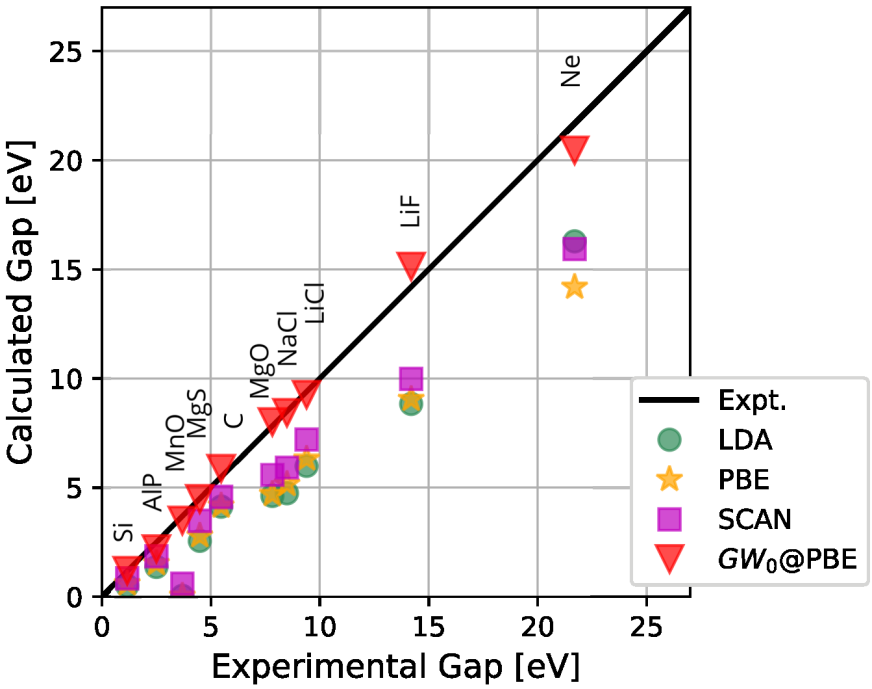

Figure 3.16: Comparison of calculated and experimental experimental band-gaps for LDA, PBE, SCAN and GW_0@PBE. The black diagonal line corresponds to the calculated line of the perfect agreement. The **k**-grids $4 \times 4 \times 4$ and *tight tier*-1 NAO basis sets are used for the calculations of LDA, PBE and SCAN in *FHI-aims*. GW_0 and experiment results are from [125] and references within. The effect of the zero-point vibrations is not considered in this thesis. The theoretical studies for this effect on band-gaps can be found in Ref. [71, 133, 134].

4 Self-interaction Correction in *FHI-aims*

In the first part of my project, I have implemented the standard SCAN functional in *FHI-aims*, which features a fast and stable SCF convergence for both molecules and solids. In agreement with the previous works [6, 11], our results confirm the systematic tendency of the SCAN functional to underestimate the IPs, the reaction barrier heights, and the band-gaps. These common weakness has been closely linked to the notorious self-interaction error (SIE) of all semi-local DFAs, including LDA, PBE, and SCAN.

In Sec. 2.4.2, several prototype model systems have been used to illustrate the heavy SIE in the SCAN functional. As introduced in Sec. 2.4.2, the self-interaction correction (SIC) algorithm proposed by Perdew and Zunger (PZ-SIC) precisely corrects the SIE of any semi-local DFAs for one-electron systems and has shown to be an effectual remedy against the many-electron SIE in semi-local DFAs [24]. In the second part of my project, I implemented an efficient and numerically stable self-consistent PZ-SIC method in *FHI-aims* for both molecules and solids.

As mentioned in Sec. 2.4.2, the PZ-SIC strategy aims at removing all the one-electron SIEs in a given DFA [24]:

$$E^{\mathrm{SIC}}[\{n_i\}] = -\sum_{i}^{N_\mathrm{e}} \left(E_{\mathrm{xc}}^{\mathrm{DFA}}[n_i] + E_{\mathrm{es}}[n_i] \right) \quad , \tag{4.1}$$

where the self-electrostatic functional $E_{\mathrm{es}}[n_i]$ and the self-XC functional $E_{\mathrm{xc}}^{\mathrm{DFA}}[n_i]$ have the same formula of the standard Hartree term and the exchange-correlation functional, respectively, but depend on the density n_i of occupied orbitals rather than the total density n. The total energy then becomes as [24]

$$E^{\text{SIC-DFA}}[n, \{n_i\}] = E^{\mathrm{DFA}}[n] + E^{\mathrm{SIC}}[\{n_i\}] \quad . \tag{4.2}$$

Here $E^{\mathrm{DFA}}[n]$ is the total energy of a given DFA in the standard KS framework [36]

$$E^{\mathrm{DFA}}[n] = T_{\mathrm{s}}[n] + E_{\mathrm{ext}}[n] + E_{\mathrm{es}}[n] + E_{\mathrm{xc}}^{\mathrm{DFA}}[n] \quad , \tag{4.3}$$

with T_{s} being the kinetic energy of the KS non-interacting system (defined in Eq. (2.11)). $E_{\mathrm{ext}}[n]$ and $E_{\mathrm{es}}[n]$ are the external potential energy and the electrostatic energy (the Hartree energy), respectively (defined in Eqs. (2.6) and (2.12)), both of which are an explicit functional of the total density n. $E_{\mathrm{xc}}^{\mathrm{DFA}}[n]$ is the exchange-correlation energy of the given DFA (defined in Eq. (2.13)).

As usual, the variational minimization of the SIC-DFA energy $E^{\text{SIC-DFA}}$ (Eq. (4.2)) should be subject to the conservation of the electron number N_e [24], which is equivalent to imposing an orthonormalization constraint on the canonical KS orbitals $\{\psi_l\}$ in the KS non-interacting systems (See Sec. 2.1):

$$\delta \left[E^{\text{SIC-DFA}} - \sum_{ij}^{N_\mathrm{e}} \epsilon_{ij} \left(\langle \psi_i | \psi_j \rangle - \delta_{ij} \right) \right] = 0 \quad . \tag{4.4}$$

However, the SIC total energy is no longer invariant under unitary transformations of the occupied orbitals. The dependence of $E^{\text{SIC}}[\{n_i\}]$ in Eq. (4.2) on the SIC orbital densities is problematic. In other words, the SIC orbitals $\{\phi_i\}$, to produce the lowest SIC-DFA energy, are not exactly necessary $\{\psi_l\}$ but a transformation of it [100]. The minimization of $E^{\text{SIC-DFA}}$ (Eq. (4.2)) is subject to another orthonormalization constraint on the SIC orbitals [30, 94], which results in the SIC Euler equation of

$$\delta\left[E^{\text{SIC-DFA}} - \sum_{ij}^{N_e} \epsilon_{ij}\left(\langle\psi_i|\psi_j\rangle - \delta_{ij}\right) - \sum_{ab}^{N_e} \lambda_{ab}\left(\langle\phi_a|\phi_b\rangle - \delta_{ab}\right)\right] = 0 \quad . \tag{4.5}$$

It has long been a problem to solve the SIC Euler equation and to find the feasible SIC orbital densities $\{n_i\}$ [24, 30, 100]. The following section introduced three constraints with clear physical meanings to guarantee the corresponding variation procedure locating the correct set of $\{n_i\}$.

4.1 Constraints in PZ-SIC

The first constraint discussed here is the total density constraint, i.e., the sum over all occupied SIC orbital densities $n_i(\mathbf{r}) = |\phi_i(\mathbf{r})|^2$ must recover the total electron density.

Constraint 1. Total Density Constraint :

$$\sum_{i}^{N_e} |\phi_i(\mathbf{r})|^2 = n(\mathbf{r}) \quad , \tag{4.6}$$

in which N_e denotes the total number of electrons in the system. As the derivation in Eq. (2.9), the ground-state density is also determined by the occupied KS orbitals

$$\sum_{l}^{N_e} |\psi_l(\mathbf{r})|^2 = n(\mathbf{r}) = \sum_{i}^{N_e} |\phi_i(\mathbf{r})|^2 \quad , \tag{4.7}$$

which becomes a bridge to the SIC orbital densities and KS orbital densities. Because of the orthonormalization conditions of SIC orbitals and KS orbitals

$$\begin{aligned} \delta_{ij} &= \langle\phi_i|\phi_j\rangle \quad , \\ \delta_{ml} &= \langle\psi_m|\psi_l\rangle \quad , \end{aligned} \tag{4.8}$$

the first constraint further restricts the occupied SIC orbitals to correspond to a unitary transformation of the occupied KS orbitals,

$$\begin{aligned} \phi_i &= \sum_{l}^{N_e} T_{il}\psi_l \quad , \\ \delta_{ml} &= \sum_{i}^{N_e} T_{im}^* T_{il} \quad . \end{aligned} \tag{4.9}$$

This constraint can confirm the exact SIC orbital for one-electron systems, where the SIC orbital is exactly the occupied KS orbital with an arbitrary phase $\phi_1 = \psi_1 e^{i\theta}$ [26]. For example, the hydrogen atom has one occupied orbital $1s$ and a lot of unoccupied orbitals such as $2s$, $2p$ and so on. The total density constraint restricts the SIC orbital of the hydrogen atom to be the $1s$ orbital $\phi_1 = 1s \cdot e^{i\theta}$, which yields the same density of the system $n = |1s|^2 = |\phi_1|^2$. Unfortunately, constraint 1 is incapable of confirming SIC orbitals for many-electron systems. Taking a model system with two occupied orbitals as an example, the canonical KS orbitals are $\{\psi_1, \psi_2\}$. By applying a unitary transformation operation on them, we can obtain another set of orbitals $\{\phi_1, \phi_2\}$:

$$\phi_1 = \frac{1}{\sqrt{2}}(\psi_1 + \psi_2); \phi_2 = \frac{1}{\sqrt{2}}(\psi_1 - \psi_2) \quad . \tag{4.10}$$

which satisfy the total density constraint and thus produce the exact same total energy for the DFA $E^{\mathrm{DFA}}[n]$, i.e., for the first part of Eq. (4.2)). By inserting these orbitals into the SIC energy (the second part of Eq. (4.2), and defined in Eq. (4.1)), one obtains

$$E^{\mathrm{SIC}}[\{\phi_1, \phi_2\}] \neq E^{\mathrm{SIC}}[\{\psi_1, \psi_2\}]. \tag{4.11}$$

Because of this uncertainty in constraint 1, Pederson *et al.* introduced the orbital potential constraint to facilitate the search for the SIC orbitals in 1988 [93].

Constraint 2. Orbital Potential Constraint :

$$\langle \phi_m | \hat{v}_m^{1e\mathrm{SIC}} | \phi_n \rangle = \langle \phi_m | \hat{v}_n^{1e\mathrm{SIC}} | \phi_n \rangle \quad . \tag{4.12}$$

Here the one-electron SIC potential $\{\hat{v}_i^{1e\mathrm{SIC}}\}$ comes from the variation of the self-electrostatic energy and self-exchange-correlation energy in gKS scheme and is the sum of the single-particle Hartree potential $\hat{v}_{\mathrm{es}}[n_i]$ and the single-particle XC potential $\hat{v}_{\mathrm{xc}}[n_i]$:

$$\hat{v}_i^{1e\mathrm{SIC}} = -\frac{\left(E_{\mathrm{xc}}^{\mathrm{DFA}}[n_i] + E_{\mathrm{es}}[n_i]\right)}{\delta n_i} = -(\hat{v}_{\mathrm{es}}[n_i] + \hat{v}_{\mathrm{xc}}[n_i]) \quad . \tag{4.13}$$

A detailed derivation of the given SIC potential is given in Appx. 7.2. Constraint 2 is derived from the full Euler equation of minimizing the SIC-DFA total energy (Eq. (4.5)) with fixing the total density.

Including Constraint 2, the SIC Euler equation is decoupled into two mutual parts in practice: 1) The first part solves the KS orbitals based on a given set of SIC orbitals $\{\phi_i\}$. 2) The second part in Eq. (4.12) then focuses on the solution of SIC orbitals $\{\phi_i\}$ with given KS orbitals. In the first part with a given set of SIC orbitals $\{\phi_i\}$, the Euler equation can be solved with respect to the electron density n and the KS orbitals $\{\psi_l\}$ only. It yields the SIC one-electron equations for the KS non-interacting systems

$$\hat{h}^{\text{SIC-DFA}}[n, \{\psi_l\}, \{\phi_i\}]\psi_l = \epsilon_l^{\text{SIC-DFA}}\psi_l \quad . \tag{4.14}$$

Here $\hat{h}^{\text{SIC-DFA}}$ is the KS non-interacting Hamiltonian of the SIC-DFA

$$\hat{h}^{\text{SIC-DFA}} = \hat{h}^{\text{DFA}} + \hat{v}^{\text{SIC}} \quad . \tag{4.15}$$

$\hat{h}^{\text{DFA}}$, defined in Eq. (2.20), is the KS one-electron Hamiltonian and includes the kinetic energy operator $\hat{t}_{\text{s}}$, the external potential $\hat{v}_{\text{ext}}$, the Hartree potential $\hat{v}_{\text{es}}$, and the XC potential $\hat{v}_{\text{xc}}$. Obviously, the only addition to the standard KS non-interacting Hamiltonian is the SIC potential operator $\hat{v}^{\text{SIC}}$, which is determined by the given set of SIC orbitals $\{\phi_i\}$

$$\hat{v}^{\text{SIC}} = \frac{\delta E^{\text{SIC}}[\{n_i\}]}{\delta n} = \sum_i^{N_{\text{e}}} \hat{v}_i^{1e\text{SIC}} |\phi_i\rangle\langle\phi_i| \quad . \tag{4.16}$$

By fixing the total density n, the Euler equation Eq. (4.5) simplifies to

$$\delta \left[E^{\text{SIC-DFA}} - \sum_{ab}^{N_{\text{e}}} \lambda_{ab} \left(\langle \phi_a | \phi_b \rangle - \delta_{ab} \right) \right] = 0. \tag{4.17}$$

Since the KS orbitals and density are fixed, the minimization of the total SIC-DFA energy (Eq. 4.2) equals to the minimization of the SIC energy itself (Eq. 4.1), resulting in the Euler equation of

$$\hat{v}_i^{1e\text{SIC}} \phi_i = \sum_a \lambda_{ai} \phi_a \quad . \tag{4.18}$$

By imposing the Hermitian property of the Lagrange multiplier $\lambda_{mn} = \lambda_{nm}^*$, we can obtain the orbital potential constraint

$$\langle \phi_m | \hat{v}_m^{1e\text{SIC}} - \hat{v}_n^{1e\text{SIC}} | \phi_n \rangle = \lambda_{nm}^* - \lambda_{mn} = 0. \tag{4.19}$$

Mathematically speaking, the full minimization of the SIC-DFA energy (Eq. (4.2)) implies solving the Euler equation with restrictions on both KS orbitals and SIC orbitals (Eq. (4.5)). This is, however, difficult to numerically find the KS orbitals and SIC orbitals simultaneously. Fortunately, this problem can be recast rigorously to the PZ-SIC one-electron equations (Eq. (4.14)) under Constraints 1 and 2 (Eqs. (4.6) and (4.12)), which can be implemented in the SCF manner of (g)KS framework.

An additional hurdle in solving SIC Euler equations (Eq. (4.5)) consists in finding the global minimization. Multiple sets of SIC orbitals has been documented to satisfy Constraint 1 and 2 even for atoms and simple diatomic molecules [93, 94]. In a vast amount of early SIC works [94], it has been determined that the SIC total energy for a many-electron system often tends to be lower when the SIC orbital densities are localized. This drives to search the right SIC orbitals from the vicinity of an initially given set of localized orbitals. Note that many localization methods have been proposed to generate the localized orbitals, such as Pipek-Mezey [139], Edmiston-Ruedenderb [31], von Niessen

[140], Foster-Boys [89], the fourth moment method [141], and so forth. Lehtola, Head-Gordon and Jónsson found that by using the SIC orbitals from different localization procedures, the SIC-PBE calculations for the total energy of Acrylic Acid differ from each other by more than 0.054 eV. Based on a comprehensive investigation for a set of molecules, they concluded that it is difficult to avoid the multiple-solution problem in the PZ-SIC calculations [30]. It is because there are too many local minima due to the unitary transformations between the KS orbitals and the SIC orbitals.

Instead of seeking for better localization procedure, Pederson, Ruzsinskzy and Perdew [95, 142] suggested generating the SIC orbitals by using Fermi-Löwdin orbitals (FLO) [143, 144] that are self-consistently localized in the space spanned by the KS orbitals. In FLO-SIC method, the transformation matrix in Eq. (4.9) is approximated by the single-particle density matrix T_{il}^F:

$$T_{il} \approx T_{il}^F = \frac{1}{n(\mathbf{r}=\mathbf{a}_i)} \cdot \psi_l^*(\mathbf{r}=\mathbf{a}_i) \tag{4.20}$$

with a set of Fermi orbital descriptors $\{\mathbf{a}_i\}$ obtained from the solution of Constraint 2 [96]. In this way, the SIC orbitals can keep localized and center on the descriptors [143, 145]. The transformation matrix T_{il} is obtained from T_{il}^F by the Löwdin orthonormalization. In the Löwdin orthonormalization, the self-consistent FLO-SIC method asks a careful rotation of T_{il} to avoid damage to Constraint 2 [97]. Pederson's constraint prevents consideration of all unitary transformations and thus simplifies the self-consistent solution of SIC orbitals. Despite not strictly solving the aforementioned multiple-solution problem, the FLO-based SIC orbitals have been demonstrated to be a good choice to evaluate the SIC-DFA energies, in particular, for the molecules with π bonds [98, 146].

Inspired by Pederson's FLO constraint, I proposed in this work to self-consistently restrict the unitary transformation of the SIC orbitals based on the Edmiston and Ruedenberg (E-R) localization method. Edmiston and Ruedenberg [31] first proposed the E-R method for localizing orbitals and maximizing the sum of self-electrostatic energies ($\max \sum_i E_{\text{es}}[n_i]$). It was documented that the orbitals localized by E-R method (named E-R orbitals) can maximize the integral of all squared orbital densities [140]

$$\max \int d^3\mathbf{r} \sum_i n_i^2(\mathbf{r}) \Leftrightarrow \min \int d^3\mathbf{r} \sum_{i\neq j} n_i(\mathbf{r}) n_j(\mathbf{r}) \quad , \tag{4.21}$$

which is equivalent to explicitly minimize the overlap of different E-R orbital densities. The E-R restriction in a matrix representation can be obtained by the variational maximization of the integral of all squared orbital densities subject to an orthonormalization constraint on the E-R orbitals with Lagrange multipliers η_{ab}:

$$\begin{aligned} &\delta \left[\sum_i \langle \phi_i | n_i | \phi_i \rangle - \sum_{ab} \eta_{ab} \left(\langle \phi_a | \phi_b \rangle - \delta_{ab} \right) \right] = 0 \\ &\Rightarrow \langle \phi_m | n_m | \phi_n \rangle = \langle \phi_m | n_n | \phi_n \rangle \end{aligned} \tag{4.22}$$

(Appx. 7.4 for detailed derivations). In the original PZ-SIC paper, Perdew and Zunger argued that E-R orbitals should be the most appropriate orbitals to define the SIC energy, which was confirmed in the latter work of Pederson and Lin [94]. The E-R localization method has been used to filter the multiple solutions of the standard PZ-SIC equation (Eq. (4.14)) under Constraints 1 and 2 (Eqs. (4.6) and (4.12)) [30]. Although taking the E-R method, like other localization methods, does not guarantee the resulting localized orbitals satisfy Constraint 2, the E-R restriction can be written in a similar form to Constraint 2 (Eqs. (4.12) and (4.22)). In this consideration, I generalized the aforementioned standard PZ-SIC equations (Eqs. (4.14), (4.9) and (4.12)) to include a new constraint.

Constraint 3. Orbital Density Constraint :

$$\langle\phi_m|n_m\hat{v}_n^{1e\mathrm{SIC}}|\phi_n\rangle = \langle\phi_m|\hat{v}_m^{1e\mathrm{SIC}}n_n|\phi_n\rangle \quad . \tag{4.23}$$

In order to transform the E-R restriction (Eq. (4.22)) as Constraint 3 into the self-consistent PZ-SIC framework, I learned from the derivation of Constraint 2 and proposed the self-consistently localized SIC (slSIC) Euler equation

$$\delta\left[\int d^3\mathbf{r}\sum_i n_i^2(\mathbf{r}) - \eta(\delta E^{\text{SIC-DFA}})\right] = 0 \quad . \tag{4.24}$$

In analogy to the derivation of Constraint 2, for a given set of KS orbitals fixed, this slSIC Euler equation can be rewritten under Constraint 2

$$\delta\left(\sum_i\int d^3\mathbf{r}n_i^2(\mathbf{r}) - \sum_{mn}\eta_{mn}\left\{\langle\phi_m|\hat{v}_i^{1e\mathrm{SIC}\dagger}|\phi_n\rangle - \sum_j\lambda_{jn}\langle\phi_m|\phi_j\rangle\right\}\right) = 0 \quad , \tag{4.25}$$

which leads to the one-electron equations

$$2n_i\phi_i = \frac{\delta(\hat{v}_i^{1e\mathrm{SIC}}n_i)}{\delta n_i}\sum_n\eta_{in}\phi_n - \sum_n\sum_a\eta_{ia}\lambda_{na}\phi_n \tag{4.26}$$

resulting in a set of localization equations as the third SIC constraint Eq. (4.23) (See Appx. 7.5 for more detailed derivations). Because this constraint is derived by the E-R method under Constraint 2, the SIC orbitals can be self-consistently localized and content to Constraint 2 simultaneously. To my best knowledge, no similar concept and algorithm has been reported in the previous works.

In the case of solids, the periodic boundary conditions (PBCs) require the KS orbitals $\{\psi_{l,\mathbf{k}}(\mathbf{r})\}$ to depend on $\mathbf{k}$-points (See Eq. (3.8)). Because KS orbitals have the invariant value in each cell $|\psi_{l,\mathbf{k}}(\mathbf{r}+\mathbf{R_L})| = |\psi_{l,\mathbf{k}}(\mathbf{r})|$, the repeated values let $\psi_{l,\mathbf{k}}$ to be considered as delocalized orbitals [42]. Due to the localized nature of the SIC orbitals, Heaton, Harrison, and Lin (HHL) suggested that it is more convenient to expand the SIC orbitals by using Wannier orbitals [27] in *real* space instead of expanding the SIC orbitals in reciprocal space. Wannier orbitals

$\{\Psi_{l,\mathbf{L}}(\mathbf{r})\}$ are the localized molecular orbitals of solids, which are generated via the Fourier transformation of the KS orbitals $\{\psi_{l,\mathbf{k}}(\mathbf{r})\}$

$$\begin{aligned}\Psi_{l,\mathbf{L}}(\mathbf{r}) &= \Psi_l(\mathbf{r}-\mathbf{R}_\mathbf{L}) \\ &= \frac{1}{\sqrt{N_\mathrm{L}}}\sum_{\mathbf{k}}^{N_\mathrm{L}} e^{-i\mathbf{k}\cdot\mathbf{R}_\mathbf{L}}\psi_{l,\mathbf{k}}(\mathbf{r}) \quad .\end{aligned} \tag{4.27}$$

Here N_L is the unit cell number (also the $\mathbf{k}$-point number) for the supercell. Note that the arbitrary phase $e^{i\theta_{l,\mathbf{k}}}$ to the $\mathbf{k}$-dependent KS orbitals can not impact the orbital densities

$$|e^{i\theta_{l,\mathbf{k}}}\psi_{l,\mathbf{k}}(\mathbf{r})|^2 = |\psi_{l,\mathbf{k}}(\mathbf{r})|^2 \quad , \tag{4.28}$$

but the Wannier orbital densities can change $|\Psi_{l,\mathbf{L}}(\mathbf{r})|^2 \neq |\Psi'_{l,\mathbf{L}}(\mathbf{r})|^2$ for

$$\Psi'_{l,\mathbf{L}}(\mathbf{r}) = \frac{1}{\sqrt{N_\mathrm{L}}}\sum_{\mathbf{k}}^{N_\mathrm{L}} e^{-i\mathbf{k}\cdot\mathbf{R}_\mathbf{L}}e^{i\theta_{l,\mathbf{k}}}\psi_{l,\mathbf{k}}(\mathbf{r}) \quad . \tag{4.29}$$

This is the reason that the multiple-solution problem for solids was reported to be more severe than that for molecules in the previous works [27, 28, 147–149].

In the early exploration of SIC in solids, it is observed that the SIC orbitals are analogous to the maximally localized Wannier orbitals (MLWOs) [150] (i.e., the SIC orbitals in Ar crystal [28]). So that one can ignore the orbital potential constraint and approximately use MLWOs as SIC orbitals to alleviate the multiple-solution issue. The MLWOs are determined by minimizing the Boy's localization functional [89]

$$\Omega = \sum_l [\langle\Psi_{l,\mathbf{0}}|\mathbf{r}^2|\Psi_{l,\mathbf{0}}\rangle - \langle\Psi_{l,\mathbf{0}}|\mathbf{r}|\Psi_{l,\mathbf{0}}\rangle^2] \quad , \tag{4.30}$$

which measures the sum of the quadratic spreads of Wannier orbitals in the unit cell. In many molecular systems, Boy's and E-R methods lead to similar localized orbitals, which agree with the viewpoint of bonding [150, 151]. The Boy's and E-R orbitals can be quite different in some molecule cases like CO_2 [152], in which the bonding can be represented as two distinct resonant structures. The E-R orbitals generally correspond to one of the possible structures, while the Boy's orbitals correspond to a mixing structure. However, Miyake and Aryasetiawan in 2008 [153] determined that the self-electrostatic energy $E_\mathrm{es}(|\Psi_{l,\mathbf{0}}|^2)$ in extensively periodic cases is maximized by MLWOs, which are essentially identical to the E-R orbitals.

It has been determined that SIC-LDA, using the MLWOs as the SIC orbitals, can give well agreement band-gaps with experimental measurements for bulk BeO, MgO, CaO, SrO, and BaO [154]. Moreover, it has been demonstrated that the PZ-SIC calculations with MLWOs improve the LDA performance for the lattice constants of cubic/hexagonal SiC polytypes [154]. To be specific, the standard LDA method underestimates the lattice constants up to 1.4%, which can be reduced to only 0.3% if using the SIC-LDA method with MLWO-type SIC orbitals. Although MLWOs conform to the localization nature of the

SIC orbitals, they only can be used as the input of solving Constraint 2. The resulting SIC orbitals are not necessary to be MLWOs.

In my project, I followed HHL's suggestions and represented the SIC orbitals as Wannier orbitals $\Phi_{i,\mathbf{I}}(\mathbf{r})$, which are indexed by i and mainly centered in the **I**th unit cell. In such a Wannier representation, Constraint 1 asks that the sum over all occupied Wannier SIC orbital densities $n_{i,\mathbf{I}} = |\phi_{i,\mathbf{I}}(\mathbf{r})|^2$ must recover the total electron density. This reads

$$\sum_{i\mathbf{I}}^{N_\mathrm{e}N_\mathrm{L}} |\phi_{i,\mathbf{I}}(\mathbf{r})|^2 = n(\mathbf{r}) = \sum_{l,\mathbf{k}}^{N_\mathrm{e}N_\mathrm{L}} |\psi_{l,\mathbf{k}}(\mathbf{r})|^2 \quad , \tag{4.31}$$

where N_e is the electron number in a unit cell, and N_L is the unit cell number for the supercell. Accordingly to the orthonormalization conditions of Wannier orbitals, the Wannier SIC orbitals can be a unitary transformation of the orbitals $\{\Psi_{l,\mathbf{L}}(\mathbf{r})\}$, like the form in Eq. (4.9) for finite systems [27]

$$\phi_{i,\mathbf{I}}(\mathbf{r}) = \sum_{l}^{N_\mathrm{e}} \sum_{\mathbf{L}}^{N_\mathrm{L}} T_{i\mathbf{I},l\mathbf{L}} \Psi_{l,\mathbf{L}}(\mathbf{r}) \quad . \tag{4.32}$$

The PZ-SIC total energy is evaluated as removing all the SIEs produced by the occupied Wannier SIC orbital densities in the DFA total energy:

$$\begin{aligned} E^{\text{SIC-DFA}}[n, \{n_{i,\mathbf{I}}\}] &= E^{\text{DFA}}[n] + E^{\text{SIC}}[\{n_{i,\mathbf{I}}\}] \quad , \\ E^{\text{SIC}}[\{n_{i,\mathbf{I}}\}] &= -\sum_{i\mathbf{I}}^{N_\mathrm{e}N_\mathrm{L}} \left(E_{\text{xc}}^{\text{DFA}}[n_{i,\mathbf{I}}] + E_{\text{es}}[n_{i,\mathbf{I}}] \right) \quad . \end{aligned} \tag{4.33}$$

Again $E_{\text{es}}[n_{i,\mathbf{I}}]$ ($E_{\text{xc}}^{\text{DFA}}[n_{i,\mathbf{I}}]$) is the self-electrostatic (self-XC) functional, but depends on the Wannier orbital density $n_{i,\mathbf{I}}$. The PZ-SIC one-electron equations have to be solved for different **k**-points in the first Brillouin zone, respectively

$$\hat{h}^{\text{SIC-DFA}}[n, \{\psi_{l,\mathbf{k}}\}, \{\phi_{i,\mathbf{I}}\}]\psi_{l,\mathbf{k}} = \epsilon_{l,\mathbf{k}}^{\text{SIC-DFA}} \psi_{l,\mathbf{k}} \quad . \tag{4.34}$$

The SIC potential operator $\hat{v}^{\text{SIC}}$ in the Hamiltonian of the SIC-DFA $\hat{h}^{\text{SIC-DFA}} = \hat{h}^{\text{DFA}} + \hat{v}^{\text{SIC}}$ depends on the cell sites

$$\hat{v}^{\text{SIC}} = \frac{\delta E^{\text{SIC}}[\{\phi_{i,\mathbf{I}}\}]}{\delta n} = \sum_{i\mathbf{I}}^{N_\mathrm{e}N_\mathrm{L}} \hat{v}_{i,\mathbf{I}}^{1e\text{SIC}} |\phi_{i,\mathbf{I}}\rangle\langle\phi_{i,\mathbf{I}}| \quad , \tag{4.35}$$

with $\hat{v}_{i,\mathbf{I}}^{1e\text{SIC}} = \delta E^{\text{SIC}}[\{n_{i,\mathbf{I}}\}]/\delta n_{i,\mathbf{I}}$ being the one-electron SIC potential under the PBCs. Correspondingly, Constraints 2 and 3 have the form of

$$\langle\phi_{m,\mathbf{M}}|\hat{v}_{m,\mathbf{M}}^{1e\text{SIC}}|\phi_{n,\mathbf{N}}\rangle = \langle\phi_{m,\mathbf{M}}|\hat{v}_{n,\mathbf{N}}^{1e\text{SIC}}|\phi_{n,\mathbf{N}}\rangle \tag{4.36}$$

and

$$\langle\phi_{m,\mathbf{M}}|n_{m,\mathbf{M}}\hat{v}_{n,\mathbf{N}}^{1e\text{SIC}}|\phi_{n,\mathbf{N}}\rangle = \langle\phi_{m,\mathbf{M}}|\hat{v}_{m,\mathbf{M}}^{1e\text{SIC}} n_{n,\mathbf{N}}|\phi_{n,\mathbf{N}}\rangle \quad . \tag{4.37}$$

Different from the constraints in finite systems (Eqs. (4.12) and (4.23)), Constraint 2 and 3 in Wannier representation consider the cell sites, which make the dimension of the SIC constraint matrix in Eqs. (4.36) and (4.37) to increase fast with the cell counts N_{L}. Moreover, the serious multiple-solution problem in periodic systems can be traced to the introduced cell site in Constraint 2, like the hcp Helium crystal and the polyethylene chain (discussed in Sec. 4.3). Before discussing the SIC multiple-solution problem, a SIC solver is proposed to solve the PZ-SIC constraints of finite and periodic systems in the next section.

4.2 Generalized SIC Solver

In order to get the lowest total energy of SIC-DFAs (Eq. (4.2)), both KS orbitals and SIC orbitals need to be determined, which are entangled with each other by the SIC one-electron equations (Eq. (4.14)) and SIC constraints (Eqs. (4.6), (4.12) and (4.23)).

To be specific, a Lagrange multiplier solver was established to solve the SIC one-electron equations (Eqs. (4.14) and (4.34)). For a given set of density n, KS orbitals $\{\psi_l\}$ and SIC orbitals $\{\phi_m\}$, the updated KS orbitals ψ'_l and the density n' can be obtained by

$$\begin{aligned} &\hat{h}^{\text{SIC-DFA}}[n, \{\psi_l\}, \{\phi_m\}]\ \psi'_l = \epsilon_l^{\text{SIC-DFA}'} \psi'_l \\ &n' = \sum_l |\psi'_l|^2 \quad . \end{aligned} \tag{4.38}$$

The updated KS orbitals $\{\psi'_l\}$ are then taken as the input to the SIC orbital solver for the updated SIC orbitals $\{\phi'_m\}$. Under the PBCs, the updated **k**-dependent KS orbitals $\psi'_{l,\mathbf{k}}$ and the density n' can be obtained by

$$\begin{aligned} &\hat{h}^{\text{SIC-DFA}}[n, \{\psi_l\}, \{\phi_{m,\mathbf{M}}\}]\ \psi'_{l,\mathbf{k}} = \epsilon_{l,\mathbf{k}}^{\text{SIC-DFA}'} \psi'_{l,\mathbf{k}} \\ &n' = \sum_{l,\mathbf{k}}^{N_{\mathrm{e}} N_{\mathrm{L}}} |\psi'_{l,\mathbf{k}}(\mathbf{r})|^2 \end{aligned} \tag{4.39}$$

with a given set of density n, KS orbitals $\{\psi_{l,\mathbf{k}}\}$ and Wannier SIC orbitals $\{\phi_{m,\mathrm{M}}\}$. Unlike the standard PZ-SIC implementation in previous works, the SIC orbital solver in my implementation results in a set of SIC orbitals that satisfy not only Constraints 1 and 2, but also Constraint 3 guaranteeing the resulting updated SIC orbitals to be localized. These updated SIC orbitals are then used as the input to the aforementioned Lagrange multiplier solver to start a new round calculation until the total density and SIC orbital densities are converged. The flowchart of our self-consistency algorithm is shown in Fig. 4.1.

The SIC orbital solver used in previous works [96, 101] solved Constraints 1 and 2 utilizing the gradient of the SIC potential, like $\partial \hat{v}_m^{1e\text{SIC}} / \partial \phi_m$. Because the Constraint 3 (Eqs. 4.23 and (4.37)) is supplemented into the standard PZ-SIC method, those solvers need the gradient of the product of the SIC orbital

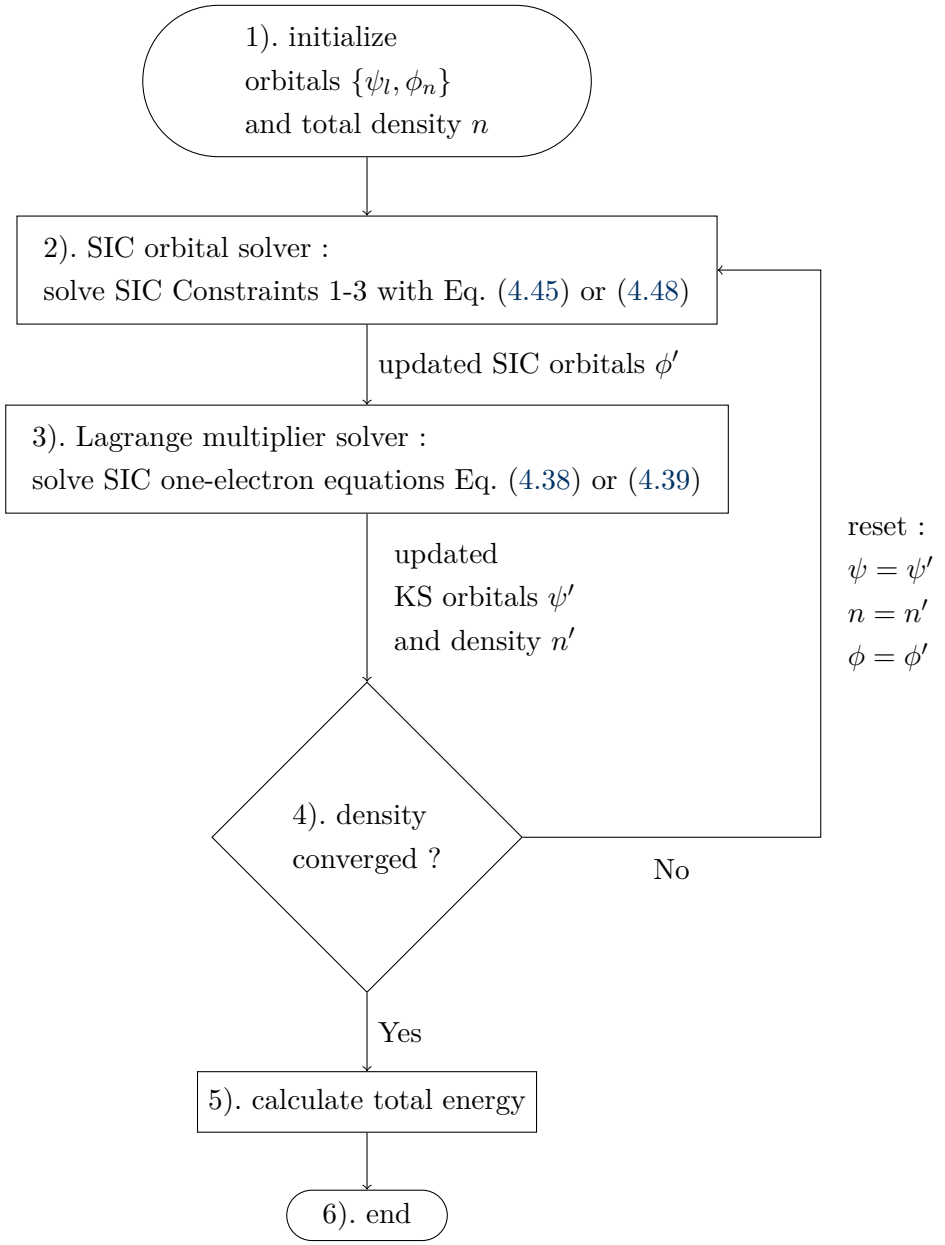

Figure 4.1: Flowchart of the self-consistent SIC calculations in *FHI-aims*.

density and potential (like $\partial(n_n \hat{v}_m^{1e\mathrm{SIC}})/\partial\phi_m$), which isn't easy to obtain and thus cannot be implemented together with those SIC solvers in my project. In this consideration, I developed a generalized SIC orbital solver. Any equation, formed as Constraints 2 or 3, can be turned into an identical eigenvalue problem and thus can be treated on an equal footing.

Finite system. Constraints 2 and 3 can be essentially written as a generalized constraint, in which an orbital-dependent operator $\hat{o}[\{\phi_i\}]$ is solved to be

hermitian conjugate

$$\hat{o}^{\dagger}[\{\phi_i\}] - \hat{o}[\{\phi_i\}] \to 0 \quad . \tag{4.40}$$

Here $\hat{o}$ in Constraint 2 is the SIC potential operator

$$\hat{o}^{v} = \hat{v}^{\mathrm{SIC}} = \sum_{i}^{N_{\mathrm{e}}} (\hat{v}_i^{1e\mathrm{SIC}} |\phi_i\rangle\langle\phi_i|) \quad , \tag{4.41}$$

and it in Constraint 3 is the multiplication of the SIC potential operator and the orbital densities

$$\hat{o}^{d} = (\hat{v}^{\mathrm{SIC}})^{\dagger} \sum_{i}^{N_{\mathrm{e}}} (n_i |\phi_i\rangle\langle\phi_i|) \quad . \tag{4.42}$$

The generalized constraint can be further written as an eigenvalue problem like

$$(\hat{o}^{\dagger}[\{\phi_i\}] - \hat{o}[\{\phi_i\}])\phi'_n = 0 \cdot \phi'_n \quad , \tag{4.43}$$

in which $(\hat{o}^{\dagger} - \hat{o})$ owns a set of eigenvectors ϕ'_n with zeroed eigenvalues. The equally zeroed eigenvalues cause a singular problem because any unitary transform C_{js} of ϕ'_i satisfies Eq. (4.43)

$$\begin{aligned} &(\hat{o}^{\dagger}[\{\phi_i\}] - \hat{o}[\{\phi_i\}])\phi''_s = 0 \cdot \phi''_s \\ &\phi''_s = \sum_i C_{is}\phi'_i \\ &\sum_s C^*_{is} C_{js} = \delta_{ij} \quad . \end{aligned} \tag{4.44}$$

To avoid this problem, I decorated the aforementioned equations (Eq. (4.43)) with a set of ordering weights with w_n :

$$(\hat{o}^{\dagger}[\{\phi_i\}] - \hat{o}[\{\phi_i\}] + \sum_i (w_i |\phi_i\rangle\langle\phi_i|))\phi'_n = w'_n \cdot \phi'_n \quad . \tag{4.45}$$

So that the corresponding eigenvalues in Eq. (4.45) become non-zero $\{w'_n\}$. These equations do not have the zeroed singular problem if $|w_{i+1}| > |w_i| \gg |\langle\phi_i|(\hat{o}^{\dagger}[\{\phi_i\}] - \hat{o}[\{\phi_i\}]|\phi_i\rangle| \geq 0$. Moreover, a hermitian conjugate $\hat{o}$ makes Eq. (4.45) to degenerate into Eq. (4.43) with $\{w'_n\} = \{w_i\}$ and $\{\phi'_n\} = \{\phi_i\}$. In this work, we numerically set $w_i = i/(|\langle\phi_i|(\hat{o}^{\dagger}[\{\phi_i\}] - \hat{o}[\{\phi_i\}]|\phi_i\rangle| + \eta_e)^{0.5}$. Here $\eta_e = 10^{-13}$ is a compensating value in case of $\langle\phi_i|(\hat{o}^{\dagger}[\{\phi_i\}] - \hat{o}[\{\phi_i\}]|\phi_i\rangle = 0$ for the double-precision floating-point.

Note that a matrix representation of the eigenvalue problem in Eq. (4.45) is used in practical calculation

$$\sum_j O_{ij} C_{jn} = C_{in} w'_n \quad , \tag{4.46}$$

where $O_{ij} = \langle\phi_i|\hat{o}^{\dagger} - \hat{o}|\phi_j\rangle + \delta_{ij} w_i$ and C_{in} is the unitary transform matrix for the SIC orbitals $\phi'_n = \sum_i C_{in}\phi_i$. In this way, the updated SIC orbitals are a

unitary transformation of the occupied KS orbitals ψ_l

$$\begin{aligned} \phi'_n &= \sum_{l}^{N_\mathrm{e}} T'_{nl}\psi_l \\ T'_{nl} &= \sum_i C_{in}T_{il} \quad . \end{aligned} \tag{4.47}$$

Consequently, the updated SIC orbitals keep satisfying Constraint 1 during solving Constraints 2 and 3. A self-consistent calculation is used to solve the equation (Eq. (4.45)) until the error of Constraints 2 and 3 is less than the requested energy accuracy, i.e., 10^{-4} eV in this work. The resulting SIC orbitals $\{\phi'_n\}$ satisfy all three constraints and deliver lower SIC-DFA total energy for a given set of KS orbitals. Fig. 4.2 shows the flowchart of the procedure in the SIC orbital solver, which is a sub-self-consistent calculation in the second step of the whole SIC self-consistent procedure in Fig. 4.1.

Periodic system. Concerning the SIC solver in the periodic systems, Constraints 2 and 3 (Eqs. (4.36) and (4.36)) can be described using a generalized eigenvalue problem with the operator $\hat{o}[\{\phi_{i,\mathbf{I}}\}]$ and weights $w_{i,\mathbf{I}}$

$$(\hat{o}^\dagger[\{\phi_{i,\mathbf{I}}\}] - \hat{o}[\{\phi_{i,\mathbf{I}}\}] + \sum_i (w_{i,\mathbf{I}}|\phi_{i,\mathbf{I}}\rangle\langle\phi_{i,\mathbf{I}}|))\phi'_{n,\mathbf{N}} = w'_{n,\mathbf{N}} \cdot \phi'_{n,\mathbf{N}} \quad , \tag{4.48}$$

which follows the same derivations in Eq. (4.45) for the finite systems but is in Wannier representation. Consequently, the weight order is set as $w_{i,\mathbf{I}} = i/(|\langle\phi_{i,\mathbf{I}}|(\hat{o}^\dagger[\{\phi_{i,\mathbf{I}}\}] - \hat{o}[\{\phi_{i,\mathbf{I}}\}]|\phi_{i,\mathbf{I}}\rangle| + \eta_e)^{0.5}$ with $\eta_e = 10^{-13}$ for the double-precision floating-point in practical calculations. The generalized operator $\hat{o}$ in Constraint 2 is the SIC potential operator using Wannier SIC orbitals

$$\hat{o}^v = \hat{v}^\mathrm{SIC} = \sum_{i\mathbf{I}}^{N_\mathrm{e}N_\mathrm{L}} \hat{v}^{1e\mathrm{SIC}}_{i,\mathbf{I}}|\phi_{i,\mathbf{I}}\rangle\langle\phi_{i,\mathbf{I}}| \quad , \tag{4.49}$$

and it in Constraint 3 reads

$$\hat{o}^d = (\hat{v}^\mathrm{SIC})^\dagger \sum_{i\mathbf{I}}^{N_\mathrm{e}N_\mathrm{L}} (n_{i,\mathbf{I}}|\phi_{i,\mathbf{I}}\rangle\langle\phi_{i,\mathbf{I}}|) \quad . \tag{4.50}$$

The matrix representation of the eigenvalue problem in Eq. (4.48) follows the one for finite systems (Eq. (4.46)) and can be written as

$$\sum_{jJ} O_{ij,\mathbf{IJ}} C_{jn,\mathbf{JN}} = C_{in,\mathbf{IN}} w'_{n,\mathbf{N}} \quad , \tag{4.51}$$

where $O_{ij,\mathbf{IJ}} = \langle\phi_{i,\mathbf{I}}|\hat{o}^\dagger - \hat{o}|\phi_{j,\mathbf{J}}\rangle + \delta_{ij}\delta_{\mathbf{IJ}}w_{i,\mathbf{I}}$ and $C_{in,\mathbf{IN}}$ is the unitary transform matrix for the Wannier SIC orbitals $\phi'_{n,\mathbf{N}} = \sum_{i,\mathbf{I}} C_{in,\mathbf{IN}}\phi_{i,\mathbf{I}}$.

As shown in Eq. (4.51), the Wannier SIC orbitals have two index i an $\mathbf{I}$, indicating that the number of SIC orbitals increase linearly with respect to the unit cell number N_L. In consequence, the dimension of the matrices $O_{ij,\mathbf{IJ}}$ and $C_{jn,\mathbf{JN}}$

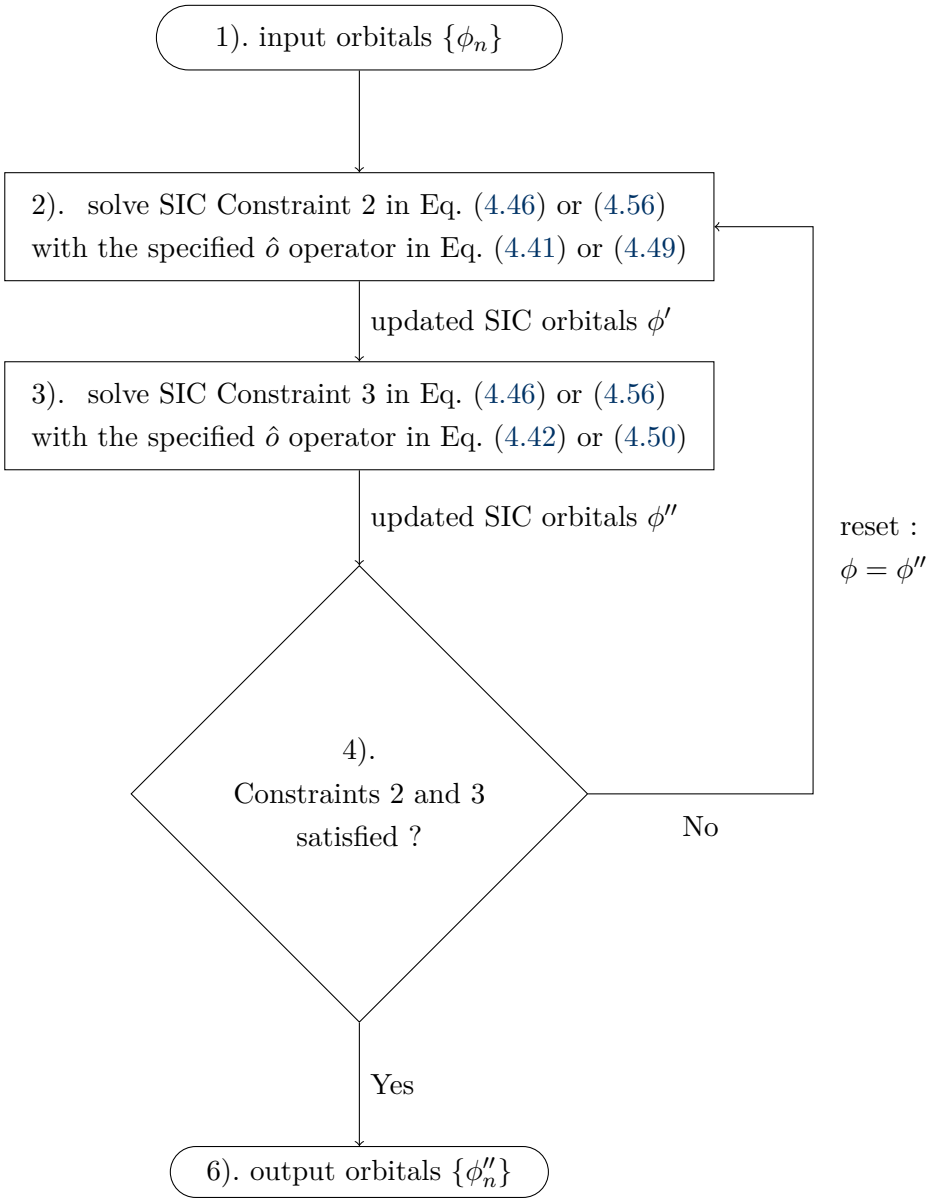

Figure 4.2: Flowchart of the self-consistent SIC orbital solver in *FHI-aims*. This is a sub-self-consistent calculation in the second step of the whole SIC self-consistent procedure shown in Fig. 4.1.

in Eq. (4.51) increases quadratically with respect to the unit cell number N_{L} (scaling with $N_{\mathrm{L}} \times N_{\mathrm{L}}$). In other words, the use of Wannier functions dramatically increases the complexity. Although each Wannier SIC orbital is localized in a single cell and is not a periodic function $|\phi_{i,\mathbf{I}+\mathbf{L}}| \neq |\phi_{i,\mathbf{I}}|$, the effective SIC potential, summing all SIC orbital potential overall cells, has the translation invariance, unlike the case for finite molecules

$$\sum_{\mathbf{M}} \hat{v}^{1e\mathrm{SIC}}_{m,\mathbf{M}+\mathbf{L}} = \sum_{\mathbf{M}} \hat{v}^{1e\mathrm{SIC}}_{m,\mathbf{M}} \quad . \tag{4.52}$$

This translation invariance under the PBCs reduces the size of Constraint 2 equations

$$\begin{aligned}
&\langle\phi_{m,\mathbf{M}}|\hat{v}^{1e\mathrm{SIC}}_{m,\mathbf{M}} - \hat{v}^{1e\mathrm{SIC}}_{n,\mathbf{N}}|\phi_{n,\mathbf{N}}\rangle \\
&= \int d^3\mathbf{r}\phi_{m,\mathbf{M}}(\mathbf{r})(\hat{v}^{1e\mathrm{SIC}}[|\phi_{m,\mathbf{M}}(\mathbf{r})|^2] - \hat{v}^{1e\mathrm{SIC}}[|\phi_{n,\mathbf{N}}(\mathbf{r})|^2])\phi_{n,\mathbf{N}}(\mathbf{r}) \\
&= \int d^3(\mathbf{r}+\mathbf{R}_\mathbf{N})\phi_{m,\mathbf{M+N}}(\mathbf{r})(\hat{v}^{1e\mathrm{SIC}}[|\phi_{m,\mathbf{M+N}}(\mathbf{r})|^2] - \hat{v}^{1e\mathrm{SIC}}[|\phi_{n,\mathbf{0}}(\mathbf{r})|^2])\phi_{n,\mathbf{0}}(\mathbf{r}) \\
&= \int d^3\mathbf{r}\phi_{m,\mathbf{M+N}}(\mathbf{r})(\hat{v}^{1e\mathrm{SIC}}[|\phi_{m,\mathbf{M+N}}(\mathbf{r})|^2] - \hat{v}^{1e\mathrm{SIC}}[|\phi_{n,0}(\mathbf{r})|^2])\phi_{n,0}(\mathbf{r}) \\
&= \langle\phi_{m,\mathbf{M+N}}|\hat{v}^{1e\mathrm{SIC}}_{m,\mathbf{M+N}} - \hat{v}^{1e\mathrm{SIC}}_{n,\mathbf{0}}|\phi_{n,\mathbf{0}}\rangle \quad .
\end{aligned} \tag{4.53}$$

Because the total density is translation invariant ($n(\mathbf{r}+\mathbf{R}_\mathbf{L}) = n(\mathbf{r})$ introduced in Eq. (3.9)), summing all the Wannier SIC orbital densities also has the translation invariance under Constraint 1 (Eq. (4.31)) and PBCs:

$$n(\mathbf{r}+\mathbf{R}_\mathbf{L}) = \sum_\mathbf{M} n_{m,\mathbf{M+L}} = \sum_\mathbf{M} n_{m,\mathbf{M}} = n(\mathbf{r}) \quad , \tag{4.54}$$

which can reduce the size of Constraint 3 equations to

$$\begin{aligned}
&\langle\phi_{m,\mathbf{M}}|\hat{v}^{1e\mathrm{SIC}}_{n,\mathbf{N}} n_{m,\mathbf{M}} - \hat{v}^{1e\mathrm{SIC}}_{m,\mathbf{M}} n_{n,\mathbf{N}}|\phi_{n,\mathbf{N}}\rangle \\
&= \langle\phi_{m,\mathbf{M+N}}|n_{m,\mathbf{M+N}}\hat{v}^{1e\mathrm{SIC}}_{n,\mathbf{0}} - \hat{v}^{1e\mathrm{SIC}}_{m,\mathbf{M+N}} n_{n,\mathbf{0}}|\phi_{n,\mathbf{0}}\rangle \quad .
\end{aligned} \tag{4.55}$$

As the vectors $\mathbf{R}_\mathbf{M} + \mathbf{R}_\mathbf{N}$ point to the same N_L sets of the lattice vectors $\{\mathbf{R}_P\}$ under the PBCs, the matrix representation of the eigenvalue problem Eq. (4.48) can be simplified as

$$\sum_s O_{ms,\mathbf{P0}} C_{ns,\mathbf{P0}} = C_{mn,\mathbf{P0}} w'_{n,\mathbf{P0}} \quad . \tag{4.56}$$

Such that the complexity of the problem is reduced from $N_\mathrm{L} \times N_\mathrm{L}$ (in Eq. (4.51)) to N_L (in Eq. (4.56)). The updated Wannier SIC orbitals are then obtained by

$$\phi'_{n,\mathbf{N}} = \frac{1}{N_\mathrm{L}} \sum_{s\mathbf{PS}} C_{ns,\mathbf{P0}} e^{i\mathbf{P}\cdot(\mathbf{R_N}-\mathbf{R_S})} \phi_{s,\mathbf{S}} \quad , \tag{4.57}$$

where the translation multiplier $e^{i\mathbf{P}\cdot(\mathbf{R_N}-\mathbf{R_S})}$ transmits the orbital $\phi'_{n,\mathbf{0}}$ in the unit cell $\mathbf{0}$ to the orbital $\phi'_{n,\mathbf{N}}$ in any given unit cell $\mathbf{N}$. The unitary transform matrix $T_{nl,\mathbf{NL}}$ from KS to SIC orbitals in the Wannier representation (shown as Eq. (4.32)) will be then updated by the unitary transformation matrix $C_{ns,\mathbf{P0}}$

$$T'_{sl,\mathbf{SL}} = \frac{1}{N_\mathrm{L}} \sum_{n\mathbf{PN}} C_{ns,\mathbf{P0}} e^{i\mathbf{P}\cdot(\mathbf{R_N}-\mathbf{R_S})} T_{nl,\mathbf{NL}} \quad . \tag{4.58}$$

The flowchart for periodic systems is shown in Fig. 4.2, which is the same for the finite systems.

Here, I took the rocksalt LiH to investigate the computational cost of my PZ-SIC implementation under the periodic boundary conditions. As shown in Fig. 4.3,

my SIC-SCAN implementation, despite being more expensive than the standard SCAN method, scales linearly with respect to the **k**-point size (or the cell size), and thus avoids the unfavorable quadratic scaling of $N_\mathrm{L} \times N_\mathrm{L}$. Having this efficient solver allows me to detect the importance of the SIC Constraint 3 and investigate the performance of the SIC method in detail.

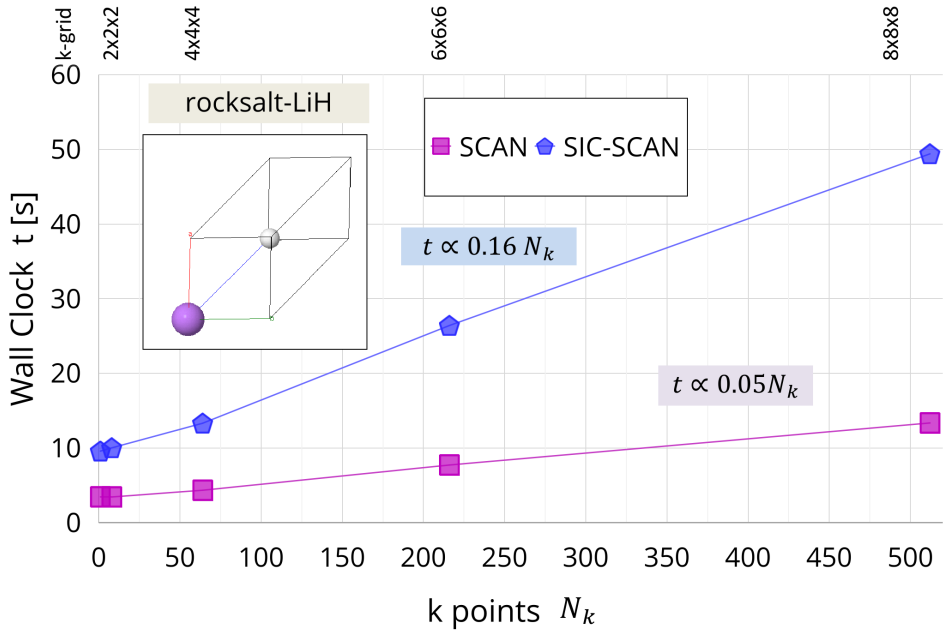

Figure 4.3: Time cost per SCF iteration for the SCAN and SIC-SCAN calculations with respect to the size of **k**-points (N_k) for the rocksalt LiH. The number of **k**-points N_k equals to that of cells N_L. The corresponding **k**-grids are listed on the top of the x-axis. Geometries are shown in the sub-graph. All the calculations were performed in *FHI-aims* using the *tier*-1 basis sets and the *tight* grid setting. The time costs are linearly scaling as $0.05N_\mathrm{k}$ and $0.16N_\mathrm{k}$ for SCAN and SIC-SCAN calculations, respectively.

4.3 Importance of Constraint 3

To illustrate the importance of Constraint 3 for a fast and stable PZ-SIC self-consistent calculation, I used three prototype finite systems (i.e., the He_2 molecule, and two atoms of B and C) as well as three periodic systems (i.e., the hcp Helium crystal, the CH_2 chain, and the rocksalt LiH).

I started with the closed-shell dihelium He_2 molecule with a bond length of 1.1 Å, where $1\sigma_g$ and $1\sigma_u$ are the occupied KS orbitals with two electrons in each of them. As shown in Fig. 4.4(a), these occupied KS orbitals delocalize over the

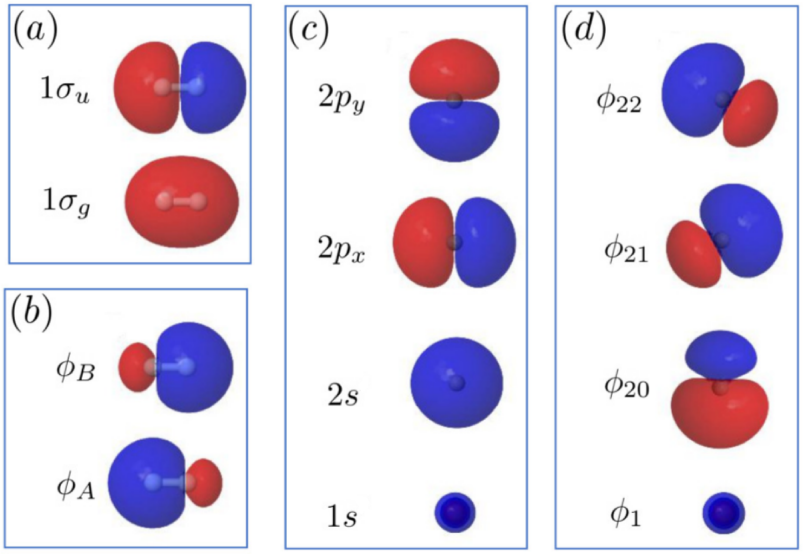

Figure 4.4: Illustration on the KS orbital densities (a) and the SIC ones (b) of the He_2 molecule, and the KS orbital densities (c) and the SIC ones (d) of the C atom. The KS orbitals and SIC orbitals are evaluated by the (SIC-)SCAN calculations in *FHI-aims* with *tight* basis sets.

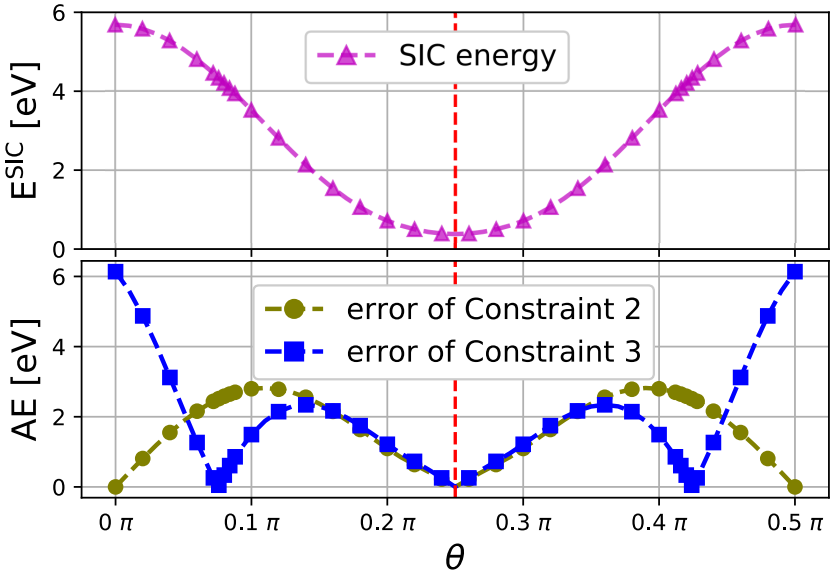

Figure 4.5: (top) SIC energy using SCAN XC for dihelium He_2 and (bottom) absolute error of Constraint 2 (Eq. (4.12)) and Constraint 3 (Eq. (4.23)) in a linear combination of the $1\sigma_g$ and $1\sigma_u$ orbitals (Eq. (4.59)). SIC and SCAN are calculated by *FHI-aims* using *tight* basis sets.

system with mirror symmetry. Consequently, all possible SIC orbitals $\{\phi_1^\theta, \phi_2^\theta\}$

can be obtained by a unitary transformation of the KS orbitals

$$\begin{pmatrix} \phi_1^\theta \\ \phi_2^\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1\sigma_g \\ 1\sigma_u \end{pmatrix} \quad . \tag{4.59}$$

Here $\theta \in [0, \pi/2]$ controls the rotation angle of the transformation, and $\theta = 0$ maps $\{\phi_{1;2}^\theta\} = \{\phi_{g;u}\}$. Since the rotation angle θ is the only variable for the unitary transformation, I plotted the SIC energy with respect to this variable on the top panel of Fig. 4.5. As we can see clearly from this figure, the SIC orbitals that minimize the SIC total energy are associated with the unitary transformation of $\theta = 0.25\pi$. With this certain unitary transformation ($\theta = 0.25\pi$), orbitals $\{\phi_1^\theta, \phi_2^\theta\}$ become to

$$\phi_{A;B} = \frac{1}{\sqrt{2}}[1\sigma_g \pm 1\sigma_u] \quad . \tag{4.60}$$

Fig. 4.4(b) shows that the orbitals $\phi_{A;B}$ are localized at each atom (A and B). It is easy to prove that all possible orbitals $\{\phi_1^\theta, \phi_2^\theta\}$ (including $\phi_{A;B}$) satisfy Constraint 1 (Eq. (4.6))

$$\begin{aligned} &2|\phi_1^\theta(\mathbf{r})|^2 + 2|\phi_2^\theta(\mathbf{r})|^2 \\ &= 2(\cos^2\theta + \sin^2\theta)(|1\sigma_g(\mathbf{r})|^2 + |1\sigma_u(\mathbf{r})|^2) \\ &= 2|1\sigma_g(\mathbf{r})|^2 + 2|1\sigma_u(\mathbf{r})|^2 \\ &= n(\mathbf{r}) \quad . \end{aligned} \tag{4.61}$$

I then plotted the absolute errors of Constraint 2 and 3 with respect to this rotation angle θ on the bottom panel of Fig. 4.5. It clearly shows that Constraint 2 itself cannot ultimately filter out the right SIC orbitals $\phi_{A;B}$ that produce the lowest SIC energy as mentioned above. To be specific, both sets of orbitals $\{\phi_{A;B}\}$ and $\{1\sigma_{g;u}\}$ satisfy Constraint 2 (Eq. (4.12))

$$\begin{aligned} \langle 1\sigma_g|\hat{v}_g^{1e\mathrm{SIC}}|1\sigma_u\rangle &= \langle 1\sigma_g|\hat{v}_u^{1e\mathrm{SIC}}|1\sigma_u\rangle \quad , \\ \langle \phi_A|\hat{v}_A^{1e\mathrm{SIC}}|\phi_B\rangle &= \langle \phi_A|\hat{v}_B^{1e\mathrm{SIC}}|\phi_B\rangle \quad , \end{aligned} \tag{4.62}$$

resulting in two solutions for the SIC orbitals. In this case, I would like to point out that Constraint 3 can be a useful and often necessary complementary to the PZ-SIC theory:

$$\begin{aligned} \langle 1\sigma_g|n_g\hat{v}_u^{1e\mathrm{SIC}}|1\sigma_u\rangle &\neq \langle 1\sigma_g|\hat{v}_g^{1e\mathrm{SIC}}n_u|1\sigma_u\rangle \quad , \\ \langle \phi_A|n_A\hat{v}_B^{1e\mathrm{SIC}}|\phi_B\rangle &= \langle \phi_A|\hat{v}_A^{1e\mathrm{SIC}}n_B|\phi_B\rangle \quad . \end{aligned} \tag{4.63}$$

It picks $\{\phi_{A;B}\}$ as a better choice for the SIC orbitals, which indeed yields the SIC-SCAN total energy about 5.29 eV lower than that of $\{1\sigma_{g;u}\}$. Besides the solution of $\{\phi_{A;B}\}$, there is another solution nearby $\theta \approx 0.4\pi$ for Constraint 3. This solution is understandable. We know that Constraint 3 (Eq. (4.25))

was imposed to minimize the overlap orbital densities under the variational of Constraint 2:

$$\delta\left(\sum_i \int d^3\mathbf{r} n_i^2(\mathbf{r})\right) - \delta\left(\text{Constraint 2}\right) = 0 \quad . \tag{4.64}$$

Constraint 3 intends to produce the SIC orbitals where Constraint 2 has the local maximal (or minimum) errors. To take the two sub-panels in Fig. 4.5 into account together, I concluded that Constraints 2 and 3 must work together in order to find the correct SIC orbitals.

The second model system I would like to introduce is the Boron atom. Same as in the first model system of He_2, the list of all possible SIC orbitals can be exhaustive associated with the electronic configuration of $1s2s(2p)^1$ in its ground state. The only one SIC orbital for the spin-down channel is the atomic orbital $2s$. In order to satisfy Constraint 1, the spin-up SIC orbitals in the second shell correspond to a unitary transformation from the occupied KS orbitals $2s; 2p$

$$\begin{pmatrix} \phi_1^\theta \\ \phi_2^\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 2s \\ 2p \end{pmatrix} \quad , \tag{4.65}$$

where $\theta \in [0, \pi/2]$ controls the rotation angle of the transformation. Also, same as in the model system of He_2, the SIC energy for SCAN (E^{SIC}) and the absolute errors of Constraint 2 and 3 were plotted with respect to the angle θ in the top panel and bottom panel of Figure 4.6, respectively. From this figure, we can obtain the same conclusion made in the first case: Constraints 2 and 3 must both be fulfilled to find the right SIC orbitals and the correct SIC energy. To be specific, Constraint 2 itself yields two solutions with the angles of $\theta = 0, \pi/4$. Only one solution ($\theta = \pi/4$) remains if I simultaneously consider both Constraints 2 and 3. Moreover, Figure 4.6-top confirms that the common solution of Constraints 2 and 3 ($\theta = \pi/4$) yields the lowest SIC energy. This energy is about 0.36 eV lower than the SIC energy of the SIC orbitals at $\theta = 0$, which only satisfies Constraint 2.

Similar observations are found in the SIC-SCAN calculations of the Carbon atom. In this case, the SIC orbitals were chiefly generated by mixing s and p orbitals associated with the electronic configuration of $1s2s(2p)^2$ in its ground state. The atomic orbitals $\{1s, 2s, 2p_x, 2p_y\}$ [94] shown in Fig. 4.4(c) satisfy Constraints 1 and 2, but not 3. I notice that the three constraints can be simultaneously satisfied by a set of occupied orbitals after a certain *real* unitary transformation from the atomic orbitals:

$$\begin{aligned} \phi_1 &= 1s \\ \phi_{20} &= \frac{1}{\sqrt{3}}2s - \frac{\sqrt{2}}{\sqrt{3}}2p_x \\ \phi_{21} &= \frac{1}{\sqrt{3}}2s + \frac{1}{\sqrt{6}}2p_x + \frac{1}{\sqrt{2}}2p_y \\ \phi_{22} &= \frac{1}{\sqrt{3}}2s + \frac{1}{\sqrt{6}}2p_x - \frac{1}{\sqrt{2}}2p_y \quad , \end{aligned} \tag{4.66}$$

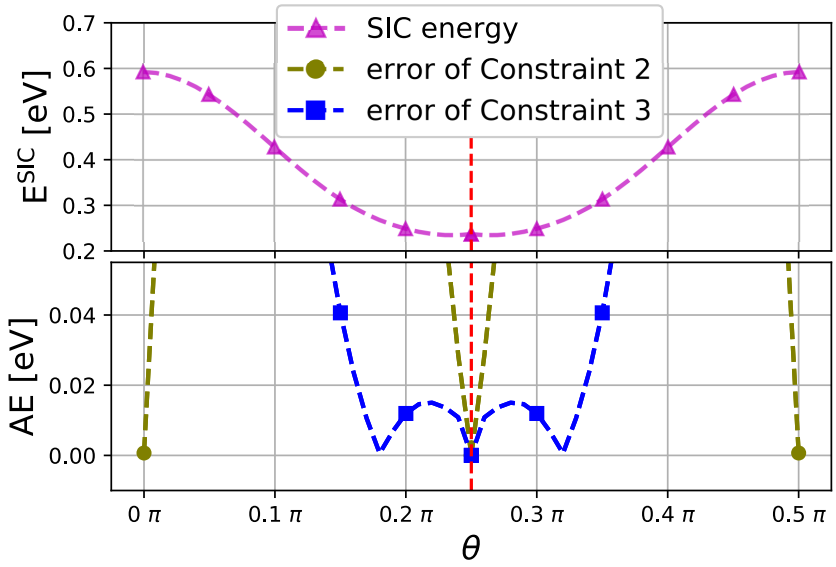

Figure 4.6: (top) SIC energy using SCAN XC for B atom and (bottom) absolute error of Constraint 2 (Eq. (4.12)) and Constraint 3 (Eq. (4.23)) in a linear combination of the $2s$ and $2p$ orbitals (Eq. (4.65)). SIC and SCAN are calculated by *FHI-aims* using *tight* basis sets.

which were visualized in Fig. 4.4(d). These SIC orbitals can be considered as a kind of *sp* hybridization with themselves. My calculation demonstrated that these SIC orbitals produce the SIC-SCAN total energy, which is about 0.86 eV lower than the atomic orbitals mentioned above. Again, the SIC-SCAN calculations for the Carbon atom lead to the same conclusion I obtained for He_2 and B, highlighting the importance and robustness of introducing Constraint 3 self-consistently for the PZ-SIC theory.

It was argued that the SIC multiple-solution problem is more serious for solids [27, 28, 147–149]. Here, I use a simple He crystal to study this problem by increasing the cell size N_{L} under periodic boundary conditions. As there are two He atoms in each unit cell (See Fig. 4.7), the occupied KS orbitals in the Wannier representation can be marked as $1\Sigma_{g,\mathbf{I}}$ and $1\Sigma_{u,\mathbf{I}}$, which have the same shapes as the occupied molecular orbitals in the closed-shell dihelium He_2 molecule (See Fig. 4.4-a)).

As in the case of the He_2 molecule, all possible SIC orbitals $\{(\Phi_{1,\mathbf{I}}, \Phi_{2,\mathbf{I}})\}$ that satisfy Constraint 1 can be generalized via unitary transformations from the Wannier KS orbitals $(1\Sigma_{g,\mathbf{I}}, 1\Sigma_{u,\mathbf{I}})$ using two rotation angles $\theta_I \in [0, 0.5\pi], I = 1, 2$.

$$\begin{aligned} \Phi_{j,\mathbf{J}} &= \sum_{\mathbf{I}=1,2} \left(a_{g,\mathbf{I}}^{j,\mathbf{J}}(\theta_1,\theta_2) 1\Sigma_{g,\mathbf{J}} + a_{u,\mathbf{I}}^{j,\mathbf{J}}(\theta_1,\theta_2) 1\Sigma_{u,\mathbf{J}} \right) \quad , \\ j &= 1,2; \quad \mathbf{J} = 1,2 \quad . \end{aligned} \tag{4.67}$$

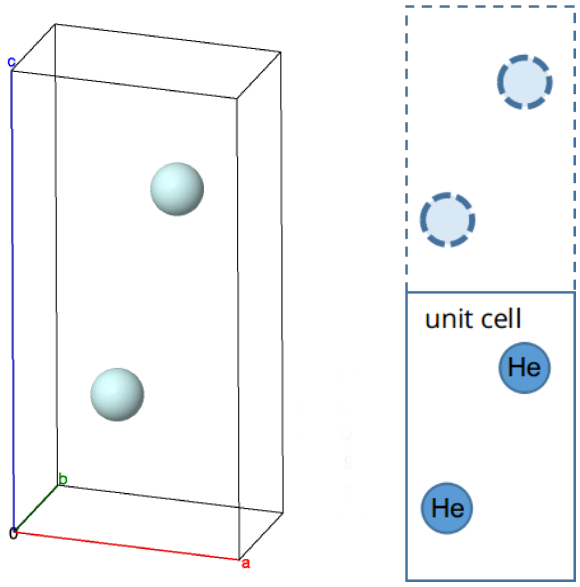

Figure 4.7: (left) The unit cell of the hcp Helium crystal with a geometry taken from Ref. [155]. (right) The schematic view of two unit cells with periodic boundary conditions. There are two Helium atoms located in the unit cell.

Here $a_{g,\mathbf{I}}^{j,\mathbf{J}}(\theta_1,\theta_2)$ is a set of rotation functions of angles θ_1,θ_2,

$$\begin{aligned} a_{g,\mathbf{I}}^{j,\mathbf{J}}(\theta_1,\theta_2) &= \frac{1}{2}(\cos\theta_1 - (-1)^j \sin\theta_1) + \frac{(-1)^{\mathbf{I}+\mathbf{J}}}{2\sqrt{2}}(\cos\theta_2 - (-1)^j \sin\theta_2) \quad , \\ a_{u,\mathbf{I}}^{j,\mathbf{J}}(\theta_1,\theta_2) &= \frac{(-1)^{\mathbf{I}+\mathbf{J}}}{2\sqrt{2}}(\cos\theta_2 + (-1)^j \sin\theta_2) \quad , \\ j = 1,2; \quad \mathbf{I} &= 1,2; \quad \mathbf{J} = 1,2 \quad . \end{aligned} \tag{4.68}$$

Fig. 4.8 shows the SIC energy for the SCAN functional using these orbitals. It suggests that the minimum value is 0.32 eV produced by the SIC orbitals with $(\theta_1 = 0.25\pi, \theta_2 = 0.50\pi)$. For these angles, the SIC orbitals have similar shapes as in the case of the closed-shell He_2 molecules (See Fig. 4.4-b)), and are thus named as $(\Phi_{A,\mathbf{I}}, \Phi_{B,\mathbf{I}})$

$$\Phi_{A;B,\mathbf{I}} = \frac{1}{\sqrt{2}}[1\Sigma_{g,\mathbf{I}} \pm 1\Sigma_{u,\mathbf{I}}] \quad . \tag{4.69}$$

I then applied the orbitals into the SIC constraints under the PBCs (See Eqs. (4.36) and (4.36)). The relevant absolute errors of Constraints 2 or 3 quantify the deviation of possible orbitals from the right ones. Focusing on Constraint 2, the top panel of Fig. 4.9 shows the absolute error of Constraints 2 on a logarithmic scale; while the bottom panel of Fig. 4.9 shows the deviation from Constraint 3.

Fig. 4.9-top proves that the correct SIC orbitals that minimize the SIE energy

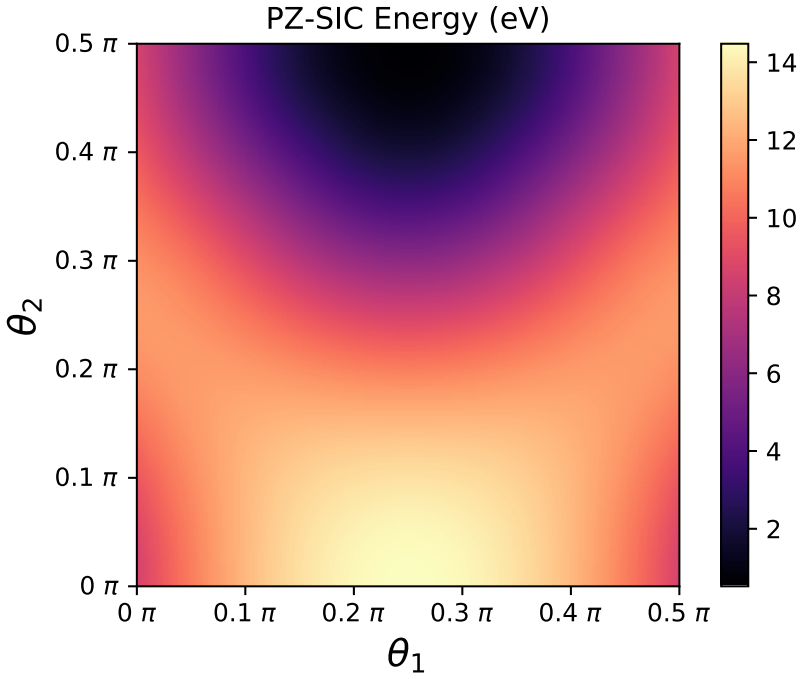

Figure 4.8: The SIC energy E^{SIC} using SCAN XC for the $1 \times 1 \times 2$ Helium cell in a linear combination of the $1\Sigma_{g,\mathbf{I}}, 1\Sigma_{u,\mathbf{I}}$ orbitals (Eq. (4.67)). Periodic boundary conditions and *tight* basis sets are used for the SCAN and SIC-SCAN calculations in *FHI-aims*.

(See Fig. 4.8) satisfy Constraint 2. However, I found another set of orbitals at $(\theta_1 = 0, \theta_2 = 0.25\pi)$ that also satisfy Constraint 2, but are associated to a saddle point on the SIC energy surface (See Fig. 4.8)

$$\begin{aligned} (\Phi_{1,\mathbf{1}}, \Phi_{2,\mathbf{1}}) &= (1\Sigma_{g,\mathbf{I}}, 1\Sigma_{u,\mathbf{I}}) \\ (\Phi_{1,\mathbf{2}}, \Phi_{2,\mathbf{2}}) &= (\Phi_{A,\mathbf{I}}, \Phi_{B,\mathbf{I}}) \quad . \end{aligned} \tag{4.70}$$

As clearly demonstrated in Fig. 4.9-bottom, this solution can be automatically filtered out if both Constraints 2 and 3 are taken into account simultaneously, highlighting again the importance of introducing Constraint 3 during the PZ-SIC self-interaction procedure to avoid the multiple-solution problem.

The next prototypical system is the polyethylene chain, $(C_2H_4)_n$, for which the geometry was provided Fig. 4.10. Unlike the aforementioned model systems, the SIC orbitals cannot be explored exhaustively. In order to uncover the importance of Constraint 3, I generated a set of initial guesses for the SIC orbitals. I constructed the initial guesses for the SIC orbitals by mixing the atom orbitals φ_i centered at the atom position $\mathbf{R}_i^{ac}$ and the Wannier orbitals $\Psi_{i,\mathbf{L}}^{\mathrm{SCAN}}$

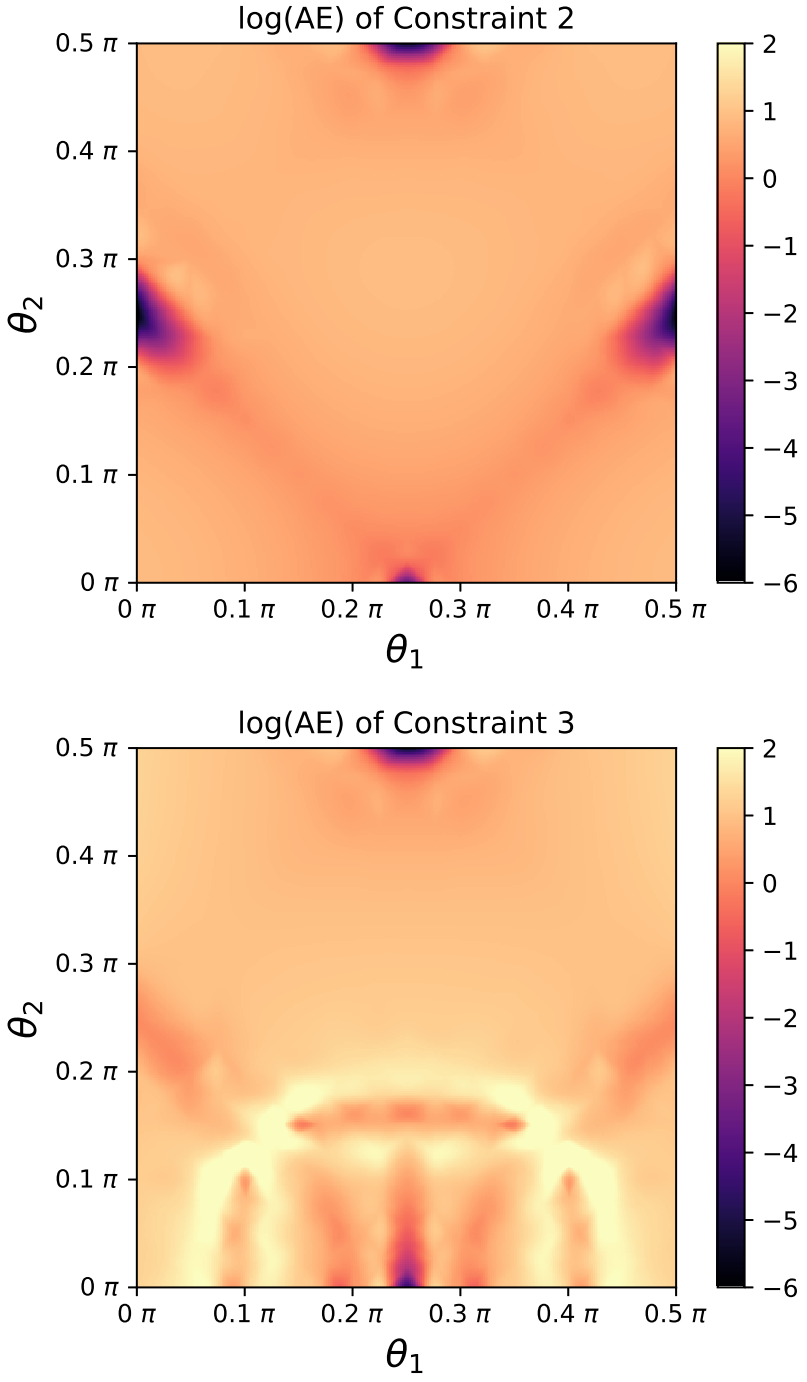

Figure 4.9: The absolute error of SIC Constraint 2 (Eq. (4.36)) (top) and Constraint 3 (Eq. (4.37)) (bottom) for the $1 \times 1 \times 2$ Helium unit cell in a linear combination of the $1\Sigma_{g,\mathbf{I}}, 1\Sigma_{u,\mathbf{I}}$ orbitals (Eq. (4.67)). Periodic boundary conditions and *tight* basis sets are used for the SCAN and SIC-SCAN calculations in *FHI-aims*.

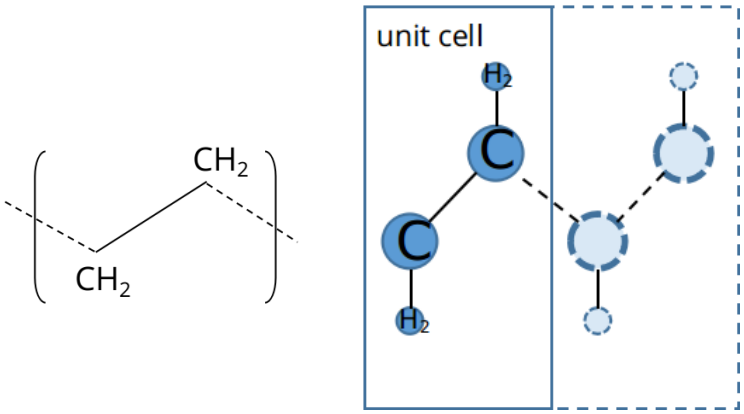

Figure 4.10: (left) The unit cell of periodic polyethylene chain, $-(C_2H_4)-$ with a geometry taken from Ref. [111], and (right) the schematic view of the periodic polyethylene chain with $1 \times 1 \times 2$ unit cell.

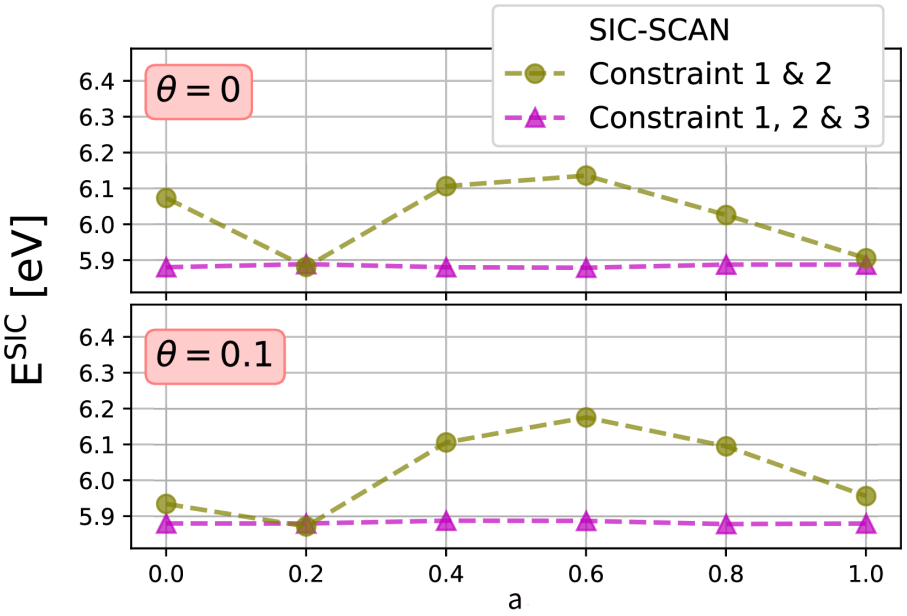

Figure 4.11: The SIC energy (using SCAN functional) with respect to the mixing factor a and the tunable phase $e^{i\theta}$ defined in Eq. (4.71) for the polyethylene chain. Two unit cells and periodic boundary conditions are used for the SIC and SCAN calculations in *FHI-aims* with *tight* basis sets.

transformed from the KS orbitals $\psi_{i,\mathbf{k}}^{\mathrm{SCAN}}$ after the standard SCAN calculations,

$$\begin{aligned}\phi_{i,\mathbf{I}}^{\mathrm{init}} =& \frac{1}{A_i}\left[\, e^{i\theta} a \varphi_i(\mathbf{r} - \mathbf{R}_i^{ac} - \mathbf{R}_\mathbf{I}) \right. \\ & \left. + (1-a)\Psi_i^{\mathrm{SCAN}}(\mathbf{r} - \mathbf{R}_\mathbf{I}) \,\right] \quad ,\end{aligned} \tag{4.71}$$

with A_i being the normalization factor, the tunable phase $e^{i\theta}$, and the mixing factor a ranging from 0 to 1. The atom orbital, surrounding the atom center $\mathbf{R}_i^{ac} + \mathbf{R}_\mathbf{I}$, is relatively more localized than the DFA molecular orbitals of solids $\Psi_{i,\mathbf{L}}^{\mathrm{SCAN}}$, which distributes in the area of the unit cell $\mathbf{L}$ [149]. With turning up a from 0 to 1, the initial guess $\phi_{i,\mathbf{I}}^{\mathrm{init}}$ varies from the localized orbital φ_i to the relatively delocalized orbital $\Psi_{i,\mathbf{L}}^{\mathrm{SCAN}}$. Searching the solution of Constraint 2 from the vicinity of the different initial guesses can lead to thus different SIC orbitals. As shown in Fig. 4.11, various SIC energies can be obtained for varying initial guesses if only Constraint 2 is taken into account, indicating a serious multiple-solution problem. The discrepancy of the offered SIC energies with each other is about 0.3 eV at most. The tunable phases $e^{i\theta}$ on $\theta = 0, 0.1$ also lead to different converged energies with only Constraint 2 used. The phase brings a new localization uncertainty to the SIC orbitals, which is discussed in the next section (Sec. 4.4). However, this case shows that the SIC solver with Constraint 3 can get rid of the impacts of phase and lead to the same SIC energy. Again, we can see that Constraint 3 guarantees to find the lowest SIC energy regardless of what initial guess was used.

Besides the SIC energy, the trace of the SIC potential $\mathrm{Tr}(v^{\hat{\mathrm{SIC}}})$ is also important

$$\mathrm{Tr}(v^{\hat{\mathrm{SIC}}}) = \sum_i \langle \phi_i | v_i^{\mathrm{1eSIC}} | \phi_i \rangle \quad . \tag{4.72}$$

This is because its accuracy determines the accuracy of the corrected KS eigenvalues

$$\sum_l \epsilon_l^{\text{SIC-DFA}} = \sum_l \epsilon_l^{\mathrm{DFA}} + \mathrm{Tr}(v^{\hat{\mathrm{SIC}}}) \quad . \tag{4.73}$$

Accordingly, it is closely relevant to the numerical convergence of many physical properties of interest, like IPs, EAs, and fundamental band-gaps. To investigate the influence of different initial guesses on the final convergence, three sets of initial guesses were selected, which were constructed according to Eq. (4.71) with the parameters $\theta = 0$ and $a = 0., 1, 0.2$, respectively (denoted as "initial 1", "initial 2" and "initial 3"). They respectively stand for the molecular orbitals of solids $\Psi_{i,\mathbf{L}}^{\mathrm{SCAN}}$, the localized atomic orbitals φ_i, and a set of localized atomic orbitals but with slight perturbation from the molecular orbitals. Fig. 4.12 plotted the trace of the SIC potential during the self-consistent procedure. The SIC self-consistent iteration stopped when the maximum constraint error is below 10^{-4} eV. The top panel of Fig. 4.12 plots the convergence of $\mathrm{Tr}(v^{\hat{\mathrm{SIC}}})$ with Constraint 2 only. The guesses of Initial 1 and 2 stop at wrong solutions within about 250 iterations. With a very slow convergence behavior, initial 3 guess reaches the lowest energy compared with those of initial 1 and initial 2. Fig. 4.12-bottom shows the SIC convergence of $\mathrm{Tr}(v^{\hat{\mathrm{SIC}}})$ starting from the aforementioned three initial guesses but using Constraint 3. Within 150 iteration steps, all of them have been converged to the desired threshold, resulting in the same values of $\mathrm{Tr}(v^{\hat{\mathrm{SIC}}})$, completely overcoming the multiple-solution problem observed when only Constraint 2 is used.

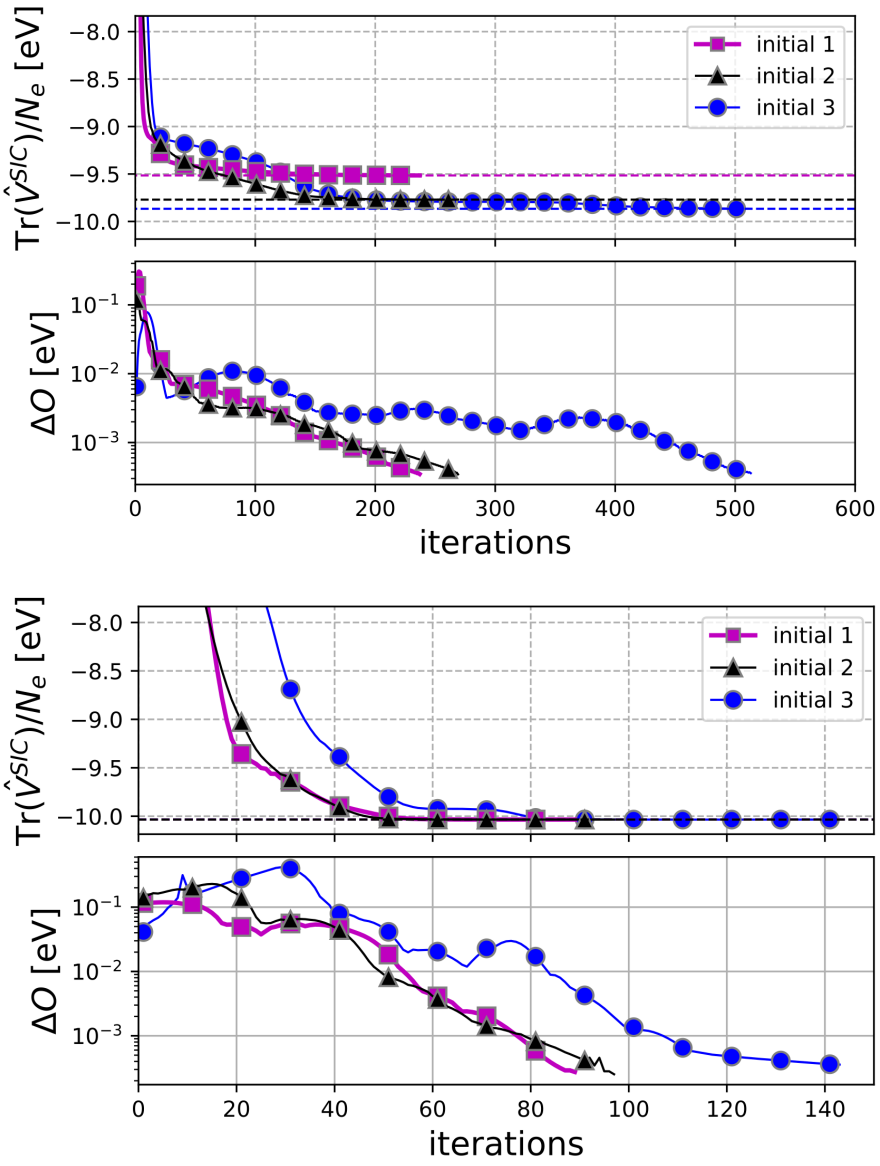

Figure 4.12: The trace of the SIC potential per electron and the max error of SIC constraints with iterations for polyethylene chain using different initial guesses of SIC orbitals. Three sets of initial guesses were selected, which were constructed according to Eq. (4.71) with the parameters $\theta = 0$ and $a = 0, 1, 0.2$, respectively. (top) Self-consistently solving Constraint 2 (Eq. (4.36)) only and (bottom) self-consistently solving Constraint 2 and 3 (Eqs. (4.36) and (4.37)) together are given, respectively. Two unit cells and periodic boundary conditions are used for the SIC and SCAN calculations in *FHI-aims* with *tight* basis sets.

The last example in this section is a more realistic system, i.e., the rocksalt

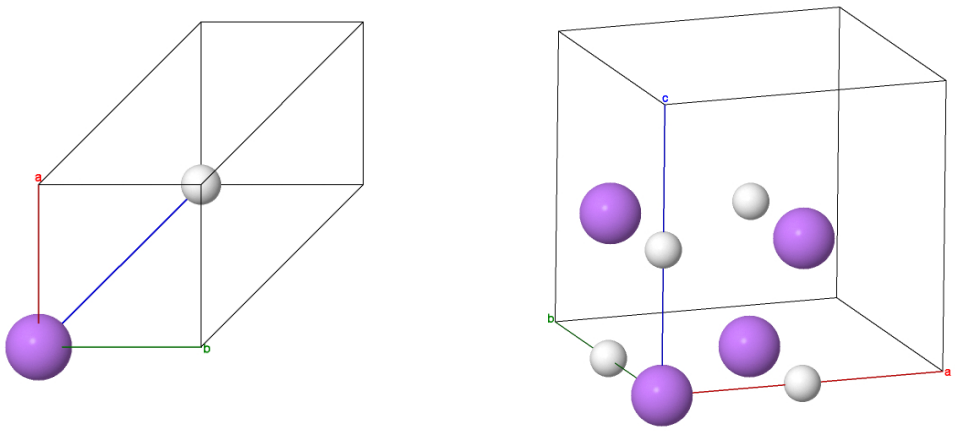

Figure 4.13: (left) The primitive unit cell and (right) the conventional cell of the rocksalt LiH crystal with geometries taken from Ref. [112].

LiH crystal. Fig. 4.13 shows geometries in the primitive unit cell (left) and the conventional cell (right). Instead of SIC energies or SIC potential, the fundamental band-gap was investigated here. In 2021, Shende *et al.* reported that the band-gap calculated by SIC-PBE using Wannier Fermi-Löwdin orbitals is about 3.9 eV for LiH crystal, which is better than the one by PBE [99]. I plotted the numerical convergence of calculated band-gaps with respect to the **k**-grid number in Fig. 4.14. The primitive unit cell was used, and all the relevant calculations were performed using *tier*-1 basis sets and the *tight* integration grid setting. The band-gap of rocksalt LiH calculated by PBE converges to 2.76 eV with **k**-grids $4 \times 4 \times 4$, which is about 2.23 eV lower than the experimental value of 4.99 eV. The initial guess of SIC orbitals is constructed by mixing the KS orbitals and atomic orbitals according to Eq. (4.71) with the parameters $\theta = 0.1$ and $a = 0.2$. However, the SIC-PBE method results in a set of very different band-gaps with different **k**-grid settings without Constraint 3 for the SIC solver. This ill-behaved k-grid convergence should be ascribed to the multiple-solution problems. I proved again here that introducing Constraint 3 indeed addresses the aforementioned problem completely. The SIC-PBE method produces a set of fundamental band-gaps, which converge fast and consistently with respect to the **k**-grid number.

I also investigated the basis-set convergence of the calculated band-gaps, as well as the influence of the unit cell (i.e., primitive cell v.s. conventional cell). A **k**-grid of $4 \times 4 \times 4$ was used. The relevant results were plotted in Fig. 4.15. We can see that the *tier*-1 basis set is good enough to produce numerically well-converged fundamental band-gaps for all the methods, including PBE, SIC-PBE, SCAN, and SIC-SCAN. It is worth noting that the band-gaps calculated with conventional cell (CC) are identical to those in the primitive cell. This observation validates my SIC implementation under the periodic boundary conditions

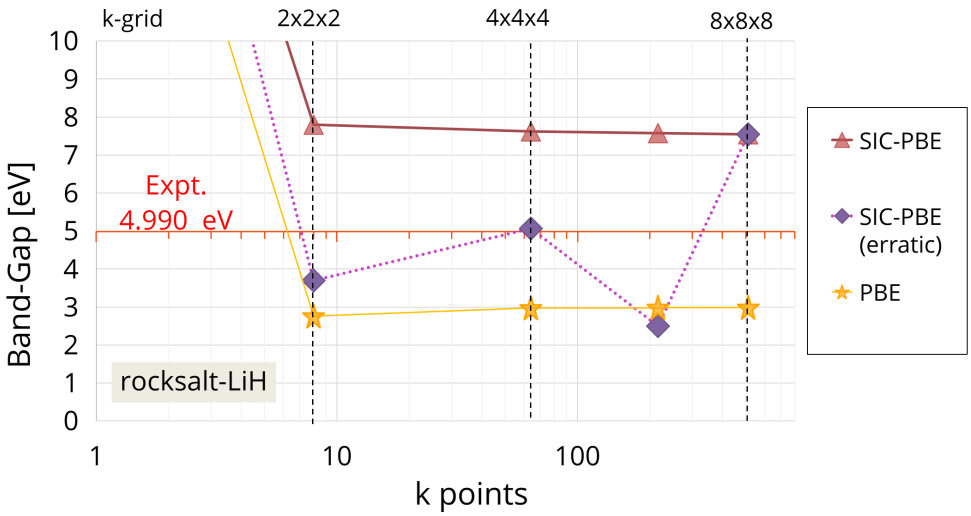

Figure 4.14: The band-gap convergence for rocksalt LiH using PBE and SIC-PBE in primitive unit cell with respect to the **k** points. The corresponding **k**-grids are listed on the top of the x-axis. The SIC-PBE (erratic) represents that the SIC calculations ignore solving the SIC Constraint 3. The PBE and SIC-PBE calculations in *FHI-aims* use the *tight* basis sets. The red line presents the experimental band-gap at T = 4.2 K taken from Ref. [156]. Note that the electron-phonon interaction can impact band-gaps at absolute zero temperature due to the zero-point vibrations [71], but the exact effect is difficult to access from experiments [132]; thus, we ignore it when calculating band-gaps in this thesis. More theoretical studies for this effect on band-gaps can be found in Ref. [133, 134].

and highlights the numerical robustness of my implementation, including size extensively.

To further validate my SIC implementation for solids, I plotted the calculated band-gaps with different **k**-grids in Fig. 4.16. The basis set of *tier*-1 with the *tight* grid setting was used. At first, we can see a good **k**-grid convergence for the calculated band-gap, no matter which DFA (PBE or SCAN) was used. It is also true for SIC-DFAs, including SIC-PBE and SIC-SCAN. Because the size of conventional cell (CC) is larger than the size of the primitive cell, the calculated band-gaps based on conventional cell show an obviously faster **k**-grid convergence than those of primitive cell. However, the calculated band-gaps for a large enough **k**-grid need to converge to a single value, which is independent of the selected cell. In this example, the difference between the band-gaps on conventional and primitive cells gradually reduced when the **k**-grid numbers increased to $2 \times 2 \times 2$. Note that this well-behaved numerical convergence is guaranteed by the periodic boundary conditions, which should be satisfied by

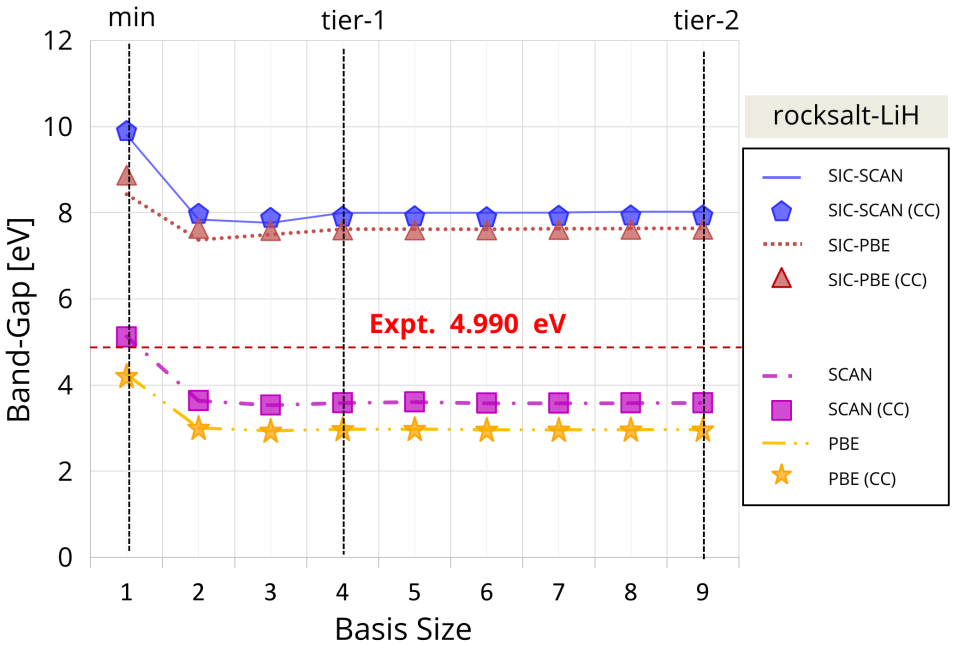

Figure 4.15: Band-gap convergence for rocksalt LiH respectively to basis size (varying from *minimal*, *tier*-1 to *tier*-2). The PBE, SCAN, SIC-PBE, and SIC-SCAN are calculated in primitive cell and conventional cell (marked as (CC)), respectively. The **k**-grids $4 \times 4 \times 4$ are used for all calculations. The red line presents the experimental band-gap at T = 4.2 K taken from Ref. [156], and the zero-point vibration is ignored.

any electronic-structure methods properly implemented. It is good to see that the SIC-PBE and SIC-SCAN results based on my SIC implementation repeat all the aforementioned convergence.

4.4 *Complex* v.s. *Real* SIC Orbitals

In the original PZ-SIC paper, the SIC orbitals were assumed to be a *real* unitary transformation of the KS orbitals [24]. This is because, in molecule systems, the phase of the KS orbitals does not impact the total density and hence does not impact the total energy of semi-local DFAs. However, the same total density can give different SIC orbitals when a *complex* unitary transformation of the KS orbitals (Eq. (4.9) is used. The SIC orbitals using *real*/*complex* unitary transformation is named as *real*/*complex* SIC orbitals in this thesis. It has been issued that the *complex* SIC orbitals can yield a lower or equal total energy than *real* ones [29]. In the recent work (2019), Shahi *et al.* [12] showed that the *real* SIC orbitals have more nodes in the orbital density than the *complex* SIC

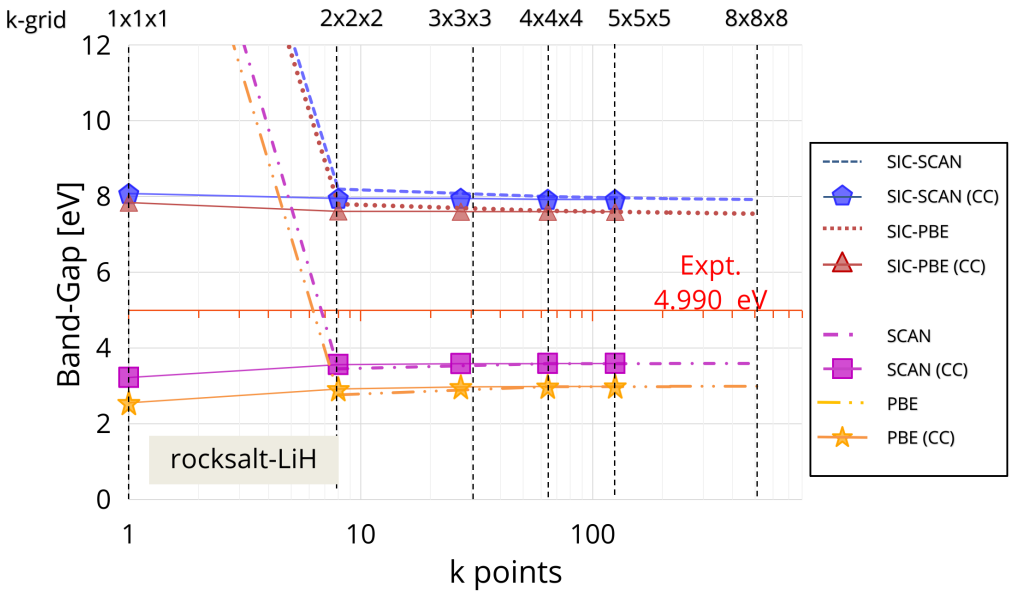

Figure 4.16: Convergence of rocksalt LiH band-gap with the **k** points (corresponding **k**-grids are on the top of the x-axis), which are calculated by PBE, SCAN, SIC-PBE and SIC-SCAN in primitive cell and conventional cell (marked as (CC)). The *tight tier*-1 NAO basis sets are used for all calculations. The red line presents the experimental band-gap at T = 4.2 K taken from Ref. [156], and the zero-point vibration is ignored.

orbitals and that the *complex* SIC orbital densities can thus be more localized than the *real* one. In the following, the G2-1 test set [102] was investigated under SIC Constraints 1-3 using *real* and *complex* SIC orbitals to detect the validity of my SIC implementation. Sequentially, the Wannier SIC orbitals with the *complex* unitary transformation matrix (defined in Eq. (4.32) and named as *complex* Wannier SIC orbitals in this thesis) were proposed and discussed by us in this section.

Fig. 4.17 shows the SIC-SCAN total energy of 55 molecules in the G2-1 test set [102] using *tight tier*-1 NAO basis sets. We found that the SIC-SCAN total energy employing *real* SIC orbitals is always higher than the SCAN total energy and that SIC using *complex* SIC orbitals seems always to reduce the SCAN total energy. In Figure 4.18, a cross-check including 9 molecules was done for our results compared with the SIC results from other published works, including the works of Jussi *et al.* [157] (marked as *ERKALE*) and Sebastian *et al.* [158] (marked as *PyFLOSIC*). The FLO-SIC method is firstly developed in the *PyFLOSIC* package, and the *complex* SIC orbitals are well implemented in *ERKALE*. Although all calculated results in my and other works are slightly different, my results using *real* SIC orbitals lie between the values in *ERKALE*

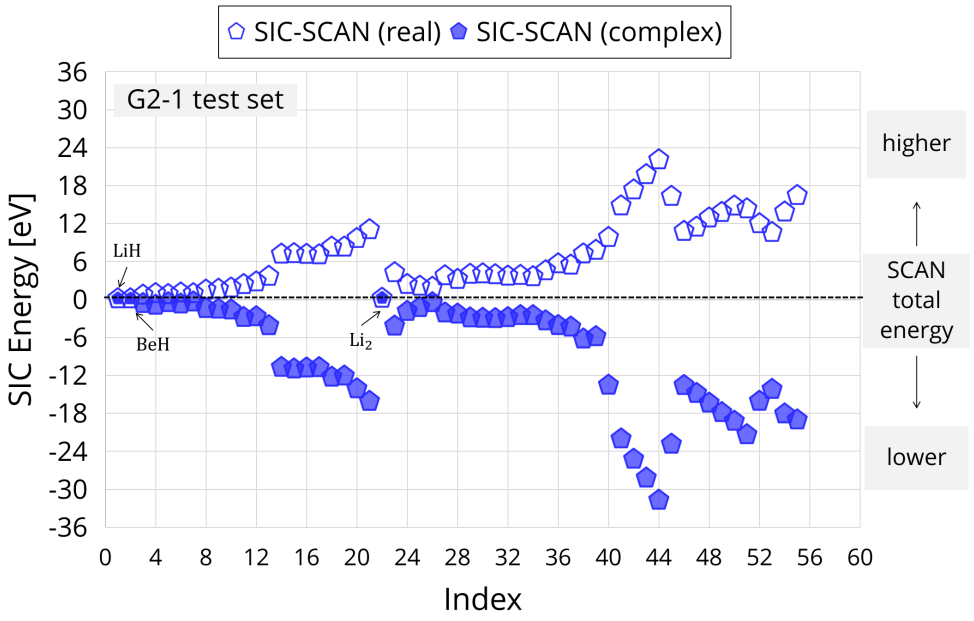

Figure 4.17: The SIC energies for the 55 molecules in the G2-1 test set [102]. SCAN and SIC-SCAN are calculated in *FHI-aims* using the *tight* basis sets. Detailed values of SCAN, SIC-SCAN are listed in Appx. 7.6. The molecules are indexed by 1. LiH 2. BeH 3. CH 4. CH_2(1A1) 5. CH_2 (3B1) 6. CH_3 7. CH_4 8. NH 9. NH_2 10. NH_3 11. OH 12. H_2O 13. HF 14. SiH_2 (1A1) 15. SiH_2(3B1) 16. SiH_3 17. SiH_4 18. PH_2 19. PH_3 20. SH_2 21. HCl 22. Li_2 23. LiF 24. C_2H_2 25. C_2H_4 26. C_2H_6 27. CN 28. HCN 29. CO 30. HCO 31. H_2CO 32. CH_3OH 33. N_2 34. N_2H_4 35. NO 36. O_2 37. H_2O_2 38. F_2 39. CO_2 40. Na_2 41. Si_2 42. P_2 43. S_2 44. Cl_2 45. NaCl 46. SiO 47. CS 48. SO 49. ClO 50. ClF 51. Si_2H_6 52. CH_3Cl 53. CH_3SH 54. HOCl 55. SO_2.

and *PyFLOSIC*. In other words, the trend of the energy cross molecules is the same as others for both *real* and *complex* SIC orbitals.

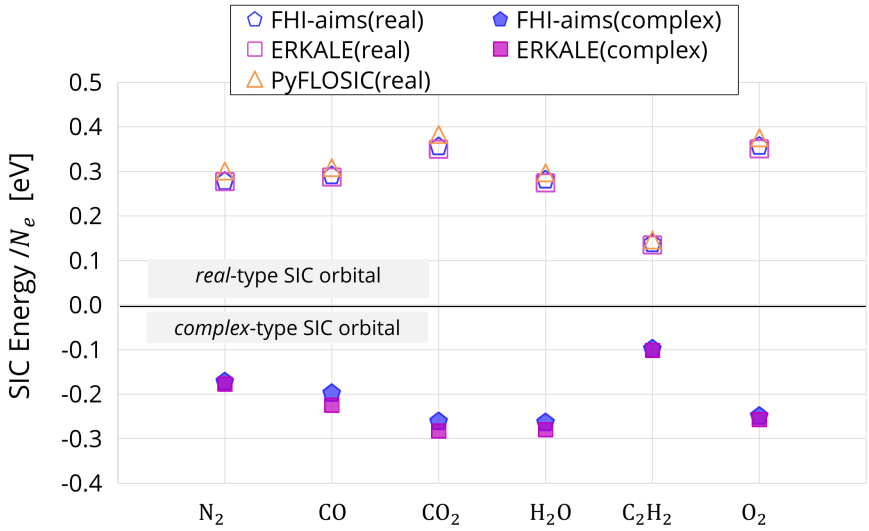

Figure 4.18: A crossing check for SIC-SCAN calculations using cc-pVTZ basis sets in *FHI-aims* and other packages. The *ERKALE* results taken from Ref. [157] use the cc-pVQZ basis sets, and the *PyFLOSIC* results taken from Ref. [158] use the DFO16 basis sets. The geometries of these molecules are taken from G2-1 test set [102].

A more in-depth view of the relative SIC energy with respect to the SCAN energy $E^{\mathrm{SIC}}/|E^{\mathrm{SCAN}}|$ can be found in Fig. 4.19. In this figure, it is clear to see that most SIC-SCAN energy points using *complex* SIC orbitals are located under the neutral line, which presents the SCAN total energy. Note that the points of LiH, BeH, and Li_2 using *complex* SIC orbitals sit above the neutral line because only *real* SIC orbitals in these molecules are found to satisfy all SIC Constraints 1-3 (Eqs.(4.6), (4.12) and (4.23)). From the view of finding the minimal SIC total energy, it is definitely better to choose *complex* SIC orbitals instead of the *real* SIC orbitals. Furthermore, the SIC-SCAN using *complex* SIC orbitals can give smaller deviations of the formation energy for the G2-1 test set, as shown in Figure 4.20. The ME (e.g., mean error) of the SIC-SCAN deviations is 0.21 eV for using *complex* SIC orbitals (SIC-SCAN(complex)) and 0.53 eV for using *real* SIC orbitals (SIC-SCAN(real)), but the ME of the SCAN deviations is only -0.04 eV. The MAEs (e.g., mean absolute errors) of the formation energy for the G2-1 test set are 0.15 eV, 0.54 eV, and 0.29 eV for the calculations of SCAN, SIC-SCAN(real), and SIC-SCAN(complex), respectively. The decline in performance is much smaller when if the *complex* SIC orbitals rather than the the *real* ones are used. However, SIC indeed worsens the SCAN prediction for both *real* and *complex* SIC orbitals, which has been observed

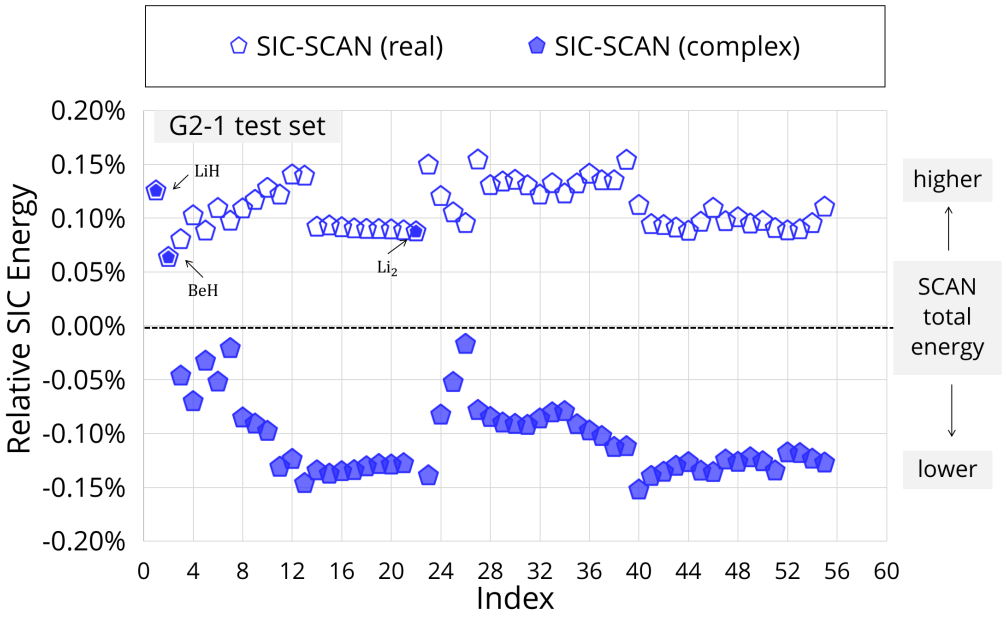

Figure 4.19: Relative SIC energy to the SCAN energies $E^{\mathrm{SIC}}/|E^{\mathrm{SCAN}}|$ of SIC-SCAN using *complex* SIC orbitals and using *real* SIC orbitals for the 55 molecules from the G2-1 test set [102]. SCAN and SIC-SCAN are calculated in *FHI-aims* using the *tight* basis sets.

earlier [12]. As we showed the SCAN calculations of the Ne-Ne^+ dissociation in Sec. 2.3, the accuracy of DFT methods relies on a lucky counterbalance between different kinds of errors in the approximations for cohesive energy of normal systems. Simply removing SIEs from SCAN calculations may break the artful counterbalance and result in poor performance.

Now we know the (Wannier) SIC orbitals must satisfy the SIC constraints instead of being maximally localized (See Sec. 4.3). But the maximally localized Wannier orbitals (MLWOs) are good approximations of the Wannier SIC orbitals and were often used in the early works of developing SIC for solids (See Sec. 4.1). In the past decades, many studies about Wannier orbitals have shown that the MLWOs are always *real*, apart from a trivial overall phase transformation [159, 160]. In 2004, Martin proved that the MLWOs are *real* in one-dimensional crystals [70]. However, we realize that the Wannier SIC orbitals can be *complex* for solids. In the case of solids, the periodic boundary conditions (PBCs) require that the calculation for a molecule sitting in a very large primitive cell must match the calculation for an isolated molecule. Consequently, the impact of applying *complex* SIC orbitals (shown in Fig. 4.18) can be passed from molecules to the periodic systems because the transformation matrix in a very large primitive cell (Eq. (4.32)) is degenerated to Eq. (4.9). In other words, the unitary transformation matrix defined in Eq. (4.32) can have a *complex* domain. Here we calculated the HOMO energy for 8 molecules sitting

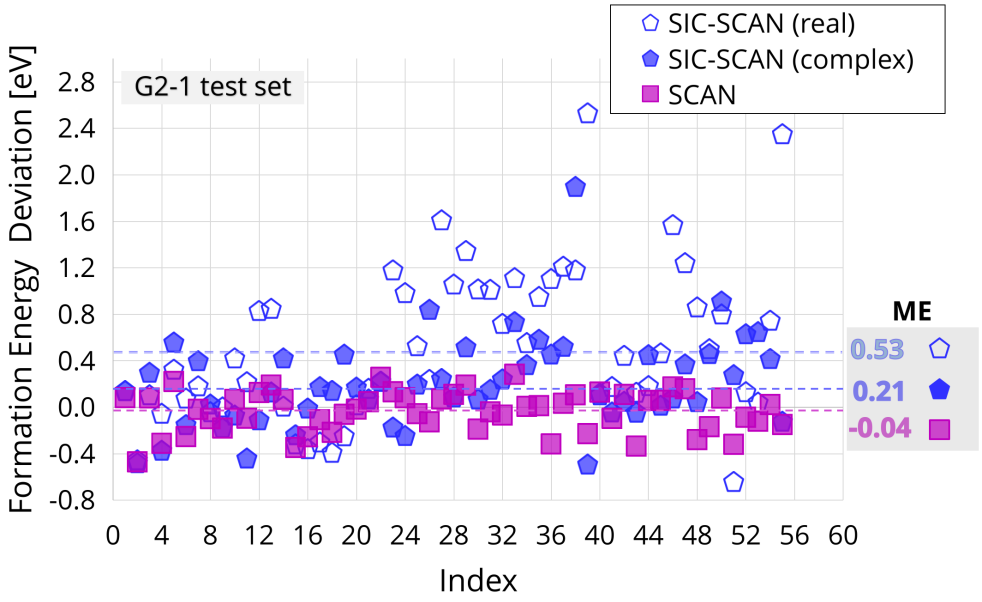

Figure 4.20: Formation energy deviation of the 55 molecules from the G2-1 test set [102], by SCAN, SIC-SCAN using *real* SIC orbitals (marked as SIC-SCAN(real)), and SIC-SCAN using *complex* SIC orbitals (marked as SIC-SCAN(complex)). The index of the molecules is the same as Fig. 4.17. SCAN and SIC-SCAN are calculated in *FHI-aims* using the *tight* basis sets. The mean errors (MEs) of SCAN, SIC-SCAN(real) and SIC-SCAN(complex) are listed on the right side and plotted as the dash lines. Detailed values can be checked in Appx. 7.6.

in a large unit cell (50Å × 50Å × 50Å) and compared with the results putting the molecules in unlimited space, as shown in Figure 4.21. The difference of the HOMO energy in these artificial periodic boundary systems can be achieved more than 0.5 eV for C_2H_2 between using *real* or *complex* Wannier SIC orbitals. These results directly show that the *complex* Wannier SIC orbitals indeed exist and may lead to lower SIC energy in periodic systems (Simultaneously see the results in Fig. 4.19).

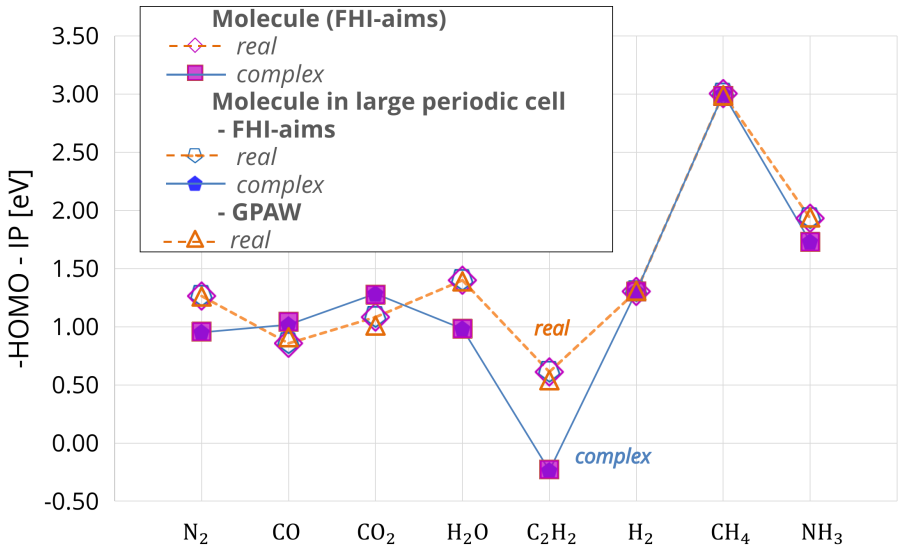

Figure 4.21: The difference between IP and negative HOMO energy calculated by SIC-PBE using *real* and *complex* SIC orbitals for 8 molecules (N_2, CO, CO_2, H_2O, C_2H_2, H_2, CH_4 and NH_3) sitting in a large unit cell (50Å × 50Å × 50Å) and in unlimited space. The calculations in *FHI-aims* use the cc-pVTZ basis sets, and the results of the package *GPAW*, taken from Ref. [99], use the DFT and beyond within the projector-augmented wave method. The *FHI-aims* results with *real* SIC orbitals are well-matched with the *GPAW* results published by Shinde *et al.* in 2021 [99].

We further calculated 5 crystals (diamond Si, diamond C, rocksalt LiH, fcc Ne, and zincblende AlP) to figure out the impact of SIC energy using *real* and *complex* Wannier SIC orbitals. As shown in Figure 4.22, the SIC energy can have as large as 8 eV difference between using *complex* SIC orbitals and *real* SIC orbitals. But the SIC energy using *complex* Wannier SIC orbitals are always lower than or equal to the SIC energy using *real* ones, as in the case of molecules. We found that the SIC energy using *real* is equal with the one using *complex* Wannier SIC orbitals in some systems (like rocksalt LiH). This is because no *complex* Wannier SIC orbitals are found to satisfy all SIC constraints in these systems. From the view of achieving minimal SIC energy, the *complex* Wannier SIC orbitals must be employed for periodic systems. In the following section, the

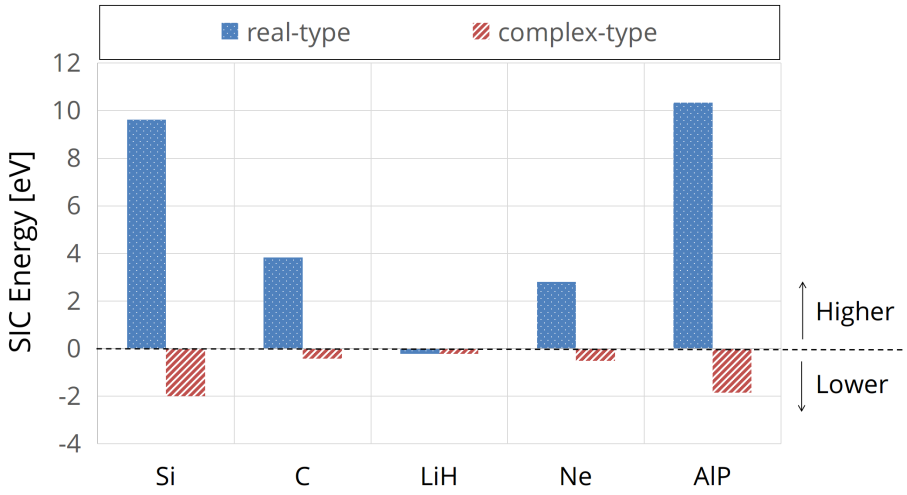

Figure 4.22: SIC energies of SIC-SCAN calculations using *complex* Wannier SIC orbitals and using *real* Wannier SIC orbitals for diamond Si, diamond C, rocksalt LiH, fcc Ne and zincblende AlP. The **k**-grids $2 \times 2 \times 2$ and *tight tier*-1 NAO basis sets are used for these calculations in *FHI-aims*.

performance of SIC on more crystal properties, like energies curve with volume, lattice constant, were investigated for the given crystal set used in Fig. 4.21.

4.5 SIC Performance for Crystal Properties

Figure 4.23 shows the varying energy curve with the volume of diamond C and fcc Ne for SCAN and SIC-SCAN(complex). In these calculations, the initial guess of SIC orbitals was constructed by mixing KS orbitals and atomic orbitals according to Eq. (4.71) with the parameters $\theta = 0.1$ and $a = 0.2$. I found that the neighboring points in the energy curve (like the diamond C volume equal to = 11.3 or 11.9 Å^3) have closing KS orbitals, and the different KS orbitals lead to imperceptibly different initial guesses for the SIC orbitals. However, without Constraint 3 (Eq. (4.37)), the SIC solver can output substantial erroneous energy differences for the neighboring points. For example, Fig. 4.24 shows the SIC-SCAN energies curve only under SIC Constraint 1 and 2. SIC-SCAN energy without Constraint 3 (Fig. 4.24) at volume 11.3 Å^3 is 1.2 eV higher than the energy at volume 11.9 Å^3, while less than 0.2 eV is found for the difference of the SIC-SCAN energy with Constraint 3 (Fig. 4.23-2)) at volume 11.3 Å^3 and volume 11.9 Å^3. It should be noted that a smooth fitted energy curve (shown in Figure 4.23-2) and -4)) can be treated as a severe test for finding the unique set of SIC orbitals. The smooth energy curves of diamond C and fcc Ne in Fig. 4.23 further confirm the necessity of Constraint 3 and the correctness of

the SIC implementation in *FHI-aims*.

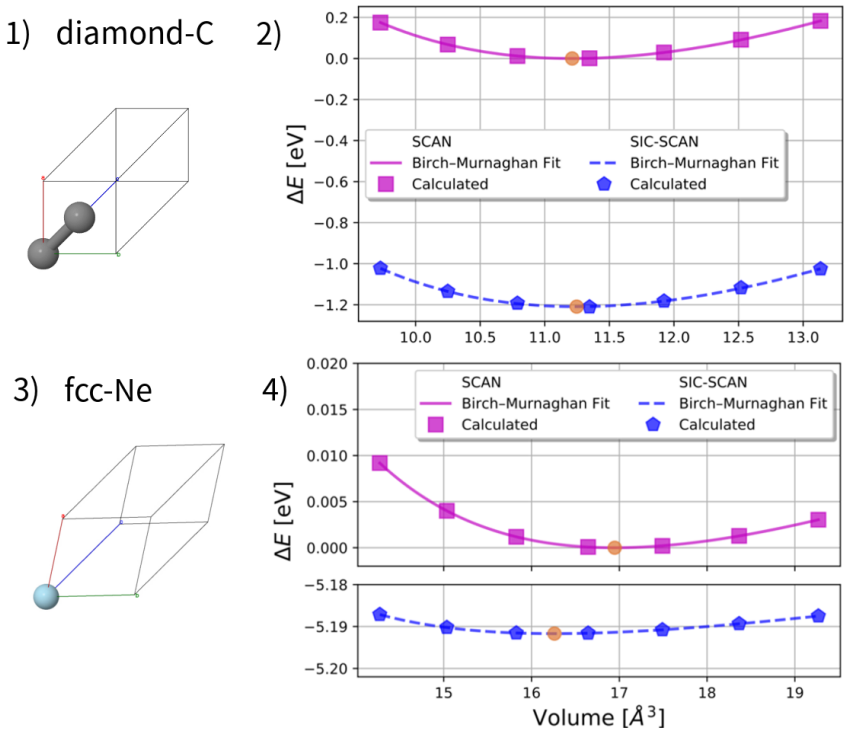

Figure 4.23: The fitted energy curve (per unit cell) of SCAN and SIC-SCAN with respect to volume for diamond C and fcc Ne . 1) and 3) are the structures of diamond C and fcc Ne. The (SIC-)SCAN energies for diamond C and fcc Ne are shown in 2) and 4). The zero energy level is set to the SCAN energy at the optimized volume. The red marker presents the position where the lowest energy exists at the optimized volume. The *complex* Wannier SIC orbitals are used in the SIC-SCAN calculations. The **k**-grids $4 \times 4 \times 4$ and *tight tier*-1 NAO basis sets are used for SCAN and SIC-SCAN calculations.

In practical calculations, varying the energy nearby the equilibrium volume can obtain the properties of the lattice constant, cohesive energy, and bulk modulus correspond to the total energy of the systems:

- The cohesive energy is calculated by the difference of isolated atom energies and the lowest energy in the varying curve, also known as the equilibrium energy;
- The position of the equilibrium energy also indicates the lattice constant of the crystal;
- The bulk modulus is defined as the ratio of the infinitesimal pressure increase to the resulting relative decrease of the volume, which can be calculated by

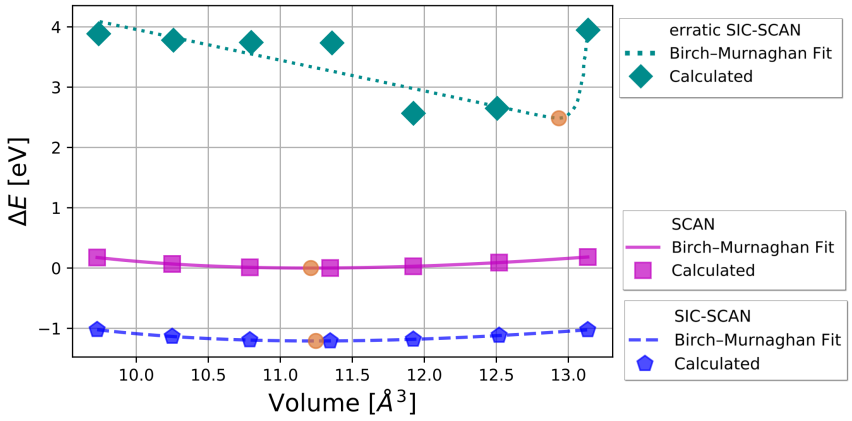

Figure 4.24: The calculated and fitted energy curve (per unit cell) of SCAN, SIC-SCAN and erratic SIC-SCAN calculations with respect to volume for diamond C. The erratic SIC-SCAN calculations are obtained with *complex* Wannier SIC orbitals but ignore solving the SIC Constraint 3. The SIC-SCAN calculations are obtained with *complex* Wannier SIC orbitals and solve all SIC Constraints 1, 2 and 3 (Eqs. (4.6), (4.12) and (4.23)). The red marker presents the position where the lowest energy exists at the optimized volume in the fitted curve. The zero energy level is set to the SCAN energy at the optimized volume. The **k**-grids $4 \times 4 \times 4$ and *tight tier*-1 NAO basis sets are used for SCAN and (erratic) SIC-SCAN calculations.

computing the second derivative of the fitted energy curve with respect to volume.

These three mechanical properties are important in designing and selecting materials [161, 162] and are often used as the benchmarks for new electronic structure methods [112, 161, 163]. Note that the interaction of electrons and phonons can slightly impact these properties at absolute zero temperature due to the zero-point vibrations [71]. The experimental reference values showed in Table 4.1 are corrected from thermal and zero-point vibrational effects. More theoretical studies for the zero-point vibrations can be found in Ref. [71, 132, 133, 164]. In this thesis, the SIC performance for the lattice constant, cohesive energy, and bulk modulus of the 5 crystals (e.g., diamond Si, diamond C, rocksalt LiH, fcc Ne, and zincblende AlP) were investigated using the *complex* Wannier SIC orbitals.

The calculated results of the lattice constant, bulk modulus, and cohesive energy for the test set are listed in Table 4.1. The absolute errors (AEs) are plotted in Figure 4.25 for convenience. SIC slightly worsens all the SCAN calculations for the lattice constant of diamond Si, diamond C, rocksalt LiH, fcc Ne, and zincblende AlP. Specifically, the prediction of SCAN has a considerably large deviation from the experimental lattice constant at 4.455 Å. It has been determined that the error for the fcc Ne is not caused by the SIEs but caused by the difficult description of the weakly bound systems in meta-GGAs, including SCAN [6, 166]. Accordingly, the SIC-SCAN calculation further enlarges the error, giving a prediction of the lattice constant at 4.02 Å (Table 4.1). On the aspect of bulk modulus, SIC breaks the performance of SCAN in diamond Si, diamond C, and rocksalt LiH, while the relative absolute errors of SIC-SCAN keep less than 15% for these cases (4.5%, 5.9%, and 14.8% for each, respectively). Moreover, SIC-SCAN gets slightly better results than SCAN for fcc Ne and zincblende AlP with respect to the experiments. I observed that the SIC method only marginally changes the KS orbitals from the SCAN calculations for diamond Si, diamond C, rocksalt LiH, fcc Ne, and zincblende AlP. As a result, SIC does not worsen the good performance of SCAN in describing the bonding characters for solids, including the covalent bonding and ionic bonding characters [6, 98, 165]. However, our calculations show that the SIC heavily worsens the cohesive energy for the SCAN calculations in our test set besides fcc Ne. The results indicate that simply removing SIEs from SCAN calculations may break the lucky energy counterbalance between different kinds of errors in the approximations of SCAN for solid systems. This conclusion is in line with the previous study, which was also observed not only in solids but in finite molecules like the Ne-Ne^+ dissociation in Sec. 2.3 and the G2-1 test set in Sec. 4.4.

As we realized the SIC-SCAN energy curve calculation for solids, the ability of SIC-SCAN to describe the phase transformation in solids is further invested in this work. The phase of Si is chosen as our prototype. It has been studied in many experimental and theoretical works [167–173] apart from the SIC application. In addition to the well-known diamond phase, a hexagonal diamond

Lattice Constant (Å)			
	SCAN	SIC-SCAN	Expt.
Si	5.45	5.47	5.43
C	3.55	3.51	3.56
LiH	4.01	4.00	4.08
Ne	4.08	4.02	4.45
AlP	5.48	5.49	5.47
Bulk modulus (GPa)			
	SCAN	SIC-SCAN	Expt.
Si	98.89	94.30	98.80
C	454.49	468.22	442.00
LiH	35.85	37.08	32.30
Ne	4.25	3.91	1.20
AlP	90.33	87.54	86.50
Cohesive Energy (eV)			
	SCAN	SIC-SCAN	Expt.
Si	4.68	4.75	4.68
C	7.54	8.00	7.27
LiH	2.47	2.45	2.49
Ne	0.06	0.03	0.03
AlP	4.34	3.83	4.32

Table 4.1: Lattice constants at equilibrium (Å), bulk modulus (GPa) and cohesive energy (eV) of the diamond Si, diamond C, rocksalt LiH, fcc Ne and zincblende AlP for SCAN and SIC-SCAN calculations against the experimental results. The calculations are done in *FHI-aims* using *tight tier*-1 NAO basis set and using **k**-grids $4 \times 4 \times 4$. The experimental reference values are collected from the main-group test set [112, 165] and are corrected from thermal and zero-point vibrational effects.

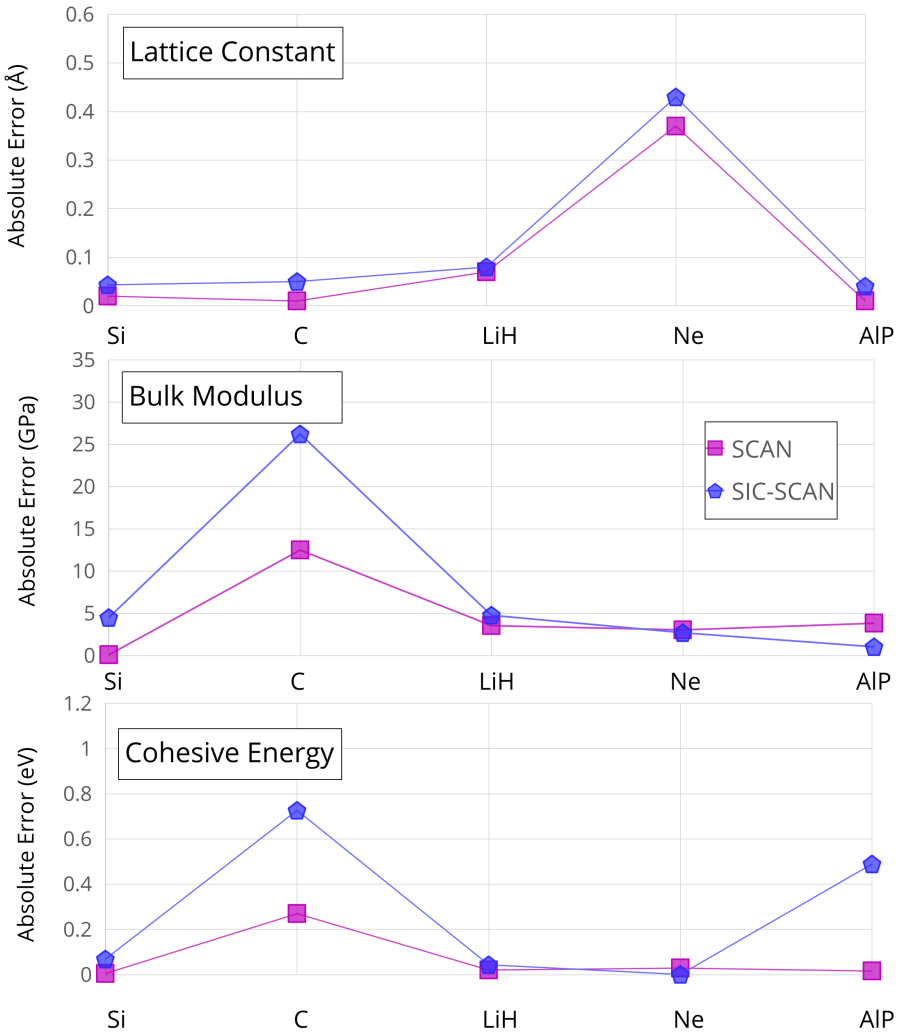

Figure 4.25: Absolute errors of lattice constant (Å), bulk modulus (GPa) and cohesive energy (eV) of SCAN and SIC-SCAN with respect to experiments for the diamond Si, diamond C, rocksalt LiH, fcc Ne and zincblende AlP. The *complex* Wannier SIC orbitals are used in the SIC-SCAN calculations. The **k**-grids $4\times4\times4$ and *tight tier*-1 NAO basis sets are used for SCAN and SIC-SCAN calculations.

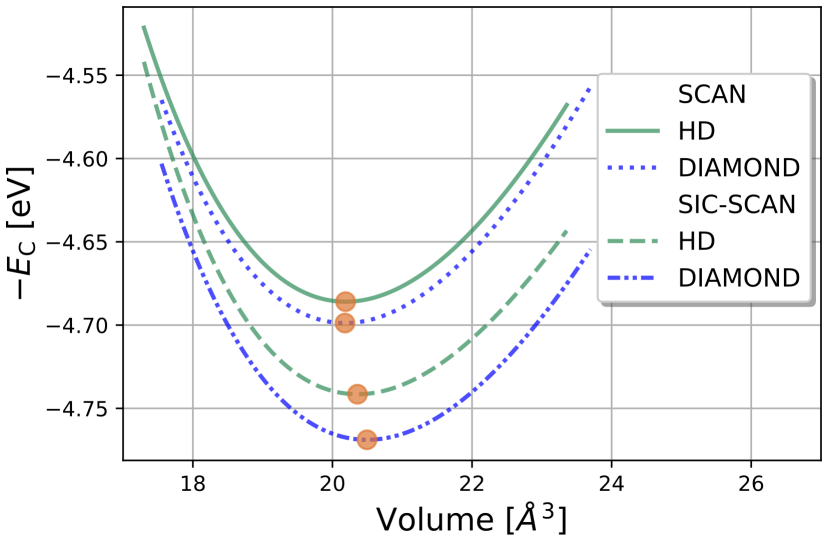

Figure 4.26: The negative cohesive energy (per unit cell) curve of SCAN and SIC-SCAN with respect to volume for diamond Si and hexagonal diamond (HD) Si. The *complex* Wannier SIC orbitals are used in the SIC-SCAN calculations. The red marker presents the position where the lowest energy exists at the optimized volume. The **k**-grids $8 \times 8 \times 8$ and *tight tier*-1 NAO basis sets are used for SCAN and SIC-SCAN calculations.

(HD) structure has been found to be metastable at room temperature and atmospheric pressure after going through high-pressure and high-temperature treatments [168]. The HD structure of Si has been indicated to be a semi-conductive phase as same as the diamond phase [167–170]. The cohesive energy curves with respect to volumes are calculated by (SIC-)SCAN and plotted in Fig. 4.26 for the diamond and HD phases. From the phase energies calculated by the (SIC-)SCAN method, it is apparent that the diamond phase has a lower optimized energy and is more stable than the HD phase. SCAN gives an energy difference between the diamond phase and HD phase at 12.8 meV, which is compatible with the experimental observations and previous results at about 12 meV[167, 174]. However, the SIC method over-enlarges the energy difference to 27.37 meV. It is worth noting that the experiment result indicated that the HD phase for Si has a band-gap at about 1.15 eV [167]. However, only a narrow indirect band-gap at 0.66 eV is given in the SCAN calculation due to the self-interaction error (SIE). The underestimation of the band-gap agrees well with the previous DFA (like LDA, PBE, and PBEsol) calculations [170, 175]. The SIC-SCAN predicts the band-gap at 1.21 eV, which is in agreement with the experiment result. In the next section, a wide of systems, including atoms, molecules, and solids, were further invested in the performance of the SIC method in predicting band gaps.

4.6 HOMO Energy and Band-Gap

As introduced in Sec. 2.3, the many-electron SIE breaks the linear behavior of the total energy of fractional-charge systems (known as the PPLB condition and defined in Eq. (2.44)) and leads to lower total energy calculated with semi-local DFAs. This lower total energy is further linked to the poor performance of the HOMO energy in predicting the ionization potential (IP) [65]. The IP measures the capability of an element to enter into chemical reactions requiring ion formation or donation of electrons. It is also generally related to the nature of the chemical bonding in the compounds formed by the elements. The prediction of IP with the negative HOMO energy can be used as a metric for benchmarking electronic methods [176].

The performance of HOMO energies for the atoms H to Ar using LDA, PBE, SCAN and SIC-SCAN functionals are shown in Fig. 4.27, in which the experimental values [77] are compared as the reference. As the SIE exists in the commonly used LDA and PBE functionals [177] (discussed in Sec. 2.3), the eigenvalues by LDA and PBE can not well-estimate the IPs. Although the gKS framework is used for the meta-GGA SCAN calculations, the performance of SCAN is only slightly better than LDA and PBE according to the mean absolute percentage error (MAPE) as 42.03% for LDA, 41.75% for PBE, and 38.5% for SCAN. SCAN still suffers serious SIEs. As shown in Fig. 4.27-top, the SIC-SCAN calculation gives better estimations with all points under 20% errors and primarily within 10% errors, whether using *real* or *complex* SIC or-

bitals. The MAPE is 5.02% (3.95%) for using *real* (*complex*) SIC orbitals, much lower than SCAN. Moreover, the HOMO energies for the SIC-SCAN(complex) calculations are better than the SIC-SCAN(real) calculations, which agrees with the SIC-PBE(complex) conclusion from the work of Simon *et al.* [29].

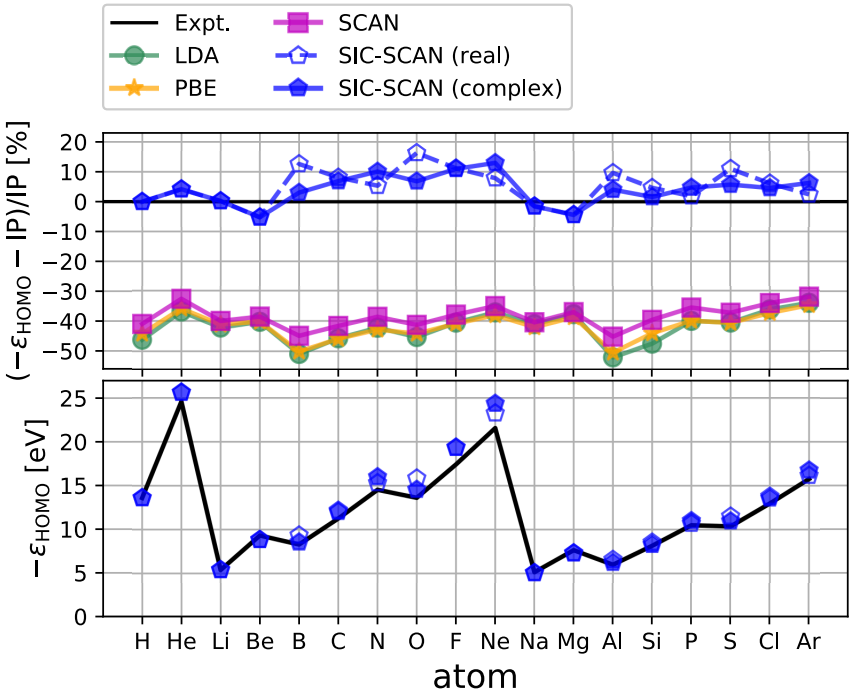

Figure 4.27: Comparison of IPs and the calculated $-\epsilon_{\text{HOMO}}$ energies by LDA, PBE, SCAN and SIC-SCAN of H to Ar atoms. The SIC-SCAN calculations using *real* and *complex* SIC orbitals are marked as SIC-SCAN(real) and SIC-SCAN(complex), respectively. (top) The relative errors of $-\epsilon_{\text{HOMO}}$ energies to IPs and (bottom) the values of $-\epsilon_{\text{HOMO}}$ energies are given. The *tight tier*-1 NAO basis sets are used for all calculations. The black line presents the experimental IPs [77].

Besides atoms, the negative HOMO energies ($-\epsilon_{\text{HOMO}}$) of 24 molecules calculated by LDA, PBE, SCAN, and SIC-SCAN are shown in Figure 4.28. The SCAN calculations heavily underestimate the $-\epsilon_{\text{HOMO}}$ energies against the IPs, while our SIC approach can offer a stable improvement to the HOMO energies by SCAN. All the percentage errors of SIC-SCAN(complex) for the test set are under 23%, and the MAPE of SIC-SCAN(complex) is 11.74%, much less than the SCAN error (34.5%) compared with the experimental IP. G_0W_0@PBE, as reported in Ref. [176], supports reliable predictions for the IPs with 3% MAPE. However, SIC slightly over-corrects the SCAN calculations for the molecule test set, and leads the $-\epsilon_{\text{HOMO}}$ energies are always higher than the experimental IPs. (The cross-comparison for the calculated HOMO energies of SCAN, SIC-SCAN, and G_0W_0@PBE are listed in Appendix Table 7.6.)

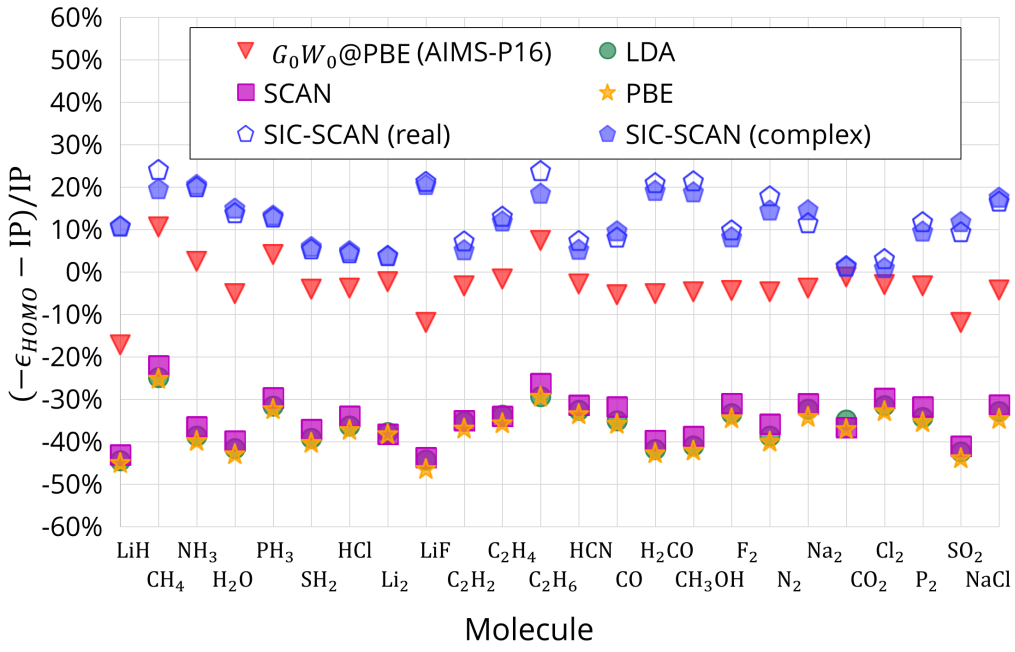

Figure 4.28: Comparison of the calculated eigenvalues of the highest occupied orbital (HOMO) energy ϵ_{HOMO} by LDA, PBE, SCAN, SIC-SCAN and G_0W_0@PBE of 24 molecules with the experimental IP [77]. The LDA, PBE, SCAN, SIC-SCAN are calculated using *tight tier-1* NAO basis sets in *FHI-aims*. The results of G_0W_0@PBE, taken from Ref. [176], are calculated using def2-QZVP basis sets in *FHI-aims*.

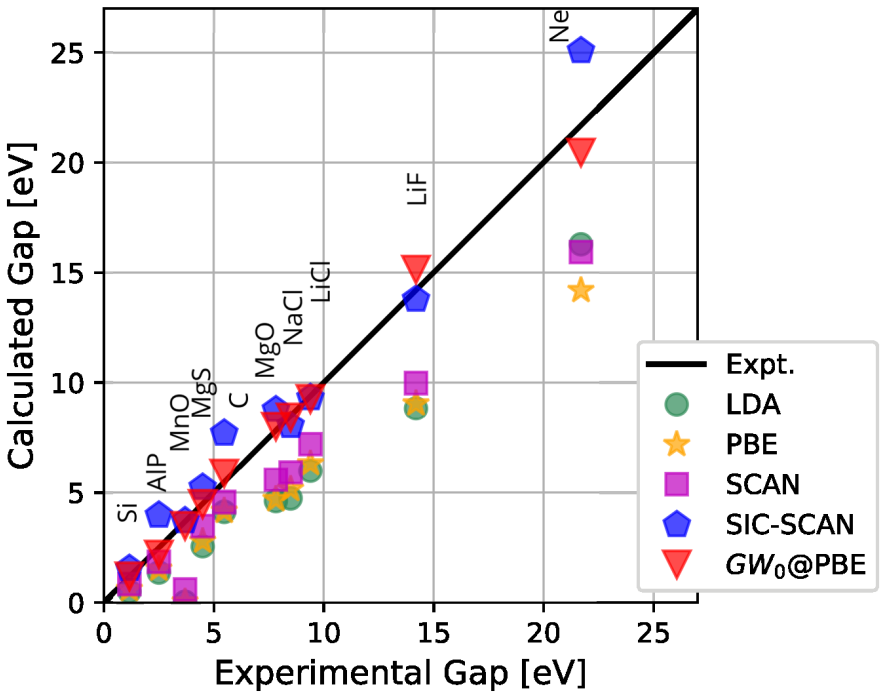

Figure 4.29: Comparison of calculated and experimental band-gaps for LDA, PBE, SCAN, SIC-SCAN, and GW_0@PBE. The black diagonal line corresponds to the calculated line of perfect agreement. The **k**-grids $5\times5\times5$ and *tight tier*-1 NAO basis sets are used for the calculations of LDA, PBE, SCAN, SIC-SCAN in *FHI-aims*. The *complex* SIC orbitals are used for SIC-SCAN. The experimental geometries are used for the LDA, PBE, SCAN and, SIC-SCAN calculations. GW_0 and experiment results are from Ref. [125] and references within.

As the worse performance of the HOMO energy will bring inaccurate descriptions of the band-gaps for semi-local DFAs, including SCAN (discussed and shown in Fig. 3.16), we consider the performance of SIC using *complex* SIC orbitals on calculating electronic band-gaps of the same set validated by SCAN. It is worth noting again that the electron-phonon interaction can impact band-gaps at absolute zero temperature due to the zero-point vibrations [71], but the exact effect is difficult to access from experiments [132]; thus, we ignore it when calculating band-gaps in this thesis. The theoretical studies for this effect on band-gaps can be found in Ref. [133, 134]. Figure 4.29 compares experimental band-gaps with computed values obtained by pure DFAs, SIC-SCAN and GW_0. Specifically, GW_0@PBE results have a small MAPE (e.g., mean absolute percentage error) of 3.22 %. LDA and PBE, which use the KS scheme, predict band-gaps for the test set lie much below the line of perfect agreement shown in Fig. 4.29. LDA and PBE have the MAPEs as 54.96 % and 44.71 %, respectively. SCAN underestimates the band-gaps with a MAPE of 32.82 %, which indicates

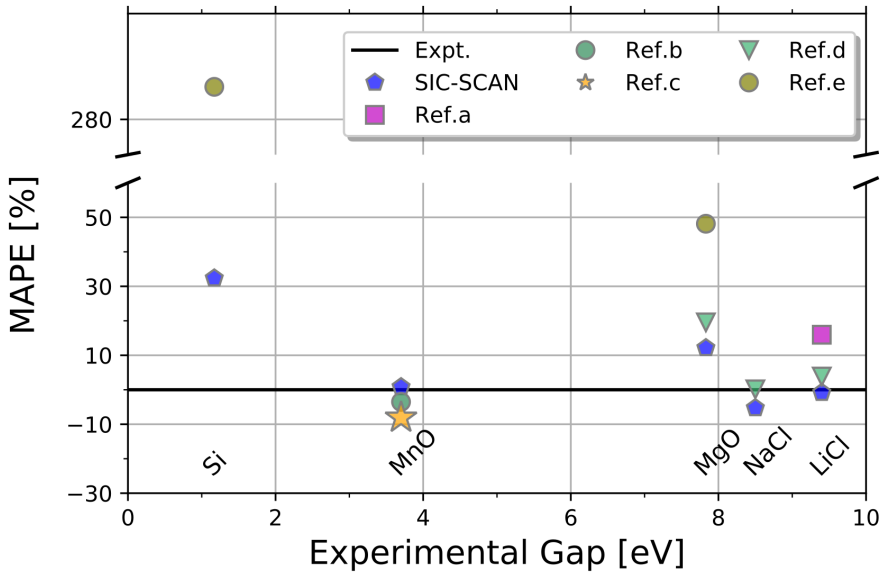

Figure 4.30: Mean absolute percentage error (MAPE) with respect to experimental band-gaps determined from SIC-SCAN by *FHI-aims* and references. The references list as Ref.a [27]; Ref.b [147]; Ref.c [148]; Ref.d [149] and Ref.e [28]. The SIC-LDA calculations are used in these references. In *FHI-aims* calculations, the *complex* SIC orbitals are used for SIC-SCAN. The **k**-grids $5\times5\times5$ and *tight tier*-1 NAO basis sets are used for LDA, PBE, SCAN, and SIC-SCAN. The calculations use the experimental geometries from Ref. [125] and references within.

that SCAN still heavily suffers from SIEs like PBE and LDA. SIC-SCAN predicts better band-gaps than pure SCAN calculations, as shown in Fig. 4.29. All the points of SIC-SCAN stand on or over the line of perfect agreement with respect to experimental values, and the MAPE of SIC-SCAN is 18.5 %. Although LDA, PBE, and SCAN calculations underestimate the band-gaps, they still can correctly predict the insulating properties for the covalent crystals (such as Si, C) and ionic crystals (such as NaCl, LiF). For example, the SCAN calculations agree with the experimental results, in which Si belongs to semiconductor and C (NaCl and LiF) is an insulator. Particularly, SCAN (also LDA and PBE) gives a gap of 0.60eV for MnO, far below the experiment data 3.70 eV. SCAN erroneously predicts MnO as a semiconductor rather than an insulator. This is because MnO, as a Mott insulator, is a strongly correlated material with strong electron repulsions for different spin channels [178, 179]. It is a challenge for semi-local DFAs to describe the strong electron repulsion consequently predict the electric properties of Mott insulators [180]. It has been proposed that DFT+U is a workable but straightforward correction method for this issue, in which an empirical electron repulsion "U" is introduced [181, 182]. For example, the band-gap of MnO calculated by GGA+U varies from 0.9 eV to 2.3 eV using

the electron repulsion "U" value from 0 eV to 7 eV [183, 184]. In comparison, GGA+U gives a band-gap around 4 eV using the "U" (5.25 eV) approximated by the linear-response approach [185]; the self-consistent GW_0 approach based on the LDA+U wave functions shows a band-gap of 3.32 eV [186]. Moreover, a systematical investigation by Yue-Chao Wang et al. in 2016 has shown that the results of DFT+U also depend on the choice of orbitals applying the "U" correction. Specifically, the band-gap of MnO can have more than 0.3 eV differences when using different implementations of PBE+U [184]. However, the SIC-SCAN approach without any empirical factor successfully predicts MnO to be an insulator with a gap of 3.73 eV, in agreement with experimental data of 3.70 eV. One explanation is that SIC can restore the electron repulsions, which are erased mainly by the semi-local approximations [24, 180].

Furthermore, we compared our results with other SIC works for Si, MnO, MgO, NaCl, and LiCl, as shown in Fig 4.30. At variance with our calculation and other references, Ref.a [27] does not use the orbital potential constraint, which gives the highest MAPE startlingly over 45%. SIC calculations in Ref.b [147] and Ref.c [148], which self-consistently solve the potential orbital constraint, give a more accurate prediction of the MnO band-gap with the MAPE less than 10%, while our results are almost in perfect agreement.

Although we have shown that SIC can improve SCAN calculations in many cases, PZ-SIC also over-corrects (meta-)GGA functionals. For example, the SIC-SCAN results in this research show that an obvious overcorrection exists for band-gap predictions (shown in Fig. 4.29). One can alleviate this issue by using the scaled-down PZ-SIC (sd-SIC) method

$$\begin{aligned} E^{\text{sdSIC-DFA}} &= E^{\text{DFA}} + E^{\text{sdSIC}} \\ E^{\text{sdSIC}}[\{n_i\}] &= -\sum_i^{N_e} X_i^\alpha \left(E_{\text{xc}}^{\text{DFA}}[n_i] + E_{\text{es}}[n_i] \right) \quad . \end{aligned} \tag{4.74}$$

The dimensionless scaling factor $0 \leq X_i^\alpha \leq 1$ is evaluated for each SIC orbital $\phi_i(\mathbf{r})$ via

$$X_i^\alpha = \int d^3\mathbf{r} (\frac{\tau^{\text{W}}}{\tau})^\alpha |\phi_i|^2 \tag{4.75}$$

with the kinetic energy density $\tau = \frac{1}{2}\sum_i |\nabla\phi_i|^2$ and von Weizsäcker kinetic energy density $\tau^{\text{W}} = \frac{|\nabla n|^2}{8n}$. The sd-SIC was proposed in 2006 [187] to originally ensure that the SIC term vanishes for the uniform density, in which $\tau^{\text{W}} = 0$ leads $X_i^\alpha = 0$. This vanished SIC term can protect the design of nonempirical XC functionals (such as LDA, PBE, and SCAN) from satisfying the uniform density constraint [188], which is one of the DFA constraints (introduced in Sec. 2.2). It was determined that the scaled-down factor $X_i^\alpha \leq 1$ can help SIC weaken the overcorrection. For example, the errors in total SIC energies of the atoms from Li to Ar are positively reduced by employing sd-SIC in LDA and PBE calculations [187]. Moreover, the sd-SIC-SCAN can provide a good description of the HOMO energies of the atoms from Li to Kr [189].

In 2020, the constant scaled-down factor $X_i^\alpha \equiv 2/3$ was suggested as an approximation to the orbital-dependent factor X_i^α. The constant sd-SIC can improve the prediction of the total energies of atoms, as well as enhance the accuracy of energies for molecules [190]. Inspired by using the constant sd-SIC for periodic systems in the work of Shinde *et al.* [99], we further investigated the band-gaps calculated by the constant sd-SIC using SCAN functional. The results were plotted in Figure 4.31. The sd-SIC-SCAN (giving the MAPE of 13.2 %) is closer to the line of perfect agreement than the SIC-SCAN (MAPE: 18.5 %), while it still lags behind the GW_0@PBE method (MAPE: 3.22 %).

Although the sd-SIC method reduces the over-correction problem, the behavior of $X_i^\alpha \leq 1$ results in that it can not exactly go back to the standard PZ-SIC method and thus is not an exact method for one-electron systems. Moreover, it has been indicated that the SIEs remain in fractional-charge systems, and the PPLB condition (defined in Eq. (2.44)) is broken once again using sd-SIC for semi-local DFAs (such as PBE and TPSS) [64].

4.7 Charge Transfer and Reaction Barrier

As discussed in Sec. 2.3, semi-local DFAs suffering from SIEs prefer a delocalized ground state and predict fractionally charged fragments in the dissociation limit of the neutral heteroatom molecules. In the example of dissociating the molecule H-F (Fig. 4.32), SCAN incorrectly predicts a positive fractional charge on the H atom, hence giving too high energies in the dissociation limit, which agrees with the analysis in Sec. 2.3. SIC successfully cancels the nonphysical charge transfer presented by SCAN and keeps the atoms electro-neutral in the dissociation limit. In this case, the energy curve of SIC-SCAN(complex) matches the curve calculated by CCSD(T) at the minimal energy point ($r \approx 0.8$ Å) and the dissociation limit ($r > 0.8$ Å). In comparison, the energy curve of SIC-SCAN(real) apparently deviates from the one of CCSD(T) in the equilibrium region ($r \approx 0.7$ eV). The incorrect dissociation curve predicted by SCAN also happens in the open-shell case Ne_2^+, as shown in Fig. 4.32, and can be offset by SIC with both *real* or *complex* SIC orbitals to approach the CCSD(T) calculations. Also, SIC does not break the mirror symmetry of the two ions $Ne^{0.5} \cdots Ne^{0.5}$ during the whole dissociation, which gives the same charge transfer as SCAN. Note that the energies in Fig. 4.32-top are zeroed out at the sum of the energies of Ne^+0 and Ne_1^+. The well-behaved energy curve of SIC-SCAN indicates that SIC successfully cancels most part of the many-electron SIEs in Ne_2^+, and SIC-SCAN nearly reproduces the degeneracy between $Ne^+ \cdots Ne$ and $Ne^{0.5} \cdots Ne^{0.5}$ (with less than 0.3 eV difference for both *real* and *complex* SIC orbitals).

Besides the dissociation curves of H-F and Ne-Ne^+, the SIE also impacts other charge transfer problems, like chemical reactions [191]. Specifically, an early investigation of the $F + H_2 \longrightarrow HF + H$ hydrogen transfer by Scuseria found that SIEs bring about 0.23 eV deviation in the reaction energy calculated by the GGA functional (BLYP) [192]. Moreover, Andersson and Grüning, in 2004,

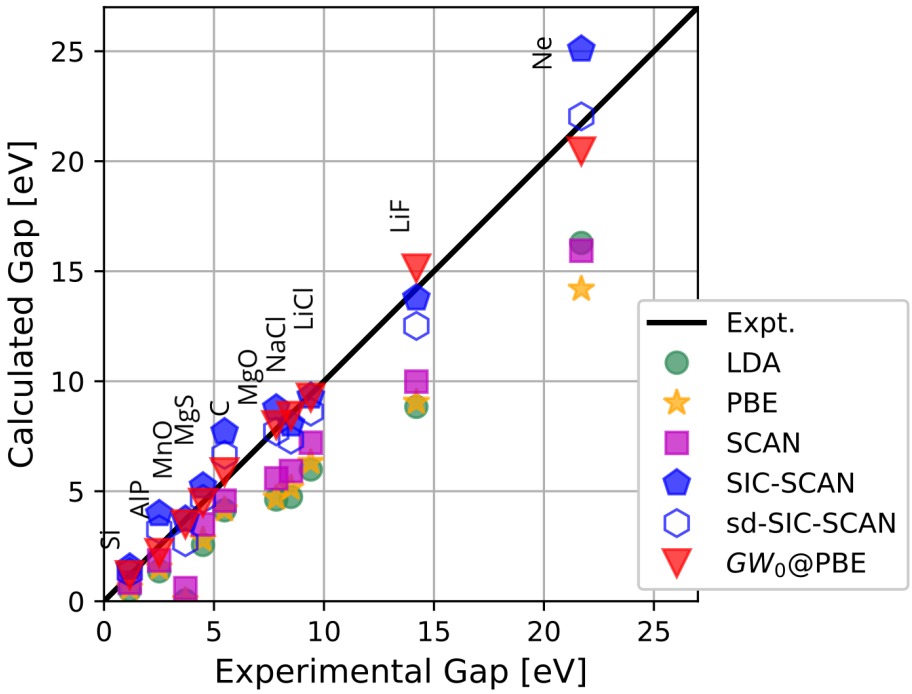

Figure 4.31: Comparison of calculated and experimental band-gaps for LDA, PBE, SCAN, (sd-)SIC-SCAN and GW_0@PBE. The black diagonal line corresponds to the calculated line of perfect agreement. The **k**-grids $5 \times 5 \times 5$ and *tight tier*-1 NAO basis sets are used for the calculations of LDA, PBE, SCAN, (sd-)SIC-SCAN in *FHI-aims*. The *complex* SIC orbitals are used for (sd-)SIC-SCAN. The scaled-down factor $X^{\alpha} \equiv 2/3$ is used in the sd-SIC-SCAN calculations. The experimental geometries are used for the LDA, PBE, SCAN, and SIC-SCAN calculations. GW_0 and experiment results are from Ref. [125] and references within. The effect of the zero-point vibrations is not considered in this thesis. The theoretical studies for this effect on band-gaps can be found in Ref. [71, 133, 134].

found that the hybrid density functional, which employs parts of Hartree-Fock exchange, performs better than other semi-local DFA functionals (including LDAs, GGAs, and meta-GGAs) for calculating hydrogen involved gas-phase reactions [193]. In 2012, Paier *et al.* showed that eliminating the one-electron SIE contained in advanced methods (i.e., the random-phase approximation) is apparently beneficial for observing hydrogen transfer reaction energies. An early investigation for 11 "difficult" reactions (including hydrogen transfer, proton transfer, and so on) indicated that SIC using *real* SIC orbitals [194] can significantly improve the reaction barriers calculated by semi-local DFAs (i.e., PBE). Consequently, I examined all of the hydrogen-transfer reactions included in the BH76 [123, 124, 195] test set using SCAN and our SIC-SCAN(complex). Fig-

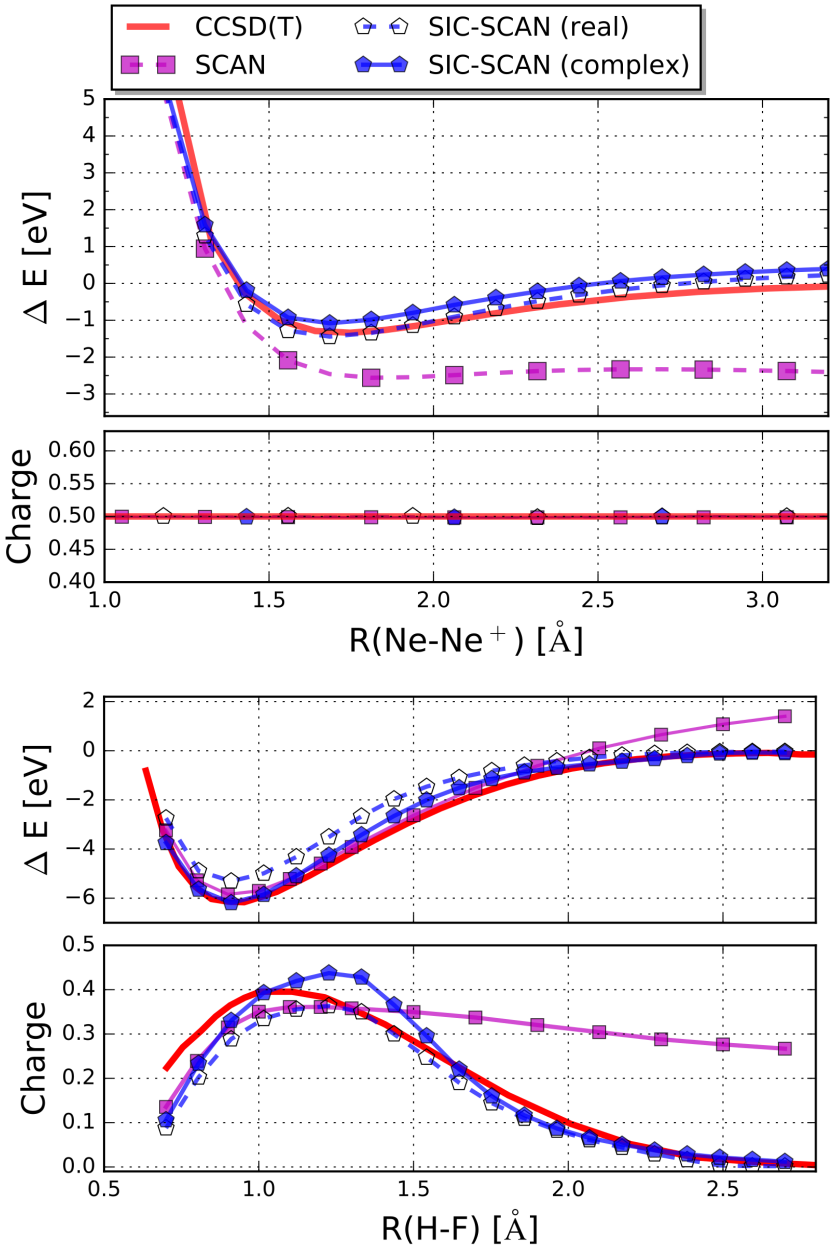

Figure 4.32: Dissociation curves of Ne-Ne$^+$ (top) and H-F (bottom) with the zero-energy level set to the total energy of isolated atoms/ions. The Mulliken charge analyses in the dissociation Ne-Ne$^+$ and H-F are studied, for which the charge transfer to the Ne atom and the H atom are provided. SCAN and SIC-SCAN with *real* SIC orbitals and *complex* SIC orbitals are calculated in *FHI-aims* using *tight* basis sets. CCSD(T) is calculated in *FHI-aims* using cc-pVTZ Gaussian-type basis sets.

ure 4.33 shows the calculated reaction barriers by (SIC-)SCAN with the ones calculated by the Weizmann-1 (W1) method, high-accuracy computational thermochemistry designed to extrapolate to the complete basis limit of a CCSD(T) [129, 130]. (The reactions and calculated reaction barriers with reference values are listed in Appx. 7.8.) SCAN underestimates the barrier heights with a large

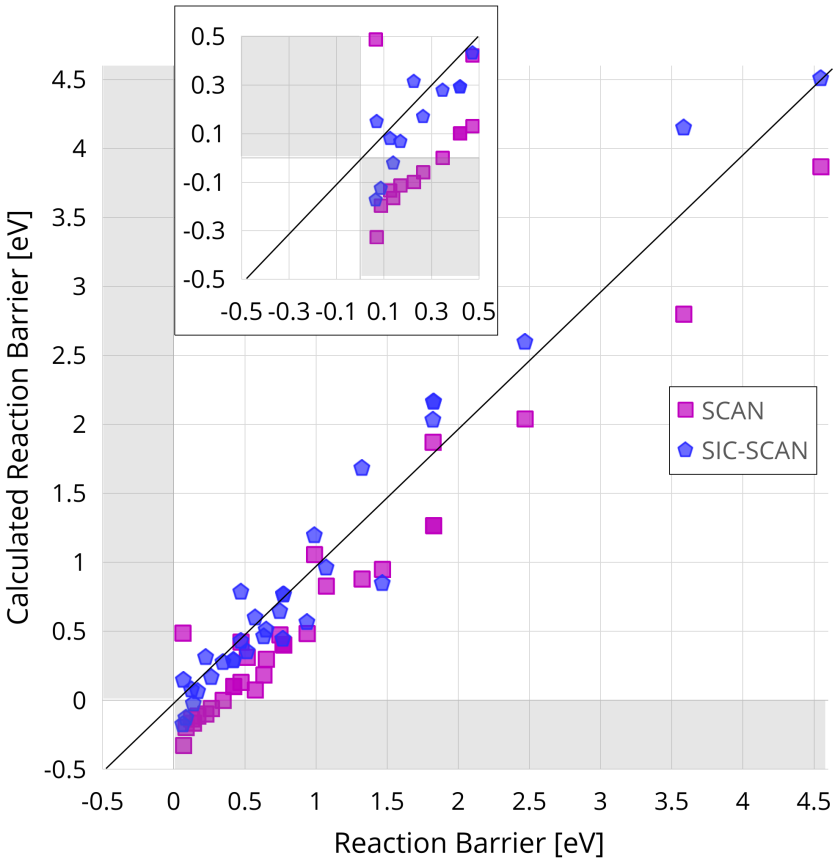

Figure 4.33: Comparison of the calculated results by (SIC-)SCAN and results by the Weizmann-1 (W1) method [129, 130] for the hydrogen-transfer reactions (16 reactions in total) included in BH76 [123, 124]. The black diagonal line corresponds to the calculated line of perfect agreement. *Tight tier*-1 NAO basis sets are used for (SIC-)SCAN calculations in *FHI-aims*. The SIC-SCAN calculation employs the *complex* SIC orbitals. (1 kcal/mol = 0.0434 eV)

MAE (0.365 eV (8.42 kcal/mol)) because of using the semi-local approximation. It is impressive to note that SIC-SCAN provides barrier heights with the MAE of 0.184 eV (4.25 kcal/mol) accuracy, while CCSD(T) is able to have an accurate measure with the MAE of 0.056 eV (1.3 kcal/mol) [196]. The positive performance of SIC-SCAN on the reaction barriers indicates that removing self-

interaction is important to the meta-GGAs, exceptionally the SCAN functional, in reorganizing the ground state of transition states.

The insert of Fig. 4.33 shows a deep-in view of the calculations to check the grey area, where SCAN and SIC-SCAN predict the improper reaction directions. There are 8 calculated points of SCAN sitting in the grey area, but the improper predictions have been reduced by SIC down to 3 reactions (e.g., the forward reactions in ($H + F_2 \rightarrow HF + F$) , ($H + CO \rightarrow HCO$) and ($H + C_2H_4 \rightarrow C_2H_5$)). The reaction ($H + F_2 \rightarrow HF + F$) is the only one where SIC-SCAN predicts the

	Reaction Barrier [eV]				
Reaction	Ref.	PBE	SIC-PBE	SCAN	SIC-SCAN
$H + F_2 \rightarrow HF + F$	1.5	-0.408	0.170	0.487	-0.214

Table 4.2: Comparison of the forward barrier heights (in eV) of the hydrogen-transfer reactions for ($H + F_2 \rightarrow HF + F$). W1 results [129, 130] are used as references and marked as "Ref." in the table. *Tight tier*-1 NAO basis sets are used for (SIC-)PBE and (SIC-)SCAN calculations. The SIC-PBE and SIC-SCAN calculations employ the *complex* SIC orbitals. (1 kcal/mol = 0.0434 eV)

transition state (HF_2H) having lower energy (about -4.93 eV; -113.59 kcal/mol) than the sum energy of reactants (H and F_2), different from the right prediction of the relationship in SCAN calculation. However, SIC reduces the enormous absolute difference of the barrier by SCAN (0.487 eV; 11.23 kcal/mol) and the W1 result (0.065 eV; 1.50 kcal/mol). The results for the reaction barrier ($H + F_2 \rightarrow HF_2H$) of the PBE, SCAN, SIC-PBE, and SIC-SCAN calculations are listed in table 4.2. The PBE calculation predicts lower energy of HF_2H than the sum energy of H and F_2, but SIC successfully turns the PBE calculation right with a positive barrier (0.170 eV; 3.92 kcal/mol). Although SIC can systematically improve SCAN performance of reaction barriers, the reaction ($H + F_2 \rightarrow HF + F$) indicates that SCAN as a semi-local DFA functional still has other kinds of errors besides SIEs, which echos the analysis when calculating the formation energy for the G2-1 test set in Sec. 4.4. For the reaction (H_2+ Cl $\rightarrow$ HCl + H), SCAN leads to the total energy of the transition structures being slightly less than the energy of the separated reagents -0.000 eV (-0.00 kcal/mol). This result suggests a barrier-less process, contrary to the results of the W1 calculation, which indicates a barrier of 0.347 eV (8.0 kcal/mol). Standard hybrid functionals also predict small or even negative activation barriers for this reaction [197]. In contrast, it was reported that the barriers of SIC-revPBE(real) are too high, overestimating about 0.174 eV (4 kcal/mol) [194]. In this work, SIC recovers the barrier calculated by SCAN to 0.276 eV (6.36 kcal/mol), resulting in an almost quantitative agreement with the W1 value. Another example is the reaction of ($H + PH_3 \rightarrow H_2$+ PH_2). SCAN predicts the supposed transition state has -0.135 eV (-3.10 kcal/mol) than the total energy of ($H + PH_3$), while

	Reaction Barrier [eV]		
Reaction	Ref.	SCAN	SIC-SCAN
H_2+ Cl → HCl + H	0.347	-0.000	0.276
H + PH_3 → H_2+ PH_2	0.126	-0.135	0.071

Table 4.3: Comparison of the forward barrier heights (in eV) of the hydrogen-transfer reactions for (H_2+ Cl → HCl + H) and (H + PH_3 → H_2+ PH_2). W1 results [129, 130] are used as references and marked as "Ref." in the table. *Tight tier*-1 NAO basis sets are used for (SIC-)SCAN calculations. The SIC-SCAN calculations employ the *complex* SIC orbitals. (1 kcal/mol = 0.0434 eV)

the W1 calculated value is as large as 0.126 eV (2.9 kcal/mol). SIC-SCAN can rectify the behavior by giving a reaction barrier as 0.071 eV (1.64 kcal/mol) is close to the W1 value.

4.8 *trans*-Polyacetylene

We further investigate a challenging problem for periodic systems, the broken symmetry in crystalline polyacetylene $(CH)_x$. It has been shown that the polyacetylene chain, as the simplest linear conjugated polymer, has two configurations, named *cis* and *trans* [198]. In 1957, Ooshika confirmed that the *trans* configuration is a stable configuration for the long-chain polyacetylene [199, 200], in which an alternation between single and double carbon bonds is formed along the chain sketched out in Fig. 4.34-(4). The dimerization Δz is used to describe the length difference of the two kinds of bonds along the chain direction z_s and z_d:

$$\Delta z = z_s - z_d \quad , \tag{4.76}$$

where z_s and z_d are the vertical distances from the center Carbon atom on both Carbon-atom sides. The X-ray studies show that each of the two carbon chains in the crystalline *trans*-polyacetylene (TPA) has a broken symmetry like the 1D long-chain polyacetylene, shown in Fig. 4.34-(1). For crystalline TPA, experiments have measured the dimerization of $\Delta z \approx 0.05$ Å. This dimerized realization has been described by the celebrated 1D Su-Schrieffer-Heeger (SSH) model [202], in which the electrons are treated in the tight-binding approximation. From the view of phonon/electron-phonon interaction, the SSH model has successfully explained the formation of the alternate bonds and the charge transfer mechanism between carbons [202]. So far, the SSH model has been one of the most basic and most important models describing band topology in condensed matter physics [198, 203]. However, early investigations in the 1990s indicated that the electron-electron interactions in crystalline TPA can also lead to dimerization [204]. From a view of the electron-electron interaction, the hybrid functional in DFT calculations can predict a dimerization in agreement with the experimental studies [204]. But the semi-local DFT calculations (i.e., LDA and BLYP functionals) have problems reproducing the correct ground state of $(CH)_x$ [205] because of suffering serious self-interaction errors [206].

This thesis inspected the SCAN and SIC-SCAN performance for the crystalline TPA from the view of the electron-electron interactions. We construct the unit cell based on the experiment measurement [201] with the lattice vectors $\mathbf{a} = 4.18$ Å, $\mathbf{b} = 7.34$ Å and $\mathbf{c} = 2.46$ Å. The angle β between the lattice vectors $\mathbf{a}$ and $\mathbf{b}$ are set as 90°. The angle between any carbon chain and the lattice vector $\mathbf{b}$ is set as $\phi = 55°$. We keep the length of $z_s + z_d$ equal to the lattice vector length $|c|$ and set 0.70 Å as the vertical height h (the middle carbon atom to the connection of the other two carbon atoms). Figure 4.35 shows the crystalline TPA energy of SCAN, SIC-SCAN, PBE0, and RPA@PBE with respect to the dimerization Δz. The lowest energy position on each curve points out the optimized dimerization for its calculation method. Here "RPA@PBE" presents the results calculated by the random phase approximation using PBE density [208]. The RPA@PBE curve forms two wells in Fig. 4.35, which means that the pattern of single and double bonds is more stable than the one of bonds

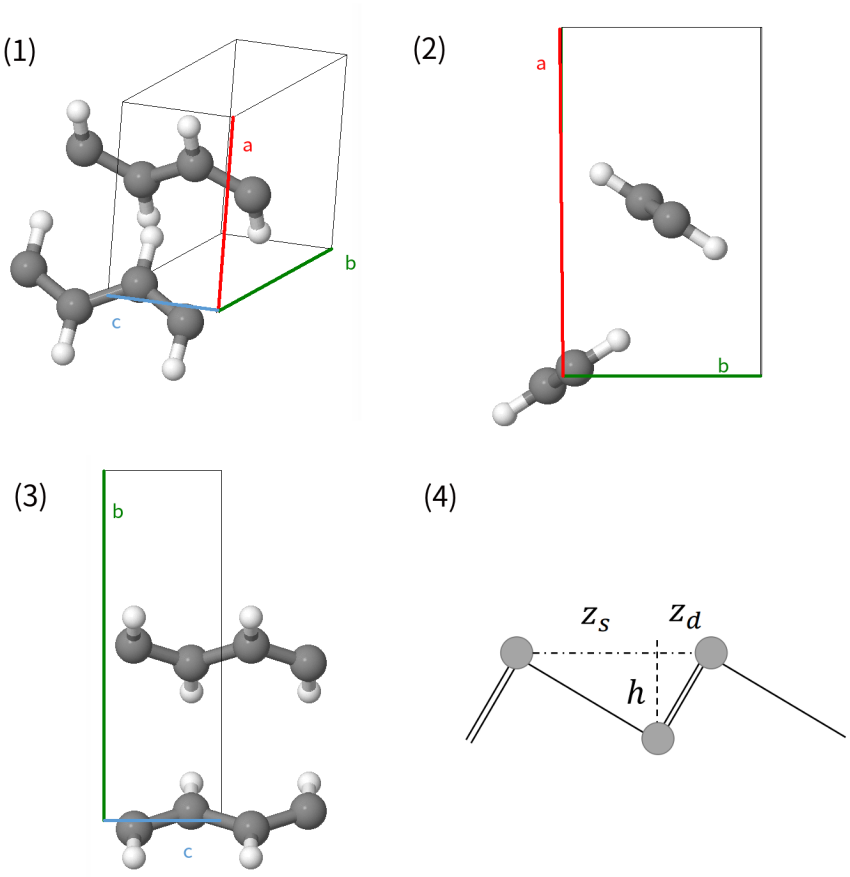

Figure 4.34: (1) Unit cell of TPA structure on the experiment measurement [201]. Carbon atoms are gray, hydrogen atoms are white. (2) The right side view and (3) the bottom side view of the unit cell of TPA. (4) Schematic view of a single $(CH)_x$ chain with the dimerization Δz, which measures the difference between the distance of single-bond and double-bond carbon atoms along the chain direction z_s and z_d.

with equal length. The SCAN calculations show that the optimized structure has the dimerization of $\Delta z = 0.026$ Å, but an inconspicuous well exists in the range of the dimerizations from $\Delta z = 0$ Å to $\Delta z = 0.026$ Å. The hybrid functional PBE0 [209], which employs a degree α of Hartree-Fock exchange admixture in PBE exchange, partly cancels the SIE and reproduces the broken symmetry. An obvious non-zero dimerization is given for both PBE0 employing $\alpha = 0.25$ and employing $\alpha = 0.50$. Specifically, PBE0 with $\alpha = 0.25$ gives the dimerization at 0.050 Å, which is relatively close to the RPA@PBE calculation [208] valued at 0.045 Å. But the optimized dimerization is sensitively shifted to 0.061 Å calculated by PBE0 ($\alpha = 0.50$). The SIC-SCAN using *complex* Wannier

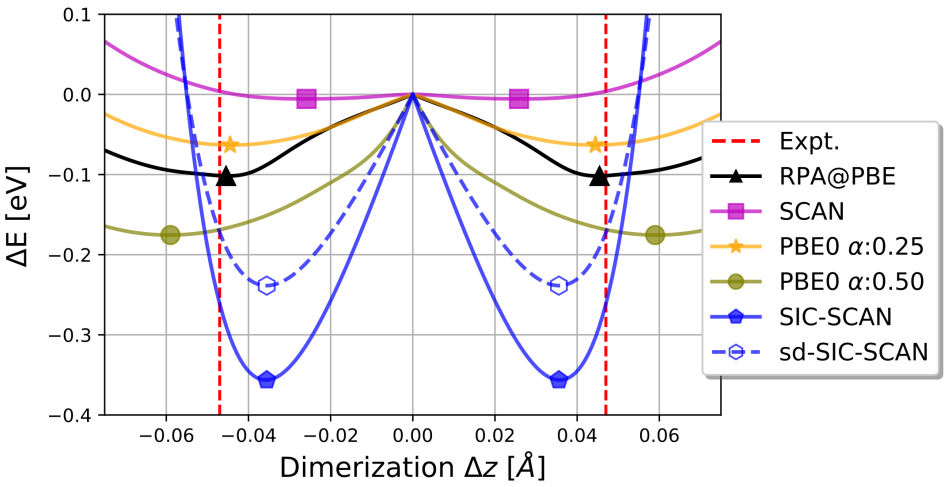

Figure 4.35: The total energy of crystalline *trans*-polyacetylene calculated by SCAN, (sd-)SIC-SCAN, PBE0, and RPA@PBE in *FHI-aims* with respect to the dimerization Δz. Red dashed lines present the experimental dimerization ($\Delta z \approx 0.05$ Å) [207]. The zero-energy level is set to the total energies of $\Delta z = 0$. The marker presents the position where the lowest energy exists at the optimized dimerization. The scaled-down factor $X_i^\alpha \equiv 2/3$ is used in the sd-SIC-SCAN calculations. The **k**-grids $14 \times 12 \times 16$ and *tight tier*-1 NAO basis sets are used for these calculations.

SIC orbitals successfully forms the well in Fig. 4.35 and gives the optimized dimerization at about 0.036 Å, which is the same as the dimerization predicted by the sd-SIC-SCAN with $X^{\alpha} \equiv 2/3$ (defined in Eq. (4.75)). Compared with the SCAN calculations, the well calculated by the SIC-SCAN is much broader and deeper. This conclusion is in line with the previous study [206], which indicates that the SIEs can impact the optimized geometry in this *trans*-Polyacetylene example.

Method	Expt.	LDA	PBE	SCAN
Indirect Band-Gap [eV]	1.1	0.03	0.05	0.12

Table 4.4: The calculated indirect band-gap of LDA, PBE, and SCAN compared with the experiment values. The **k**-grids $14 \times 12 \times 16$ and *tight tier*-1 NAO basis sets are used in the calculations. Experimental values are taken from Ref. [207].

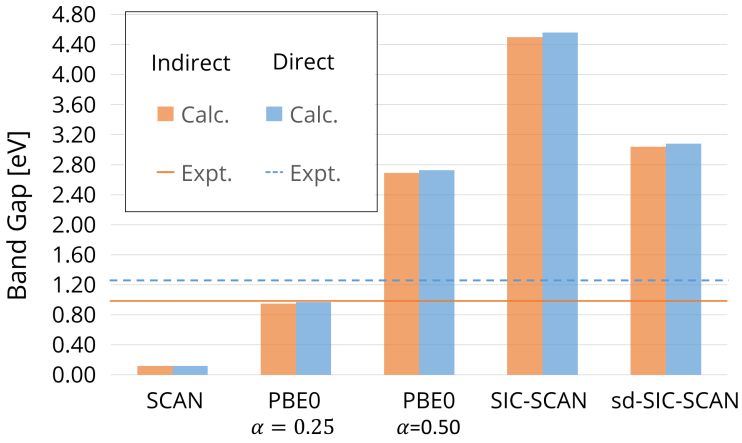

Figure 4.36: Band-gap of crystalline *trans*-polyacetylene at optimized dimmer calculated by SCAN, PBE0 and (sd-)SIC-SCAN in *FHI-aims*. The experimental geometry [207] is used in the calculations. The scaled-down factor $X^{\alpha} \equiv 2/3$ is used in the sd-SIC-SCAN calculations. The **k**-grids $14 \times 12 \times 16$ and *tight tier*-1 NAO basis sets are used. The blue dash line and orange line present the direct and indirect experimental band-gaps, which are from Ref. [207].

Finally, we investigate and characterize the band-gap of crystalline TPA using the experimental geometry [207]. An indirect optical absorption gap of 1.1 eV and a direct gap of 1.4 eV were reported by Fincher *et al.* [207]. However, semi-local DFAs, including LDA, PBE, and SCAN, show that $(CH)_x$ has a very

small band-gap and is a metal or a negligible-gap semiconductor (The calculated indirect band-gaps are shown in Table 4.4). Figure 4.36 shows that PBE0 with $\alpha = 0.25$ gives a slightly underestimated indirect band-gap (0.95 eV), but PBE0 with $\alpha = 0.50$ overestimates the band-gap. SCAN suffering serious SIEs not only gives an evidently narrowed gap as 0.12 eV but also confuses the direct and indirect band-gap of the same value, as shown in Fig. 4.36. SIC-SCAN excessively opens the indirect band-gap to 4.50 eV, rectifying the degeneration of direct and indirect band-gap. Although the scaled-down factor X^{α} can alleviate the overcorrection of SIC, the sd-SIC-SCAN still predicts a larger band-gap compared with the experimental result.

As introduced at the beginning of this section, crystalline TPA is a challenging periodic system. It has been determined that many aspects (such as the lattice constants [210]) can impact its calculated properties. As shown in Fig. 4.35 and Fig. 4.36, the calculations have large differences in energy or band-gap among varied calculated methods. The calculation of SIC-SCAN in this work only implies that reducing the SIE can be one but not the only important aspect of inspecting crystalline TPA from density functional theory.

5 Summary

The state-of-the-art, non-empirical meta-GGA (meta-generalized gradient approximation) SCAN functional provides a notable and consistent improvement over the non-empirical GGA PBE for various kinds of chemical interactions for both molecules and solids. However, it has been well-documented that it suffers from a heavier numerical instability than PBE and inherits the notorious self-interaction error from PBE. In the first part of this work, an efficient and stable prescription for accurate full-potential meta-GGA calculations (including SCAN) of periodic and finite systems has been implemented in *FHI-aims*. I revisited the numerical instability problem of the SCAN functional in the numerical integration framework, *FHI-aims*. From the example of the Ge atom, it is apparent that the SCAN exchange-correlation (XC) potential is much more rugged than PBE, and a denser grid setting is required to converge the numerical integration over the potential. However, I found that the SCAN XC potential is sufficiently captured via the *tight* grid setting in *FHI-aims*, which drove me to generalize the standard electron density mixing algorithm to the kinetic energy density variant. The implementation of the kinetic energy density in the mixing algorithm in *FHI-aims* enables a fast and stable convergence towards a standard SCAN self-consistent solution for fundamental systems, like fcc Si, hexagonal ZnO, bcc Fe, and fcc Fe.

The second part of my work considered the self-interaction correction algorithm, proposed by Perdew and Zunger (PZ-SIC), to alleviate the self-interaction error in the SCAN functional. I have implemented the PZ-SIC algorithm in *FHI-aims* with many promising features. 1) The orbital density constraint, building up together with the total density constraint and the orbital potential constraint, is proposed by us for self-consistently localizing the SIC orbitals and eventually achieving minimal PZ-SIC energy. We have shown that the orbital density constraint can restrict multiple possible SIC orbitals to a single choice, 1.1) by exhaustively checking all candidate SIC orbitals in the systems of B, C, He_2 and hcp-Helium, 1.2) and by verifying the independency of the initial SIC orbitals for the CH-chain and rocksalt LiH. 2) My PZ-SIC implementation enables searching for the SIC orbitals in both *real* and *complex* domains. My work for molecules and solids demonstrates the importance of constructing the (Wannier) SIC orbitals in the *complex* domain. I confirmed that the *complex* SIC orbitals always lead to lower SIC energy than *real* ones in the systems of 55 molecules in the G2-1 test set [102] and solids (including diamond Si, diamond C, rocksalt LiH, fcc Ne, and zincblende AlP). 3) The lowest PZ-SIC energy achievement allows me to explore in my thesis the performance of the SIC-SCAN functional for a wide range of properties with chemical and physical interest. 3.1) Our PZ-SIC approach significantly improves the SCAN performance, as demonstrated for the highest occupied orbital (HOMO) energies ϵ_{HOMO} of H-Ar atoms and 24 molecules. 3.2) Moreover, verifying our PZ-SIC approach in 10 fundamental solids shows that SIC-SCAN is systemically much better than SCAN in pre-

dicting the band-gaps. 3.3) Good energy curves and corrected charge transfers compared with CCSD(T) are obtained by our PZ-SIC approach in the dissociations of H-F and Ne-Ne^+. 3.4) The corrected results in the dissociations drove us to calculate the hydrogen-transfer reaction barriers, which are demonstrated as the SIE disaster area [191, 193, 194]. SIC can effectively rectify the inaccurate estimations of SCAN for the hydrogen-transfer reaction barriers in the BH76 test set[123, 124]. 3.5) In my thesis, I further investigated the ground state of crystalline polyacetylene, which is a famous problem for periodic systems and can not be reproduced by the semi-local DFT calculations (i.e., LDA and BLYP functionals) [205, 206]. The broken symmetry in the crystalline polyacetylene can be captured by our PZ-SIC approach from the view of electron-electron interactions.

6 References

[1] A. Teale and et al., ChemRxiv: **2022-13j2v** (2022).

[2] C. Draxl and M. Scheffler, "Big data-driven materials science and its fair data infrastructure", in *Handbook of materials modeling: methods: theory and modeling* (Springer, Cham, 2020), pp. 49–73.

[3] C. Carbogno, K. S. Thygesen, B. Bieniek, C. Draxl, L. M. Ghiringhelli, A. Gulans, O. T. Hofmann, K. W. Jacobsen, S. Lubeck, J. J. Mortensen, M. Strange, E. Wruss, and M. Scheffler, NPJ Comput. Mater. **8** (2022) 10.1038/s41524-022-00744-4.

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

[5] J. P. Perdew and W. Yue, Phys. Rev. B **33**, 8800 (1986).

[6] E. B. Isaacs and C. Wolverton, Phys. Rev. Mater. **2**, 063801 (2018).

[7] J. Sun, A. Ruzsinszky, and J. P. Perdew, Phys. Rev. Lett. **115**, 036402 (2015).

[8] Y. Zhang, J. Sun, J. P. Perdew, and X. Wu, ArXiv: **1702.04817** (2017).

[9] H. Peng, Z. Yang, J. P. Perdew, and J. Sun, Phys. Rev. X **6**, 041005 (2016).

[10] J. Sun, R. C. Remsing, Y. Zhang, Z. Sun, A. Ruzsinszky, H. Peng, Z. Yang, A. Paul, U. Waghmare, X. Wu, M. L. Klein, and J. P. Perdew, Nat. Chem. **8**, 831 (2016).

[11] Y. Yamamoto, C. M. Diaz, L. Basurto, K. A. Jackson, T. Baruah, and R. R. Zope, J. Chem. Phys. **151**, 154105 (2019).

[12] C. Shahi, P. Bhattarai, K. Wagle, B. Santra, S. Schwalbe, T. Hahn, J. Kortus, K. A. Jackson, J. E. Peralta, K. Trepte, 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. **150**, 174102 (2019).

[13] E. R. Johnson, R. A. Wolkow, and G. A. DiLabio, Chem. Phys. Lett. **394**, 334 (2004).

[14] E. R. Johnson, A. D. Becke, C. D. Sherrill, and G. A. DiLabio, J. Chem. Phys. **131**, 034111 (2009).

[15] S. E. Wheeler and K. N. Houk, J. Chem. Theory Comput. **6**, 395 (2010).

[16] S. Dasgupta and J. M. Herbert, J. Comput. Chem. **38**, 869 (2017).

[17] Z.-h. Yang, H. Peng, J. Sun, and J. P. Perdew, Phys. Rev. B **93**, 205205 (2016).

[18] *VASP manual : LMIXTAU*, https://www.vasp.at/wiki/index.php/LMIXTAU, Accessed: 2021-12-15.

[19] A. P. Bartók and J. R. Yates, J. Chem. Phys. **150**, 161101 (2019).

[20] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, Comput. Phys. Commun. **180**, 2175 (2009).

[21] A. J. Cohen, P. Mori-Sánchez, and W. Yang, Chem. Rev. **112**, 289 (2012).

[22] H. Jiang, R. I. Gomez-Abal, P. Rinke, and M. Scheffler, Phys. Rev. Lett. **102**, 126403 (2009).

[23] M. Lorenz, D. Usvyat, and M. Schütz, J. Chem. Phys. **134**, 094101 (2011).

[24] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[25] A. Ruzsinszky, J. P. Perdew, G. I. Csonka, O. A. Vydrov, and G. E. Scuseria, J. Chem. Phys. **126**, 104102 (2007).

[26] P. Mori-Sánchez, A. J. Cohen, and W. Yang, J. Chem. Phys. **125**, 201102 (2006).

[27] R. A. Heaton, J. G. Harrison, and C. C. Lin, Phys. Rev. B **28**, 5992 (1983).

[28] M. Stengel and N. A. Spaldin, Phys. Rev. B **77**, 155106 (2008).

[29] S. Klüpfel, P. Klüpfel, and H. Jónsson, Phys. Rev. A **84**, 050501 (2011).

[30] S. Lehtola, M. Head-Gordon, and H. Jónsson, J. Chem. Theory Comput. **12**, 3195 (2016).

[31] C. Edmiston and K. Ruedenberg, Rev. Mod. Phys. **35**, 457 (1963).

[32] M. Born and R. Oppenheimer, Annalen der Physik **389**, 457 (1927).

[33] L. H. Thomas, Math. Proc. Camb. Philos. Soc. **23**, 542 548 (1927).

[34] E. Fermi, Rend. Accad. Naz. Lincei **6**, 32 (1927).

[35] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

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

[37] V. Sahni and A. Solomatin, Adv. Quantum Chem. **33**, edited by P.-O. Löwdin, 241 (1998).

[38] J. P. Perdew and K. Schmidt, AIP Conf. Proc. **577**, 1 (2001).

[39] U von Barth and L Hedin, J. Phys. C: Solid State Phys. **5**, 1629 (1972).

[40] S. H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980).

[41] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).

[42] J. P. Perdew, Adv. Quantum Chem., 113 (1990).

[43] N. Harrison, NATO Sci. **187**, 45 (2003).

[44] O. Gunnarsson and R. O. Jones, J. Chem. Phys. **72**, 5357 (1980).

[45] M. S. Hybertsen and S. G. Louie, Adv. Quantum Chem. **21**, 155 (1990).

[46] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **80**, 891 (1998).

[47] J. P. Perdew, Phys. Rev. Lett. **55**, 1665 (1985).

[48] S. K. Ghosh and R. G. Parr, Phys. Rev. A **34**, 785 (1986).

[49] J. P. Perdew, S. Kurth, A. Zupan, and P. Blaha, Phys. Rev. Lett. **82**, 2544 (1999).

[50] S. Pittalis and S. Kurth, *The optimized effective potential method*, Vol. 32 (Computational Condensed Matter Physics, 2006).

[51] A. Görling and M. Levy, J. Chem. Phys. **106**, 2675 (1997).

[52] F. Zahariev, S. S. Leang, and M. S. Gordon, J. Chem. Phys. **138**, 244108 (2013).

[53] A. Seidl, A. Görling, P. Vogl, J. A. Majewski, and M. Levy, Phys. Rev. B **53**, 3764 (1996).

[54] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, Phys. Rev. Lett. **91**, 146401 (2003).

[55] B. Xiao, J. Sun, A. Ruzsinszky, J. Feng, R. Haunschild, G. E. Scuseria, and J. P. Perdew, Phys. Rev. B **88**, 184103 (2013).

[56] A. Ruzsinszky, J. Sun, B. Xiao, and G. I. Csonka, J. Chem. Theory Comput. **8**, 2078 (2012).

[57] J. P. Perdew, A. Ruzsinszky, J. Sun, and K. Burke, J. Chem. Phys. **140**, 18A533 (2014).

[58] J. Sun, J. P. Perdew, and A. Ruzsinszky, Proc. Natl. Acad. Sci. **112**, 685 (2015).

[59] J. Sun, B. Xiao, and A. Ruzsinszky, J. Chem. Phys. **137**, 051101 (2012).

[60] E. Fermi and E. Amaldi, *Le orbite ∞ s degli elementi* (Reale Accademia d'Italia, 1934).

[61] X. Zheng, A. J. Cohen, P. Mori-Sánchez, X. Hu, and W. Yang, Phys. Rev. Lett. **107**, 026403 (2011).

[62] K. Park, M. R. Pederson, and C. S. Hellberg, Phys. Rev. B **69**, 014416 (2004).

[63] M. R. Pederson, R. A. Heaton, and C. C. Lin, J. Chem. Phys. **82**, 2688 (1985).

[64] A. Ruzsinszky, J. P. Perdew, G. I. Csonka, O. A. Vydrov, and G. E. Scuseria, J. Chem. Phys. **125**, 194112 (2006).

[65] V. Atalla, I. Y. Zhang, O. T. Hofmann, X. Ren, P. Rinke, and M. Scheffler, Phys. Rev. B **94**, 035140 (2016).

[66] P. Mori-Sánchez, A. J. Cohen, and W. Yang, Phys. Rev. Lett. **100**, 146401 (2008).

[67] C. Ambrosch-Draxl and J. O. Sofo, Comput. Phys. Commun. **175**, 1 (2006).

[68] F. Caruso, D. Novko, and C. Draxl, Phys. Rev. B **101**, 035128 (2020).

[69] P. Rinke, A. Qteish, J. Neugebauer, C. Freysoldt, and M. Scheffler, New J. Phys. **7**, 126 (2005).

[70] R. M. Martin, *Electronic structure: basic theory and practical methods* (Cambridge University Press, Apr. 2004).

[71] G. D. Mahan, *Many-particle physics* (Springer US, 1990).

[72] L. X. Benedict, E. L. Shirley, and R. B. Bohn, Phys. Rev. Lett. **80**, 4514 (1998).

[73] D. Golze, M. Dvorak, and P. Rinke, Front. Chem. **7** (2019) 10.3389/fchem.2019.00377.

[74] C. D. Clark, P. J. Dean, P. V. Harris, and W. C. Price, Proc. R. Soc. A: Math. Phys. Eng. Sci. **277**, 312 (1964).

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

[76] O. A. Vydrov, G. E. Scuseria, and J. P. Perdew, J. Chem. Phys. **126**, 154109 (2007).

[77] R. Johnson, *Computational chemistry comparison and benchmark database, nist standard reference database 101*, 2002.

[78] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. **157**, 479 (1989).

[79] J. D. Watts, J. Gauss, and R. J. Bartlett, J. Chem. Phys. **98**, 8718 (1993).

[80] T. Shen, Z. Zhu, I. Y. Zhang, and M. Scheffler, J. Chem. Theory Comput. **15**, 4721 (2019).

[81] B. Bakker, M. Goes, N. Hoebe, H. van Ramesdonk, J. Verhoeven, M. Werts, and J. Hofstraat, Coord. Chem. Rev. **208**, 3 (2000).

[82] D. Reynolds, K. Bajaj, C. Litton, E. Smith, P. Yu, W. Masselink, F. Fischer, and H. Morkoc, Solid State Commun. **52**, 685 (1984).

[83] K. Hestroffer, D. Sperlich, S. Dadgostar, C. Golz, J. Krumland, W. T. Masselink, and F. Hatami, Appl. Phys. Lett. **112**, 192107 (2018).

[84] C. Li, X. Zheng, N. Q. Su, and W. Yang, Natl. Sci. Rev. **5**, 203 (2017).

[85] C. Li, X. Zheng, A. J. Cohen, P. Mori-Sánchez, and W. Yang, Phys. Rev. Lett. **114**, 053001 (2015).

[86] N. Q. Su, C. Li, and W. Yang, Proc. Natl. Acad. Sci. **115**, 9678 (2018).

[87] N. Q. Su, A. Mahler, and W. Yang, J. Phys. Chem. Lett. **11**, 1528 (2020).

[88] Y. Mei, C. Li, N. Q. Su, and W. Yang, J. Phys. Chem. A **123**, 666 (2018).

[89] J. M. Foster and S. F. Boys, Rev. Mod. Phys. **32**, 300 (1960).

[90] Y. Mei, Z. Chen, and W. Yang, J. Phys. Chem. Lett. **11**, 10269 (2020).

[91] Y. Mei, N. Yang, and W. Yang, J. Chem. Phys. **154**, 054302 (2021).

[92] P. W. Ayers, R. C. Morrison, and R. G. Parr, Mol. Phys. **103**, 2061 (2005).

[93] M. R. Pederson, R. A. Heaton, and C. C. Lin, J. Chem. Phys. **80**, 1972 (1984).

[94] M. R. Pederson and C. C. Lin, J. Chem. Phys. **88**, 1807 (1988).

[95] M. R. Pederson, A. Ruzsinszky, and J. P. Perdew, J. Chem. Phys. **140**, 121103 (2014).

[96] M. R. Pederson, J. Chem. Phys. **142**, 064112 (2015).

[97] Z. Yang, M. R. Pederson, and J. P. Perdew, Phys. Rev. A **95**, 052505 (2017).

[98] M. R. Pederson and T. Baruah, Adv. At. Mol. Opt. Phys., 153 (2015).

[99] R. Shinde, S. S. Yamijala, and B. M. Wong, J. Phys.: Condens. Matter **33**, 115501 (2020).

[100] S. Goedecker and C. J. Umrigar, Phys. Rev. A **55**, 1765 (1997).

[101] O. A. Vydrov and G. E. Scuseria, J. Chem. Phys. **121**, 8187 (2004).

[102] L. A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople, J. Chem. Phys. **94**, 7221 (1991).

[103] L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, J. Chem. Phys. **106**, 1063 (1997).

[104] B. J. Lynch and D. G. Truhlar, J. Phys. Chem. A **107**, 8996 (2003).

[105] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, and J. Sun, J. Phys. Chem. Lett. **11**, 8208 (2020).

[106] Y. Yao and Y. Kanai, J. Chem. Phys. **146**, 224105 (2017).

[107] N. Mardirossian and M. Head-Gordon, Mol. Phys. **115**, 2315 (2017).

[108] J. W. Furness and J. Sun, Phys. Rev. B **99**, 041119 (2019).

[109] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, and J. Sun, J. Chem. Phys. **156**, 034109 (2022).

[110] I. Y. Zhang, X. Ren, P. Rinke, V. Blum, and M. Scheffler, New J. Phys. **15**, 123033 (2013).

[111] H. Shang, C. Carbogno, P. Rinke, and M. Scheffler, Comput. Phys. Commun. **215**, 26 (2017).

[112] I. Y. Zhang, A. J. Logsdail, X. Ren, S. V. Levchenko, L. Ghiringhelli, and M. Scheffler, New J. Phys. **21**, 013025 (2019).

[113] A Marek, V Blum, R Johanni, V Havu, B Lang, T Auckenthaler, A Heinecke, H.-J. Bungartz, and H Lederer, J. Phys. Condens. Matter **26**, 213201 (2014).

[114] T. Auckenthaler, V. Blum, H.-J. Bungartz, T. Huckle, R. Johanni, L. Krämer, B. Lang, H. Lederer, and P. Willems, Parallel Comput. **37**, 783 (2011).

[115] P. Kůs, A. Marek, S. Köcher, H.-H. Kowalski, C. Carbogno, C. Scheurer, K. Reuter, M. Scheffler, and H. Lederer, Parallel Comput. **85**, 167 (2019).

[116] R. Stratmann, G. E. Scuseria, and M. J. Frisch, Chem. Phys. Lett. **257**, 213 (1996).

[117] V. Havu, V. Blum, P. Havu, and M. Scheffler, J. Comput. Phys. **228**, 8367 (2009).

[118] F. Knuth, C. Carbogno, V. Atalla, V. Blum, and M. Scheffler, Comput. Phys. Commun. **190**, 33 (2015).

[119] D. Mejia-Rodriguez and S. B. Trickey, Phys. Rev. A **96**, 052512 (2017).

[120] P. Pulay, Chem. Phys. Lett. **73**, 393 (1980).

[121] G. P. Kerker, Phys. Rev. B **23**, 3082 (1981).

[122] D. Mejía-Rodríguez and S. B. Trickey, Phys. Rev. B **102**, 121109 (2020).

[123] Y. Zhao, B. J. Lynch, and D. G. Truhlar, J. Phys. Chem. A **108**, 2715 (2004).

[124] Y. Zhao, N. Gonz lez-Garc a, and D. G. Truhlar, J. Phys. Chem. A **109**, 2012 (2005).

[125] M. Shishkin and G. Kresse, Phys. Rev. B **75**, 235102 (2007).

[126] M. R. Pederson and K. A. Jackson, Phys. Rev. B **41**, 7453 (1990).

[127] K. Jackson and M. R. Pederson, Phys. Rev. B **42**, 3276 (1990).

[128] D. Porezag and M. R. Pederson, Phys. Rev. A **60**, 2840 (1999).

[129] J. A. Pople, M. Head-Gordon, and K. Raghavachari, J. Chem. Phys. **87**, 5968 (1987).

[130] J. M. L. Martin and G. de Oliveira, J. Chem. Phys. **111**, 1843 (1999).

[131] M. J. Frisch and et. al., *Gaussian 03, Revision C.02*, Gaussian, Inc., Wallingford, CT, 2004.

[132] F. Karsai, M. Engel, E. Flage-Larsen, and G. Kresse, New J. Phys. **20**, 123008 (2018).

[133] M. Zacharias, M. Scheffler, and C. Carbogno, Phys. Rev. B **102** (2020).

[134] F. Brown-Altvater, G. Antonius, T. Rangel, M. Giantomassi, C. Draxl, X. Gonze, S. G. Louie, and J. B. Neaton, Phys. Rev. B **101**, 165102 (2020).

[135] G. Kresse and J. Hafner, Phys. Rev. B **49**, 14251 (1994).

[136] G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993).

[137] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).

[138] G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996).

[139] J. Pipek and P. G. Mezey, J. Chem. Phys. **90**, 4916 (1989).

[140] W. von Niessen, J. Chem. Phys. **56**, 4290 (1972).

[141] I.-M. Høyvik, B. Jansik, and P. Jørgensen, J. Chem. Phys. **137**, 224114 (2012).

[142] T. Hahn, S. Liebing, J. Kortus, and M. R. Pederson, J. Chem. Phys. **143**, 224104 (2015).

[143] W. L. Luken and D. N. Beratan, Theor. Chim. Acta **61**, 265 (1982).

[144] P.-O. Löwdin, Adv. Quantum Chem. **5**, 185 (1970).

[145] T. Hahn, S. Schwalbe, J. Kortus, and M. R. Pederson, J. Chem. Theory Comput. **13**, 5823 (2017).

[146] M. R. Pederson, T. Baruah, D. you Kao, and L. Basurto, J. Chem. Phys. **144**, 164117 (2016).

[147] Z. Szotek, W. M. Temmerman, and H. Winter, Phys. Rev. B **47**, 4029 (1993).

[148] M. Arai and T. Fujiwara, Phys. Rev. B **51**, 1477 (1995).

[149] C. D. Pemmaraju, T. Archer, D. Sánchez-Portal, and S. Sanvito, Phys. Rev. B **75**, 045101 (2007).

[150] N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt, Rev. Mod. Phys. **84**, 1419 (2012).

[151] L. Pauling, J. Am. Pharm. Assoc. **30**, 30 (1941).

[152] L. D. Brown, D. A. Kleier, and W. N. Lipscomb, J. Am. Chem. Soc. **99**, 6787 (1977).

[153] T. Miyake and F. Aryasetiawan, Phys. Rev. B **77**, 085122 (2008).

[154] B. Baumeier, P. Krüger, and J. Pollmann, Phys. Rev. B **73**, 195205 (2006).

[155] A. F. Schuch, E. R. Grilly, and R. L. Mills, Phys. Rev. **110**, 775 (1958).

[156] M. J. van Setten, V. A. Popa, G. A. de Wijs, and G. Brocks, Phys. Rev. B **75**, 035204 (2007).

[157] J. Lehtola, M. Hakala, A. Sakko, and K. Hämäläinen, J. Comput. Chem. **33**, 1572 (2012).

[158] S. Schwalbe, L. Fiedler, J. Kraus, J. Kortus, K. Trepte, and S. Lehtola, J. Chem. Phys. **153**, 084104 (2020).

[159] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).

[160] V. P. Smirnov, R. A. Evarestov, and D. E. Usvyat, Int. J. Quantum Chem. **88**, 642 (2002).

[161] K. Lejaeghere and et al., Science **351**, aad3000 (2016).

[162] M. Råsander and M. A. Moram, J. Chem. Phys. **143**, 144104 (2015).

[163] C. Pisani, L. Maschio, S. Casassa, M. Halo, M. Sch tz, and D. Usvyat, J. Comput. Chem. **29**, 2113 (2008).

[164] A. Gross and M. Scheffler, Jour. Vac. Sci. Techn. **15**, 1624 (1997).

[165] Y. Wang, Y. Li, J. Chen, I. Y. Zhang, and X. Xu, JACS Au **1**, 543 (2021).

[166] P. Haas, F. Tran, and P. Blaha, Phys. Rev. B **79**, 085104 (2009).

[167] T. B. Shiell, L. Zhu, B. A. Cook, J. E. Bradby, D. G. McCulloch, and T. A. Strobel, Phys. Rev. Lett. **126**, 215701 (2021).

[168] F. P. Bundy and J. S. Kasper, J. Chem. Phys. **46**, 3437 (1967).

[169] J. A. Van Vechten, Phys. Rev. B **7**, 1479 (1973).

[170] M. T. Yin and M. L. Cohen, Phys. Rev. B **26**, 5668 (1982).

[171] R. G. Hennig, A. Wadehra, K. P. Driver, W. D. Parker, C. J. Umrigar, and J. W. Wilkins, Phys. Rev. B **82**, 014101 (2010).

[172] C. Raffy, J. Furthmüller, and F. Bechstedt, Phys. Rev. B **66**, 075201 (2002).

[173] U Monteverde, M. A. Migliorato, J Pal, and D Powell, J. Phys. Condens. Matter **25**, 425801 (2013).

[174] S. Pandolfi, C. Renero-Lecuna, Y. Le Godec, B. Baptiste, N. Menguy, M. Lazzeri, C. Gervais, K. Spektor, W. A. Crichton, and O. O. Kurakevych, Nano Lett. **18**, PMID: 30102550, 5989 (2018).

[175] N. Bao, F. Guo, D. Kang, Y. Feng, H. Wang, and J. Dai, J. Appl. Phys. **129**, 145701 (2021).

[176] M. J. van Setten, F. Caruso, S. Sharifzadeh, X. Ren, M. Scheffler, F. Liu, J. Lischner, L. Lin, J. R. Deslippe, S. G. Louie, C. Yang, F. Weigend, J. B. Neaton, F. Evers, and P. Rinke, J. Chem. Theory Comput. **11**, 5665 (2015).

[177] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. Lett. **56**, 2415 (1986).

[178] J. H. de Boer and E. J. W. Verwey, Proc. Phys. Soc. **49**, 59 (1937).

[179] J. Hubbard and B. H. Flowers, Proc. R. Soc. A: Math. Phys. Eng. Sci. **276**, 238 (1963).

[180] B. Brandow, Adv. Phys. **26**, 651 (1977).

[181] S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys, and A. P. Sutton, Phys. Rev. B **57**, 1505 (1998).

[182] H. Jiang, R. I. Gomez-Abal, P. Rinke, and M. Scheffler, Phys. Rev. B **82**, 045108 (2010).

[183] M. Forti, P. Alonso, P. Gargano, and G. Rubiolo, Procedia Mater. Sci. **1**, 230 (2012).

[184] Y.-C. Wang, Z.-H. Chen, and H. Jiang, J. Chem. Phys. **144**, 144106 (2016).

[185] A. Floris, S. de Gironcoli, E. K. U. Gross, and M. Cococcioni, Phys. Rev. B **84**, 161102 (2011).

[186] H. Jiang, Phys. Rev. B **97**, 245132 (2018).

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

[188] B. Santra and J. P. Perdew, J. Chem. Phys. **150**, 174106 (2019).

[189] Y. Yamamoto, S. Romero, T. Baruah, and R. R. Zope, J. Chem. Phys. **152**, 174112 (2020).

[190] F. W. Aquino, R. Shinde, and B. M. Wong, J. Comput. Chem. **41**, 1200 (2020).

[191] I. Y. Zhang and X. Xu, WIREs Comput. Mol. Sci. **11**, e1490 (2021).

[192] G. E. Scuseria, The Journal of Chemical Physics **97**, 7528 (1992).

[193] S. Andersson and M. Grüning, J. Phys. Chem. A **108**, 7621 (2004).

[194] S. Patchkovskii and T. Ziegler, J. Chem. Phys. **116**, 7806 (2002).

[195] L. Goerigk and S. Grimme, J. Chem. Theory Comput. **6**, 107 (2010).

[196] P. Verma, A. Perera, and R. J. Bartlett, Chem. Phys. Lett. **524**, 10 (2012).

[197] B. J. Lynch and D. G. Truhlar, J. Phys. Chem. A **105**, 2936 (2001).

[198] M. Atala, M. Aidelsburger, J. T. Barreiro, D. Abanin, T. Kitagawa, E. Demler, and I. Bloch, Nat. Phys. **9**, 795 (2013).

[199] Y. Ooshika, J. Phys. Soc. Jpn. **12**, 1238 (1957).

[200] Y. Ooshika, J. Phys. Soc. Jpn. **12**, 1246 (1957).

[201] H. Kahlert, O. Leitner, and G. Leising, Synth. Met. **17**, 467 (1987).

[202] W. P. Su, J. R. Schrieffer, and A. J. Heeger, Phys. Rev. Lett. **42**, 1698 (1979).

[203] L. Wang, M. Troyer, and X. Dai, Phys. Rev. Lett. **111**, 026802 (2013).

[204] S. Hirata, H. Torii, and M. Tasumi, Phys. Rev. B **57**, 11994 (1998).

[205] P. Vogl and D. K. Campbell, Phys. Rev. B **41**, 12797 (1990).

[206] T. Körzdörfer, R. M. Parrish, J. S. Sears, C. D. Sherrill, and J.-L. Brédas, J. Chem. Phys. **137**, 124305 (2012).

[207] C. R. Fincher, C. E. Chen, A. J. Heeger, A. G. MacDiarmid, and J. B. Hastings, Phys. Rev. Lett. **48**, 100 (1982).

[208] X. Ren, P. Rinke, C. Joas, and M. Scheffler, J. Mater. Sci. **47**, 7447 (2012).

[209] J. P. Perdew, M. Ernzerhof, and K. Burke, J. Chem. Phys. **105**, 9982 (1996).

[210] F. Knuth, *Strain and stress: derivation, implementation, and application to organic crystals*, 2015.

7 Appendix

7.1 Exact Constraints for meta-GGAs

As the SCAN potential is sensitive to the kinetic energy density, updated SCAN functionals are proposed frequently. There are two non-empirical versions of the variant SCAN (named rSCAN and r^{2}SCAN) widely accepted. The rSCAN functional is a regularized form of the SCAN functional, in which the dimensionless kinetic energy variable (defined in Eq. (2.39)) is optimized, and the SCAN's numerical stability is therefore improved [19]. The r^{2}SCAN is proposed to restore exact constraint adherence to rSCAN, maintain rSCAN's numerical stability and the accuracy of SCAN [105]. Table 7.1 summarizes the cross-comparison for the satisfied constraints of SCAN, rSCAN, and r^{2}SCAN.

Constraint	SCAN	rSCAN	r^{2}SCAN
uniform density limit	√	×	√
uniform density scaling	√	×	√
GE2	√	×	√
GE4x	√	×	×

Table 7.1: Summary of the adherence of the known exact constraints in rSCAN and r2SCAN, compared with SCAN.

A brief summary of the 17 constraints valid in SCAN is listed as follows:

- for the exchange term,
  1. negativity
  2. spin-scaling
  3. uniform density scaling
  4. (GE4x) fourth-order gradient expansion
  5. non-uniform density scaling
  6. tight bound for two-electron density
- for the correction term,
  7. non-positivity
  8. (GE2) second-order gradient expansion
  9. uniform density scaling to the high-density limit

10 uniform density scaling to the low-density limit

11 zero correlation energy for any one-electron spin-polarized density

12 non-uniform density scaling

- for both the exchange and correction terms,

13 size extensively

14 Lieb-Oxford bound

15 Lieb-Oxford bound for two-electron densities

16 weak dependence upon relative spin polarization in the low-density limit

17 static linear response of the uniform electron gas

7.2 Derivation of SIC Potential

The SIC potential $\hat{v}^{\mathrm{SIC}}$ is defined by the SIC energy (E^{SIC} defined in Eq. (4.1)) variation with respect to the density

$$\begin{aligned}\hat{v}^{\mathrm{SIC}}(r) &= \frac{\delta E^{\mathrm{SIC}}[\{n_i\}]}{\delta n(r)} \\ &= \sum_i^{N_\mathrm{e}} \int d^3r' \frac{\delta E^{\mathrm{SIC}}[\{n_i\}]}{\delta n_i(r')} \frac{\delta n_i(r')}{\delta n(r)} \\ &= \sum_i^{N_\mathrm{e}} \int d^3r' \hat{v}_i^{1e\mathrm{SIC}} \frac{\delta n_i(r')}{\delta n(r)} \quad .\end{aligned} \tag{7.1}$$

Here the one-electron SIC potential $\hat{v}_i^{1e\mathrm{SIC}}$ is the sum of the single-particle Hartree potential $\hat{v}_{\mathrm{es}}[n_i]$ and the single-particle XC potential $\hat{v}_{\mathrm{xc}}[n_i]$. In the generalized Kohn-Sham framework [52], the SIC potential is evaluated by acting on the SIC orbitals

$$\hat{v}^{\mathrm{SIC}}(r)\phi_i(r) = \hat{v}^{\mathrm{SIC}}(r) \frac{\delta n(r)}{\delta \phi_i^*(r')} = \hat{v}_i^{1e\mathrm{SIC}} \phi_i(r) \quad . \tag{7.2}$$

The SIC potential operator in Eq. (4.16) is written as

$$\hat{v}^{\mathrm{SIC}} = \sum_i^{N_\mathrm{e}} \hat{v}_i^{1e\mathrm{SIC}} |\phi_i\rangle\langle\phi_i| \quad , \tag{7.3}$$

which gives $\hat{v}^{\mathrm{SIC}}\phi_i(r) = \sum_j^{N_\mathrm{e}} \hat{v}_j^{1e\mathrm{SIC}} |\phi_j\rangle\langle\phi_j|\phi_i\rangle = \hat{v}_i^{1e\mathrm{SIC}}\phi_i(r)$.

7.3 Derivation of Constraint 2

By fixing the total density n, the Euler equation Eq. (4.5) is simplified to

$$\delta\left[E^{\text{SIC-DFA}}-\sum_{ab}^{N_\text{e}}\lambda_{ab}\left(\langle\phi_a|\phi_b\rangle-\delta_{ab}\right)\right]=0. \tag{7.4}$$

Given the total density ($\delta n = 0$), a set of one-electron equations for the SIC potential can be obtained by solving the Euler equation:

$$\begin{aligned}
&\hat{h}^{\text{DFA}}(\sum_i \delta n_i)+\sum_i \hat{v}_i^{1e\text{SIC}}\delta n_i \\
&-\sum_{ab}\lambda_{ab}(\delta\phi_a^*\cdot\phi_b+\phi_a^*\cdot\delta\phi_b)=0 \\
\Leftrightarrow&\sum_i \hat{v}_i^{1e\text{SIC}}(\delta\phi_i^*\cdot\phi_i+\phi_i^*\cdot\delta\phi_i) \\
&-\sum_{ai}\lambda_{ai}(\delta\phi_a^*\cdot\phi_i+\phi_a^*\cdot\delta\phi_i)=0 \\
\Leftrightarrow&\{\sum_i \delta\phi_i^*(v_i^{1e\text{SIC}}\phi_i-\sum_a\lambda_{ai}\phi_a)\}+\text{H.C.}=0 \\
\Leftrightarrow&\hat{v}_i^{1e\text{SIC}}\phi_i=\sum_a\lambda_{ai}\phi_a \quad .
\end{aligned} \tag{7.5}$$

With the Hermitian property of the Lagrange multiplier $\lambda_{ab}=\lambda_{ba}^*$, Constraint 2 shown in Eq. (4.12) is obtained by

$$\begin{aligned}
&\langle\phi_m|\hat{v}_m^{1e\text{SIC}}-\hat{v}_n^{1e\text{SIC}}|\phi_n\rangle \\
&=\langle\phi_m|\hat{v}_m^{1e\text{SIC}}|\phi_n\rangle-\langle\phi_m|\hat{v}_n^{1e\text{SIC}}|\phi_n\rangle \\
&=\sum_a\lambda_{na}^*\langle\phi_a|\phi_n\rangle-\sum_a\lambda_{an}\langle\phi_m|\phi_a\rangle \\
&=\lambda_{nm}^*-\lambda_{mn} \\
&=0 \quad .
\end{aligned} \tag{7.6}$$

Furthermore, the Hermitian property of the SIC-DFA KS non-interacting Hamiltonian can be proven by

$$\begin{aligned}
&\langle\phi_m|\hat{h}^{\text{SIC-DFA}\dagger} - \hat{h}^{\text{SIC-DFA}}|\phi_n\rangle \\
&=\langle\phi_m|\hat{h}^{\text{SIC-DFA}\dagger}|\phi_n\rangle - \langle\phi_m|\hat{h}^{\text{SIC-DFA}}|\phi_n\rangle \\
&=\langle\phi_m|(\hat{t}_\text{s} + \hat{v}_\text{ext} + \hat{v}_\text{es} + \hat{v}_\text{xc} + \hat{v}^\text{SIC})^\dagger|\phi_n\rangle \\
&\quad - \langle\phi_m|\hat{t}_\text{s} + \hat{v}_\text{ext} + \hat{v}_\text{es} + \hat{v}_\text{xc} + \hat{v}^\text{SIC}|\phi_n\rangle \\
&=\langle\phi_m|\hat{t}_\text{s} + \hat{v}_\text{ext} + \hat{v}_\text{es} + \hat{v}_\text{xc} + \hat{v}^{\text{SIC}\dagger}|\phi_n\rangle \\
&\quad - \langle\phi_m|\hat{t}_\text{s} + \hat{v}_\text{ext} + \hat{v}_\text{es} + \hat{v}_\text{xc} + \hat{v}^\text{SIC}|\phi_n\rangle \\
&=\langle\phi_m|\hat{v}^{\text{SIC}\dagger}|\phi_n\rangle - \langle\phi_m|\hat{v}^\text{SIC}|\phi_n\rangle \\
&=\sum_j^{N_\text{e}}\langle\phi_m|\phi_j\rangle\langle\phi_j|\hat{v}_j^{1e\text{SIC}}|\phi_n\rangle - \langle\phi_m|\sum_i^{N_\text{e}}\hat{v}_i^{1e\text{SIC}}|\phi_i\rangle\langle\phi_i|\phi_n\rangle \\
&=\langle\phi_m|\sum_j^{N_\text{e}}\delta_{jm}\hat{v}_j^{1e\text{SIC}}|\phi_n\rangle - \langle\phi_m|\sum_i^{N_\text{e}}\hat{v}_i^{1e\text{SIC}}\delta_{in}|\phi_n\rangle \\
&=\langle\phi_m|\hat{v}_m^{1e\text{SIC}}|\phi_n\rangle - \langle\phi_m|\hat{v}_n^{1e\text{SIC}}|\phi_n\rangle \\
&=0 \quad .
\end{aligned} \tag{7.7}$$

7.4 Derivation of Edmiston and Ruedenberg (E-R) Restriction

The Dirac notation for the integral of the squared orbital density is

$$\int n_i^2(\mathbf{r})d^3\mathbf{r} = \langle\phi_i|n_i|\phi_i\rangle. \tag{7.8}$$

The request for maximizing the integral of all squared orbital densities (mentioned in Eq. (4.21)) is written as

$$\delta\int n_i^2(\mathbf{r}) = 0 \quad . \tag{7.9}$$

It is indicated that the maximum of total inner density overlaps implies the minimization of the overlaps between two different densities $\sum_{m\neq n}\langle\phi_m|n_n|\phi_m\rangle$ [140]. Subjecting to the orthonormalization condition for the SIC occupied orbitals

$$\langle\phi_a|\phi_b\rangle - \delta_{ab} = 0 \quad , \tag{7.10}$$

the maximization results in the Euler equation:

$$\delta\left[\sum_i\langle\phi_i|n_i|\phi_i\rangle - \sum_{ab}\eta_{ab}\left(\langle\phi_a|\phi_b\rangle - \delta_{ab}\right)\right] = 0. \tag{7.11}$$

This Euler equation gives a set of one-electron equations for the E-R orbitals:

$$n_i\phi_i = \sum_a\frac{1}{2}\eta_{ai}\phi_a \quad . \tag{7.12}$$

These one-electron equations produce the orbital density constraint

$$\begin{aligned}
&\langle\phi_m|n_m - n_n|\phi_n\rangle \\
&= \langle\phi_m|n_m|\phi_n\rangle - \langle\phi_m|n_n|\phi_n\rangle \\
&= \sum_b \frac{1}{2}\eta^*_{bm}\langle\phi_b|\phi_n\rangle - \sum_a \frac{1}{2}\eta_{an}\langle\phi_m|\phi_a\rangle \\
&= \sum_b \frac{1}{2}\eta^*_{bm}\delta_{bn} - \sum_a \frac{1}{2}\eta_{an}\delta_{ma} \\
&= \frac{1}{2}(\eta^*_{nm} - \eta_{mn}) = 0 \quad .
\end{aligned} \tag{7.13}$$

Moreover, Constraint 1 (Eq. (4.6)) ensures that summing the inner density overlaps is invariant under a unitary transformation U_{ma} of the SIC orbitals $(\phi'_m = \sum_a U_{ma}\phi_a)$:

$$\begin{aligned}
&\int d^3\mathbf{r} n^2(\mathbf{r}) = \int d^3\mathbf{r}(\sum_n n'_n)(\sum_m n'_m) \\
&= \sum_{mn}\langle\phi'_m|n'_n|\phi'_m\rangle \\
&= \sum_{mn}\sum_a U^*_{ma}\sum_b U^*_{bn}\sum_c U_{nc}\sum_d U_{md}\langle\phi_a|\phi^\dagger_b\phi_c|\phi_d\rangle \\
&= \sum_{ad}\sum_m\sum_a U^*_{ma}U_{md}\sum_{bc}\sum_n U^*_{bn}\sum_c U_{nc}\langle\phi_a|\phi^\dagger_b\phi_c|\phi_d\rangle \\
&= \sum_{ad}\delta_{ad}\sum_{bc}\delta_{bc}\langle\phi_a|\phi^\dagger_b\phi_c|\phi_d\rangle \\
&= \sum_{ab}\langle\phi_a|\phi^\dagger_b\phi_b|\phi_a\rangle \\
&= \sum_{ab}\langle\phi_a|n_\mathrm{b}|\phi_a\rangle \quad .
\end{aligned} \tag{7.14}$$

7.5 Derivation of Constraint 3

Constraint 3 comes from maximizing the overlap matrix of the SIC orbital densities under the potential orbital constraint:

$$\max\sum_i \left(\int d^3\mathbf{r} n_i^2(\mathbf{r})|_{\hat{v}_i^{1e\mathrm{SIC}\dagger}=\hat{v}_i^{1e\mathrm{SIC}}}\right) \quad . \tag{7.15}$$

Here we mark $v_i^{1e\mathrm{SIC}}$ as v_i for simplification.

Using the one-electron equations of SIC potential in Eq. (7.5), the self-consistently localized Lagrangian with the multipliers η_{mn} is constructed as

$$L = \sum_i \int d^3\mathbf{r} n_i^2(\mathbf{r}) - \sum_{mn}\eta_{mn}\left\{\langle\phi_m|\hat{v}_i^{1e\mathrm{SIC}\dagger}|\phi_n\rangle - \sum_j \lambda_{jn}\langle\phi_m|\phi_j\rangle\right\} \quad . \tag{7.16}$$

The variation is

$$\begin{aligned}
\delta L =& \sum_i 2n_i\delta\phi_i^*\phi_i + 2n_i\phi_i^*\delta\phi_i \\
&- \sum_{mn}\eta_{mn}(\delta\phi_m^* v_m\phi_n + \phi_m^*\delta v_m\phi_n + \phi_m^*\ v_m\delta\phi_n) \\
&- \sum_j \lambda_{jn}(\delta\phi_m^*\phi_j + \phi_m^*\delta\phi_j)) \\
=& \sum_i 2n_i\delta\phi_i^*\phi_i \\
&- \sum_{mn}\eta_{mn}(\delta\phi_m^* v_m\phi_n + \delta\phi_m^*\phi_m^*\frac{\delta v_m}{\delta n_m}\phi_n\phi_m \\
&- \sum_j \lambda_{jn}\delta\phi_m^*\phi_j) \\
&+ \sum_i 2n_i\phi_i^*\delta\phi_i \\
&- \sum_{mn}\eta_{mn}(\phi_m^* v_m\delta\phi_n + \phi_m^*\phi_m^*\frac{\delta v_m}{\delta n_m}\phi_n\delta\phi_m \\
&- \sum_j \lambda_{jn}\phi_m^*\delta\phi_j) \quad .
\end{aligned} \tag{7.17}$$

The zero of the variation $\delta L = 0$ leads to

$$\begin{aligned}
2n_i\phi_i &= \sum_n \eta_{in} v_i\phi_n \\
&+ \sum_n \eta_{in}\phi_i^*\frac{\delta v_i}{\delta n_i}\phi_n\phi_i - \sum_n\sum_a \eta_{ia}\lambda_{na}\phi_n \\
\leftrightarrow 2n_i\phi_i &= (v_i + n_i\frac{\delta v_i}{\delta n_i})\sum_n \eta_{in}\phi_n - \sum_n\sum_a \eta_{ia}\lambda_{na}\phi_n \\
\leftrightarrow 2n_i\phi_i &= \frac{\delta(v_i n_i)}{\delta n_i}\sum_n \eta_{in}\phi_n - \sum_n\sum_a \eta_{ia}\lambda_{na}\phi_n \quad .
\end{aligned} \tag{7.18}$$

Considering $\frac{\delta(v_i n_i)}{\delta n_i} \approx (\lambda_{ii} + v_i)$, a further simplification is obtained :

$$\langle j|2n_i|i\rangle\lambda_{ii} = \lambda_{ii}\eta_{ij} - \sum_a \eta_{ia}\lambda_{ja} \quad , \tag{7.19}$$

where $|i\rangle$ is a mark of $|\phi_i\rangle$. The matrix elements η_{ij} can be evaluated by recursively solving this equation. A matrix $\bar{\bar{D}}_{ji} = \langle j|2n_i|i\rangle$ and $\tilde{\lambda}_{ij} = \lambda_{ii}\delta_{ij}$ are

introduced to illustrate this recursive progress.

$$
\begin{aligned}
&\langle j|2n_i|i\rangle\lambda_{ii} = \lambda_{ii}\eta_{ij} - \sum_a \eta_{ia}\lambda_{ja} \\
\Leftrightarrow &\bar{\bar{D}}\tilde{\lambda} = \eta^T\tilde{\lambda} - \lambda\eta^T \\
\Leftrightarrow &\eta^T = \bar{\bar{D}} + \lambda\eta^T\tilde{\lambda}^{-1} \\
\Leftrightarrow &\eta^T = \bar{\bar{D}} + \lambda(\bar{\bar{D}} + \lambda\eta^T\tilde{\lambda}^{-1})\tilde{\lambda}^{-1} \\
&= \bar{\bar{D}} + \lambda\bar{\bar{D}}\tilde{\lambda}^{-1} + \lambda^2\eta^T\tilde{\lambda}^{-2} \\
&= \lim_{x\to\infty}\sum_{a=0}^{x-1}\lambda^a\bar{\bar{D}}\tilde{\lambda}^{-a} \\
&= \lim_{x\to\infty}(\sum_{a=0}^{x-1}\lambda^a\bar{\bar{D}}\tilde{\lambda}^{-a} + \lambda^x\eta^T\tilde{\lambda}^{-x}) \\
&= \lim_{x\to\infty}(\sum_{a=0}^{x-1}\lambda^a\bar{\bar{D}}\tilde{\lambda}^{-a} + (\lambda\tilde{\lambda}^{-1})^x\eta^T) \\
\Leftrightarrow &\lim_{x\to\infty}(1-(\lambda\tilde{\lambda}^{-1})^x)\eta^T = \sum_{a=0}^{x-1}\lambda^a\bar{\bar{D}}\tilde{\lambda}^{-a} \\
\Leftrightarrow &\eta^T = \lim_{x\to\infty}\left(\tilde{\lambda}(\tilde{\lambda}-\lambda)^{-1}\right)^x\sum_{a=0}^{x-1}\lambda^a\bar{\bar{D}}\tilde{\lambda}^{-a} \quad .
\end{aligned}
\tag{7.20}
$$

If the expansion of η^T is cut off at $x = 1$, we get

$$
\eta^T = \tilde{\lambda}(\tilde{\lambda}-\lambda)^{-1}\bar{\bar{D}} \quad . \tag{7.21}
$$

The η^T is removed by taking the property $\eta^T = \eta$ into the equation above:

$$
\begin{aligned}
&(\tilde{\lambda}(\tilde{\lambda}-\lambda)^{-1}\bar{\bar{D}})^\dagger = \tilde{\lambda}(\tilde{\lambda}-\lambda)^{-1}\bar{\bar{D}} \\
\Leftrightarrow &\bar{\bar{D}}^\dagger(\tilde{\lambda}-\lambda)^{-1}\tilde{\lambda} = \tilde{\lambda}(\tilde{\lambda}-\lambda)^{-1}\bar{\bar{D}} \\
\Leftrightarrow &(\tilde{\lambda}-\lambda)\tilde{\lambda}^{-1}\bar{\bar{D}}^\dagger = \bar{\bar{D}}\tilde{\lambda}^{-1}(\tilde{\lambda}-\lambda) \quad .
\end{aligned}
\tag{7.22}
$$

Considering that $\bar{\bar{D}}$ and $\tilde{\lambda}$ are Hermitian matrices, we can further simplify this equation to get Constraint 3:

$$
\begin{aligned}
&(\tilde{\lambda}-\lambda)\tilde{\lambda}^{-1}\bar{\bar{D}}^\dagger = \bar{\bar{D}}\tilde{\lambda}^{-1}(\tilde{\lambda}-\lambda) \\
&-\lambda\tilde{\lambda}^{-1}\bar{\bar{D}}^\dagger = -\bar{\bar{D}}\tilde{\lambda}^{-1}\lambda \\
\Leftrightarrow &\lambda\bar{\bar{D}}^\dagger = \bar{\bar{D}}\lambda \\
\Leftrightarrow &\langle\phi_m|n_m\hat{v}_n|\phi_n\rangle = \langle\phi_m|\hat{v}_m n_n|\phi_n\rangle \quad .
\end{aligned}
\tag{7.23}
$$

7.6 Data : Computed and Experimental Formation Energy for G2-1 Test Set

The calculated formation energy for LDA, PBE, and SCAN with respect to the reference values according to the G2-1 theory [102] are presented in Table 7.2.

The total energy for all systems within the G2-1 set using SCAN and SIC-SCAN are given in Table 7.3 and Table 7.4.

	Expt.	LDA	PBE	SCAN
LiH	-2.51	-2.65	-2.33	-2.43
BeH	-2.16	-2.59	-2.41	-2.64
CH	-3.64	-3.95	-3.67	-3.56
CH_2(1A1)	-8.26	-9.16	-8.44	-8.57
CH_2(3B1)	-7.85	-8.61	-7.75	-7.63
CH_3	-13.34	-14.67	-13.45	-13.59
CH_4	-18.22	-20.04	-18.22	-18.24
NH	-3.62	-4.04	-3.84	-3.72
NH_2	-7.88	-8.97	-8.18	-8.06
NH_3	-12.92	-14.60	-13.10	-12.85
OH	-4.63	-5.38	-4.77	-4.72
H_2O	-10.09	-11.57	-10.16	-9.97
HF	-6.13	-6.98	-6.17	-5.94
SiH_2(1A1)	-6.59	-7.16	-6.48	-6.52
SiH_2(3B1)	-5.69	-6.28	-5.78	-6.04
SiH_3	-9.79	-10.63	-9.72	-10.05
SiH_4	-13.98	-15.02	-13.68	-14.09
PH_2	-6.63	-7.49	-6.72	-6.85
PH_3	-10.49	-11.69	-10.39	-10.55
SH_2	-7.93	-8.95	-7.91	-7.95
HCl	-4.65	-5.15	-4.73	-4.60
Li_2	-1.06	-1.02	-0.88	-0.80
LiF	-5.98	-6.73	-6.04	-5.85
C_2H_2	-17.60	-19.93	-18.00	-17.52
C_2H_4	-24.44	-27.41	-24.80	-24.49
C_2H_6	-30.89	-34.41	-31.10	-31.01
CN	-7.87	-9.47	-8.57	-7.80
HCN	-13.56	-15.62	-14.16	-13.45
CO	-11.26	-12.97	-11.66	-11.07
HCO	-12.10	-14.43	-12.80	-12.29
H_2CO	-16.22	-18.82	-16.75	-16.26
CH_3OH	-22.24	-25.43	-22.56	-22.31
N_2	-9.91	-11.55	-10.56	-9.63

N_2H_4	-19.01	-22.33	-19.66	-19.00
NO	-6.63	-8.61	-7.46	-6.62
O_2	-5.24	-7.58	-6.23	-5.55
H_2O_2	-11.67	-14.56	-12.24	-11.64
F_2	-1.70	-3.29	-2.30	-1.59
CO_2	-16.90	-20.53	-18.05	-17.13
Na_2	-0.74	-0.74	-0.78	-0.61
Si_2	-3.28	-3.87	-3.63	-3.37
P_2	-5.08	-6.17	-5.29	-4.97
S_2	-4.45	-5.80	-5.01	-4.79
Cl_2	-2.58	-3.41	-3.04	-2.53
NaCl	-4.29	-4.33	-4.18	-4.23
SiO	-8.37	-9.68	-8.59	-8.19
CS	-7.46	-8.73	-7.79	-7.30
SO	-5.46	-7.22	-6.14	-5.74
ClO	-2.84	-4.42	-3.62	-3.01
ClF	-2.71	-3.97	-3.24	-2.63
Si_2H_6	-23.04	-25.04	-22.72	-23.36
CH_3Cl	-17.15	-19.29	-17.45	-17.23
CH_3SH	-20.57	-23.23	-20.76	-20.69
HOCl	-7.17	-9.07	-7.71	-7.15
SO_2	-11.24	-14.58	-12.20	-11.39

Table 7.2: Formation energy (in eV) of the G2-1 test set for LDA, PBE, and SCAN calculated in *FHI-aims* using *tight tier*-1 NAO basis set. Geometries and experimental values are taken from Ref. [102] and references within.

	SCAN	SIC-SCAN (real)	SIC-SCAN (complex)
H	-0.50016177	-0.49944475	-0.49944475
He	-2.90493667	-2.89814732	-2.89814732
Li	-7.47828238	-7.47107607	-7.47108942
Be	-14.64835692	-14.63992632	-14.63999588
B	-24.63934851	-24.62536876	-24.63938461
C	-37.83752135	-37.80809375	-37.86395821

N	-54.5872058	-54.52967642	-54.63929257
O	-75.07124621	-74.99085186	-75.15816454
F	-99.74522364	-99.62962522	-99.88992635
Ne	-128.9493868	-128.7829113	-129.1406733
Na	-162.2743854	-162.090964	-162.5201797
Mg	-200.0709478	-199.8666666	-200.3675556
Al	-242.369569	-242.1350979	-242.7174744
Si	-289.3885255	-289.1187077	-289.791548
P	-341.292263	-340.9772652	-341.7521286
S	-398.150199	-397.7931029	-398.6711863
Cl	-460.1941605	-459.7877979	-460.7814757
Ar	-527.5856474	-527.1196201	-528.2367977

Table 7.3: Total energy (in Hartree) of atoms H-Ar calculated by SCAN and SIC-SCAN. The SIC-SCAN calculations using *real* and *complex* SIC orbitals are marked as SIC-SCAN(real) and SIC-SCAN(complex), respectively. The calculations are done in *FHI-aims* using *tight tier*-1 NAO basis set. (1 Hartree = 27.2114 eV)

	SCAN	SIC-SCAN (real)	SIC-SCAN (complex)
LiH	-219.5316868	-219.2546198	-219.2549627
BeH	-414.8379633	-414.5706462	-414.5974284
CH	-1046.75817	-1045.909905	-1047.236968
CH_2(1A1)	-1065.371266	-1064.272405	-1066.111285
CH_2(3B1)	-1064.430293	-1063.483299	-1064.769884
CH_3	-1084.005003	-1082.813714	-1084.56143
CH_4	-1102.264771	-1101.182451	-1102.486629
NH	-1502.689554	-1501.043723	-1503.96088
NH_2	-1520.636088	-1518.84928	-1522.009302
NH_3	-1539.03726	-1537.056698	-1540.530822
OH	-2061.074128	-2058.551443	-2063.761318
H_2O	-2079.929437	-2076.995705	-2082.489912
HF	-2733.689087	-2729.863524	-2737.666846
SiH_2(1A1)	-7908.204564	-7900.880739	-7918.774802
SiH_2(3B1)	-7907.718993	-7900.306401	-7918.53571

SiH_3	-7925.342156	-7918.042819	-7936.002204
SiH_4	-7942.991178	-7935.763224	-7953.594738
PH_2	-9320.869991	-9312.430263	-9332.981938
PH_3	-9338.183418	-9329.737161	-9350.119495
SH_2	-10869.11699	-10859.32287	-10883.06427
HCl	-12540.41814	-12529.22528	-12556.35565
Li_2	-407.7795692	-407.4197584	-407.420828
LiF	-2923.473664	-2919.081491	-2927.517642
C_2H_2	-2103.906418	-2101.359912	-2105.630563
C_2H_4	-2138.103632	-2135.838423	-2139.209054
C_2H_6	-2171.84303	-2169.758284	-2172.195099
CN	-2522.742495	-2518.833828	-2524.701882
HCN	-2542.003581	-2538.667859	-2544.142141
CO	-3083.395237	-3079.247284	-3086.149566
HCO	-3098.225891	-3094.010155	-3101.031578
H_2CO	-3115.804896	-3111.722517	-3118.656333
CH_3OH	-3149.071235	-3145.215175	-3151.765218
N_2	-2980.33876	-2976.375188	-2982.718947
N_2H_4	-3044.152617	-3040.393641	-3046.548561
NO	-3534.717658	-3530.023498	-3537.927283
O_2	-4091.034768	-4085.234982	-4094.990019
H_2O_2	-4124.339933	-4118.744824	-4128.540173
F_2	-5429.864016	-5422.49673	-5435.944668
CO_2	-5132.195316	-5124.256602	-5137.90579
Na_2	-8831.808887	-8821.849069	-8845.192821
Si_2	-15752.30311	-15737.34404	-15774.1802
P_2	-18578.57724	-18561.09669	-18603.66184
S_2	-21672.68087	-21652.78281	-21700.74111
Cl_2	-25046.93759	-25024.69406	-25078.505
NaCl	-16942.03315	-16925.5753	-16964.74317
SiO	-9925.396329	-9914.469476	-9938.823334
CS	-11870.83195	-11859.22838	-11885.51356
SO	-12882.42234	-12869.37471	-12898.63568
ClO	-14567.95901	-14554.04184	-14585.67475
ClF	-15238.97729	-15224.04943	-15258.06022
Si_2H_6	-15853.94547	-15839.46319	-15875.16065

CH_3Cl	-13609.85292	-13597.71554	-13625.77518
CH_3SH	-11938.65656	-11927.8963	-11952.70255
HOCl	-14585.70773	-14571.71658	-14603.63769
SO_2	-14930.81614	-14914.2202	-14949.69407

Table 7.4: Total energy (in eV) of the G2-1 test set calculated by SCAN and SIC-SCAN. The SIC-SCAN calculations using *real* and *complex* SIC orbitals are marked as SIC-SCAN(real) and SIC-SCAN(complex), respectively. The calculations are done in *FHI-aims* using *tight tier*-1 NAO basis set. Geometries are taken from Ref. [102] and references within. (1 Hartree = 27.2114 eV)

7.7 Data : Computed and Experimental HOMO Energy

The calculated HOMO energy of 24 molecules for LDA, PBE, SCAN, SIC-SCAN, and G_0W_0@PBE with respect to experimental results are presented in table 7.5 and table 7.6.

Mol.	Expt.	LDA	PBE	SCAN	G_0W_0@PBE
LiH	7.9	4.3838	4.3518	4.4935	6.54
CH_4	12.6	9.4615	9.4420	9.8147	13.93
NH_3	10.07	6.1747	6.0698	6.3946	10.32
H_2O	12.62	7.3645	7.2144	7.5973	11.97
PH_3	9.87	6.7492	6.6883	6.9429	10.27
SH_2	10.46	6.3572	6.2590	6.5766	10.03
HCl	12.74	8.1158	8.0116	8.4095	12.25
Li_2	5.11	3.1750	3.1607	3.1555	4.99
LiF	11.3	6.2893	6.0643	6.3595	9.95
C_2H_2	11.4	7.3798	7.2032	7.4072	11.02
C_2H_4	10.51	6.9618	6.7731	6.9334	10.33
C_2H_6	11.52	8.1320	8.1460	8.4901	12.37
HCN	13.6	9.1658	9.0550	9.3269	13.21
CO	14.01	9.0954	9.0126	9.5531	13.25
H_2CO	10.89	6.3338	6.2493	6.5667	10.33
CH_3OH	10.84	6.3939	6.2889	6.6369	10.33

N_2	15.58	10.372	10.223	10.7517	14.89
F_2	15.7	9.6238	9.4366	10.066	14.96
CO_2	13.78	9.3235	9.0876	9.4994	13.25
Na_2	4.89	3.1831	3.0844	3.0970	4.83
P_2	10.53	7.2115	7.0859	7.3928	10.21
Cl_2	11.48	7.5414	7.4281	7.8308	11.1
NaCl	9.2	5.2983	5.1707	5.4222	8.1
SO_2	12.35	8.2859	8.0854	8.4792	11.82

Table 7.5: Calculated negative HOMO energy ($-\epsilon_{\mathrm{HOMO}}$) (in eV) of 24 molecules against the experimental ionization potential (IP). The calculations of LDA, PBE, and SCAN are done in *FHI-aims* using *tight tier*-1 NAO basis set. Geometries are taken from the G2-1 test set [102]. The calculated *GW* results are from Ref. [176]. The experimental IPs are taken from Ref. [77].

Mol.	Expt.	SCAN	SIC-SCAN (real)	SIC-SCAN (complex)	G_0W_0@PBE
LiH	7.9	4.4935	8.7523	8.7507	6.54
CH_4	12.6	9.8147	15.637	15.066	13.93
NH_3	10.07	6.3946	12.089	12.141	10.32
H_2O	12.62	7.5973	14.375	14.511	11.97
PH_3	9.87	6.9429	11.143	11.186	10.27
SH_2	10.46	6.5766	11.025	11.083	10.03
HCl	12.74	8.4095	13.311	13.373	12.25
Li_2	5.11	3.1555	5.3052	5.3048	4.99
LiF	11.3	6.3595	13.705	13.624	9.95
C_2H_2	11.4	7.4072	12.228	11.991	11.02
C_2H_4	10.51	6.9334	11.884	11.766	10.33
C_2H_6	11.52	8.4901	14.266	13.651	12.37
HCN	13.6	9.3269	14.603	14.320	13.21
CO	14.01	9.5531	15.152	15.361	13.25
H_2CO	10.89	6.5667	13.175	12.977	10.33
CH_3OH	10.84	6.6369	13.162	12.880	10.33
N_2	15.58	10.751	17.117	16.860	14.89
F_2	15.7	10.066	18.510	17.983	14.96

CO_2	13.78	9.4994	15.372	15.793	13.25
Na_2	4.89	3.0970	4.9537	4.9602	4.83
P_2	10.53	7.3928	10.855	10.632	10.21
Cl_2	11.48	7.8308	12.834	12.579	11.1
NaCl	9.2	5.4222	10.068	10.287	8.1
SO_2	12.35	8.4792	14.404	14.519	11.82

Table 7.6: Calculated negative HOMO energy ($-\epsilon_{\mathrm{HOMO}}$) (in eV) of 24 molecules against the experimental ionization potential (IP). The calculations of SCAN and SIC-SCAN are done in *FHI-aims* using *tight tier*-1 NAO basis set. The SIC-SCAN calculations using *real* and *complex* SIC orbitals are marked as SIC-SCAN(real) and SIC-SCAN(complex), respectively. Geometries are taken from the G2-1 test set [102]. The calculated *GW* results are taken from Ref. [176]. The experimental IPs are taken from Ref. [77].

7.8 Data : Hydrogen-Transfer Reactions

We examined all hydrogen-transfer reactions included in the BH76 [123, 124, 195] test set with the SIC-SCAN method. The results are shown in Table 7.7.

Reaction		Ref.	SCAN	SIC-SCAN
$H + N_2O \longrightarrow OH + N_2$	V_f	17.7	9.3593	17.0142
	V_b	82.6	64.4961	102.246
$H + HF \longrightarrow HF + H$	V_f	42.1	29.1821	49.5751
	V_b	42.1	29.1821	49.5751
$H + HCl \longrightarrow HCl + H$	V_f	17.8	9.2535	17.6535
	V_b	17.8	9.2535	17.6535
$H + CH_3F \longrightarrow HF + CH_3$	V_f	30.5	20.2660	38.6226
	V_b	56.9	47.0115	59.7744
$H + F_2 \longrightarrow HF + F$	V_f	1.5	11.2299	-4.9326
	V_b	104.8	89.1248	102.978
$H + N_2 \longrightarrow HN_2$	V_f	14.6	4.2386	10.2319
	V_b	10.9	9.7294	17.8880
$H + CO \longrightarrow HCO$	V_f	3.2	-3.7977	-0.3135
	V_b	22.8	24.3449	27.5432

$H + C_2H_4 \longrightarrow C_2H_5$	V_f	2	-4.5371	-2.8421
	V_b	42	43.1095	46.8258
$H + HCl \longrightarrow H_2 + Cl$	V_f	6.1	-1.3678	3.82873
	V_b	8	-0.0006	6.35705
$OH + H_2 \longrightarrow H_2O + H$	V_f	5.2	-2.2772	7.36300
	V_b	21.6	11.1609	13.2525
$CH_3 + H_2 \longrightarrow CH_4 + H$	V_f	11.9	7.1925	8.34285
	V_b	15	6.8811	11.9220
$H + H_2 \longrightarrow H\,H_2$	V_f	9.7	2.3180	6.85034
	V_b	9.7	2.3180	6.85034
$F + H_2 \longrightarrow HF + H$	V_f	1.6	-7.5273	2.95872
	V_b	33.8	21.8723	19.2582
$H + PH_3 \longrightarrow H_2 + PH_2$	V_f	2.9	-3.0964	1.64202
	V_b	24.7	19.0854	22.0598
$H + OH \longrightarrow H_2 + O$	V_f	10.9	3.0086	10.8601
	V_b	13.2	1.7343	15.0209
$H + H_2S \longrightarrow H_2 + HS$	V_f	3.9	-2.6232	2.13884
	V_b	17.2	10.9347	15.5090

Table 7.7: Comparison of the barrier heights (in kcal/mol) of the hydrogen-transfer reactions in the BH76 test set [123, 124, 195]. V_f is the forward reaction barrier height, and V_b is the backward reaction barrier height. The calculations of SCAN and SIC-SCAN are done in *FHI-aims* using *tight tier*-1 NAO basis set. The *complex* SIC orbitals are used in the SIC-SCAN calculations. The geometries and Weizmann-1 (W1) results are taken from Ref. [123]. (1 kcal/mol = 0.0434 eV)

Declaration

I declare that I have completed the thesis independently using only the aids and tools specified. I have not applied for a doctor's degree in the doctoral subject elsewhere and do not hold a corresponding doctor's degree. I have taken due note of the Faculty of Mathematics and Natural Sciences PhD Regulations, published in the Official Gazette of Humboldt-Universität zu Berlin no. 42/2018 on 11/07/2018.

Berlin (Germany), September 7, 2022,
Sheng Bi