University
of Bremen

BCCMS
Bremen Center for Computational Materials Science

Method development for the DFT calculation of charge-assisted surface reactions in a periodic model

Dissertation to obtain the academic degree of doctor of
natural science (Dr. rer. nat.)

by

Meisam Farzalipour Tabriz

Department 1

Physics and Electrical Engineering

University of Bremen

May 2023

Supervisor: Prof. Dr. Peter Deák

Second examiner: Prof. Dr. Michael Sentef

Date of defense: 03.06.2024

Abstract

Many important processes of great technological importance can be modeled by properly designed charged systems. A common example is photocatalysis, where photon absorption promotes electrons from the valence band of a semiconductor photocatalyst to its conduction band, generating electrons and holes. These electron-hole pairs are usually involved in separate reactions, in a macroscopic distance from each other. To avoid the quick recombination of the carriers in models of limited size, it is best to simulate the behavior of the hole and the electron in separate, charged systems.

Solid surfaces are most commonly investigated using slab models, which are created by applying periodic boundary conditions to a layer of gas or vacuum phase over a bulk-like region, resulting in a solid surface model that is periodic in two dimensions and is repeated artificially in the third. For studying charged models under periodic boundary conditions, a compensating background charge is also required to avoid the divergence of the Coulomb energy. The slab models are very convenient from a computational point of view. However, the interactions between the periodic images of the localized charge and between the localized charge and its neutralizing background can cause significant errors in the total energy, which somehow need to be accounted for.

Komsa and Pasquarello have proposed a correction scheme that is suitable for bulk and slab models. In this method, the error in the total energy is estimated by modeling the localized extra charge with a Gaussian function and comparing its energy in the periodic and isolated cases. I implemented this correction scheme in a user-friendly and robust standalone code called SLABCC. I have also extended the method to handle mediums with anisotropic dielectric tensors as well as cases where the extra charge is localized at multiple sites. SLABCC is automated for simple cases while being flexible enough for advanced users to handle non-trivial ones. In slab models with a large vacuum between the layers, *a posteriori* charge correction methods may not be adequate, and self-consistent correction may be needed to eliminate the spurious effects. I introduce a self-

consistent potential correction method, co-developed by myself, that is capable of dealing with those cases.

Density functional theory is widely used for investigating the properties of defects in bulk and on the surface of solids. Local or semi-local approximations in the density functional lead to an underestimation of band-gap, which may cause incorrect occupation of defect states and incorrect formation energies. The methods also underestimate the localization of defect states, missing the formation of small polarons. These problems can be avoided by using higher-level approximations (GW or a correctly chosen hybrid functional), but the computational cost of GW methods prohibits their use in calculating surface defects in a periodic slab model, and even hybrid functionals make the study of surface chemical reactions, with a large number of variables, impractical.

Lany and Zunger have suggested a convenient (low-cost) solution for solving the band-gap and charge delocalization problems by applying a correction scheme to the standard local or semi-local approximations. Most importantly, the linearity of the total energy as a function of the fractional occupation numbers is restored, leading to the fulfillment of the generalized Koopmans' theorem. This method works well in the bulk, but it is not accurate on the surface due to the different screening environment. I show that, by making the atom- and angular-momentum-dependent parameters of the Lany-Zunger polaron correction also coordination-dependent, it is possible to correctly describe charge trapping in small polaron states on the anatase (101) and rutile (110) surfaces at a low computational cost. Using this technique, I have found a two-dimensionally localized state in anatase, which has led to a study on estimating the real size of polarons in anatase-TiO_2.

Finally, I use the developed framework to investigate the photocatalytic CO oxidation on the anatase (101) surface. For the restoration of the pristine surface, I propose a mechanism to eliminate the surface oxygen vacancies by including electron-scavenging oxygen molecules in the gas phase. With the proposed mechanism, it is possible to achieve a complete catalytic cycle for the oxidation of CO over the anatase (101) surface.

To my dear friend, Péter

Acknowledgements

This work would not have been possible without help from many others whom I will forever be indebted to.

The completion of my dissertation would have not been possible without the constant guidance and relentless support of Prof. Dr. Peter Deák who, in addition to scientific thinking, taught me indispensable lessons that will guide me through life.

I would like to express my deepest gratitude to Prof. Dr. Thomas Frauenheim who supported me all throughout my studies and helped me overcome numerous obstacles.

I am also grateful to Dr. Bálint Aradi for teaching me many valuable lessons on scientific software development including many key details on best practices. I must also thank Dr. Christoph Koehler whose diligence and technical knowledge plays a crucial role in success of each Bremen Center for Computational Materials Science (BCCMS) member and has been a constant source of inspiration for me.

I gratefully acknowledge Prof. Dr. Georg Kresse, Prof. Dr. Cesare Franchini, Prof. Dr. Ersin Yurtsever, Dr. Stephan Lany, Dr. Michael Lorke, Dr. Hannu-Pekka Komsa, Dr. Jolla Kullgren, Dr. Maurício Chagas da Silva, and Dr. Suzanne Ramsay for kindly sharing their invaluable insights and extensive knowledge with me.

It is an honor to acknowledge the financial support of University of Bremen, and Research Training Group ”Quantum Mechanical Materials Modeling” which enabled me to successfully complete my dissertation.

This journey would not be possible without the support of Dr. Omid Mohammad Moradi, Dr. Pedram Salehpoor, Dr. Ahmad Esmaielzadeh Kandjani, Dr. Mahsa Barzegar Vishlaghi, Dr. Hooman Bahmani Jalali, Dr. Michele Compostella, Dr. Mohammad Bahmani, Dr. Elham Khorasani, Kristoffer Rehling, and Robert Hager without whom I would not be in the position that I am in today.

And finally, I cannot begin to express my gratitude to my parents for their endless sacrifice and going against the social norms to provide me with the opportunities that they did not have for themselves. I hope to live a life which honors the time and efforts they have spent for me.

Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	**Meaning**
CB	conduction band
CBM	conduction band minimum
CI-NEB	climbing image nudged elastic band
DFT	density functional theory
EA	electron affinity
GGA	generalized gradient approximation
gKT	generalized Koopmans' theorem
HEG	homogeneous electron gas
IP	ionization potential
LAPW	linearized augmented-plane-wave
LDA	local density approximation
MEP	minimum energy path
MPI	message passing interface
MSE	mean squared error
NEB	nudged elastic band
NLEP	non-local external potential
PEB	plain elastic band
PES	potential energy surface
SCF	self-consistent field
SIE	self-interaction error
STM	scanning tunneling microscopy
STS	scanning tunneling spectroscopy
VB	valence band
VBM	valence band maximum
XC	exchange-correlation

Chapter 1

Introduction

Photocatalysts are used in a wide variety of applications, such as environmental cleaning, solar energy conversion, and organic synthesis [1]. A considerable amount of effort has been dedicated to investigating photocatalytic reactions and their mechanisms, both from an experimental and theoretical perspective. Charge carrier species generated from photoexcitation of photocatalysts may become trapped in bulk or surface defects, or they may self-trapped at low temperatures. Small polarons are charge carriers trapped by strong lattice relaxation at specific sites [2].

Theoretical methods used in the study of photocatalytic reactions should capture all the relevant physical phenomena while being computationally affordable. Density functional theory is widely used for investigating the reactions on the surface of solids. Local or semi-local approximations in density functional theory treat correlation and exchange incompletely, which, in the case of semiconductors, results in an underestimation of band-gap energy, which may cause incorrect occupation of defect states and wrong formation energies. Hybrid functionals, which incorporate a portion of Hartree-Fock type non-local exchange, are more successful in reproducing the physical properties. However, the computational cost of using hybrid functionals makes them impractical for studying surface reactions.

Lany and Zunger have proposed a computationally inexpensive method to correct for the gap error of semi-local approximations, using non-local, empirical external potentials and occupation dependent non-empirical localizing potentials to correct the delocalization error. The latter method works well in the bulk but is not accurate enough on the surface, so I extended it by taking into account the undercoordination of surface atoms. Using this technique, I have found a two-dimensionally localized state in anatase which has led to a study on estimating the real size of polarons in anatase-TiO_2.

The supercell approach is commonly used in ab initio calculations of charged defects in the bulk or on the surface of materials. In the supercell method, the charged defect is placed in a large cell, which is then periodically repeated. However, this approach suffers from one main drawback: the spurious interaction between the charged defect and its periodic images. Several *a posteriori* methods have been proposed for estimating the error in the total energy of models under periodic boundary conditions due to localized extra charges [3]. Komsa and Pasquarello have proposed a correction scheme that is also suitable for slab models. Although calculating the corrections is not trivial and involves a lot of trial and error to find the right parameters. This issue is compounded when the extra charge is localized on multiple sites or on the surface. I have implemented their method in user-friendly code to integrate it into my workflow. My goal is to develop a code that can make accurate corrections for an extra charge with minimal user input. I will show the shortcomings of *a posteriori* correction methods in surface models, which proves the necessity of having a self-consistent correction. I have participated in developing such a method.

CO adsorption and oxidation have been widely studied on various catalysts due to their industrial importance. The interaction of CO and CO_2 with TiO_2 is crucial in many heterogeneous catalytic and photocatalytic reactions, e.g., low-temperature CO oxidation [4], CO hydrogenation [5], water gas shift reaction [6], NO reduction by CO [7], etc. Although TiO_2 is commonly used as a support for noble metal catalysts in CO oxidation [8, 9], it can also be used as a photocatalyst to directly adsorb and oxidize CO molecules [10].

Nevertheless, the details of the photocatalytic CO oxidation reaction, though simple at first glance, have not been fully understood yet. Despite the higher activity of anatase in CO oxidation [10], this reaction has been more widely studied on rutile surfaces, where it has been shown that it cannot be made cyclic, i.e., it is not a real catalytic process [11] as CO_2 is primarily produced from chemisorbed O_2 residing in a surface oxygen vacancy and the oxidation leads to a defect-free surface [12]. Recent theoretical investigations on the anatase (101) surface, however, indicated the possibility of a real catalytic cycle [13], following a different mechanism than the reaction on the rutile surface. The oxidation of CO on the rutile (110) surface is an electron-mediated process requiring the presence of an oxygen vacancy Eq. (1) [12, 14]. In contrast, on the anatase (101) surface, this reaction is hole-assisted, creating an oxygen vacancy Eq. (2) [13].

$$[\text{R-TiO}_2\text{:V}_\text{O}] + O_2 + CO \rightarrow [\text{R-TiO}_2] + CO_2 \qquad (1)$$

$$[\text{A-TiO}_2] + \text{CO} \rightarrow [\text{A-TiO}_2\text{:V}_\text{O}] + \text{CO}_2 \qquad (2)$$

Full cyclic reactions (with surface restoration) on the anatase (101) surface have not been fully studied yet. My ultimate goal was to develop a framework for investigating this reaction and identifying paths that could lead to a real photocatalytic (fully cyclic) reaction.

Chapter 2

State of the art

2.1 Photocatalysis

Photocatalysis, or more precisely, photo-induced/photo-activated catalysis [15], commonly refers to heterogeneous catalytic reactions in which the reaction is induced through photon activation rather than thermal activation [16].

The main steps of the photocatalytic reactions are:

- Absorption of photons by the catalyst (not by the reactants)
- Creation and separation of electron-hole pairs
- Carrier transfer reactions with adsorbed species (photo-oxidation and photo-reduction)

In this discussion, I consider heterogeneous photocatalytic reactions on semiconductor surfaces. The irradiation of a semiconductor by photons with an energy greater than its band gap promotes electrons from their valence band (VB) to the conduction band (CB). The created electron-hole pairs (e^-/h^+) may recombine again without participating in any chemical reaction, or they may interact with the adsorbed species on the catalyst surface, causing electron transfer from a donor to the VB and from the CB to an acceptor. These steps are shown schematically in Figure 1.

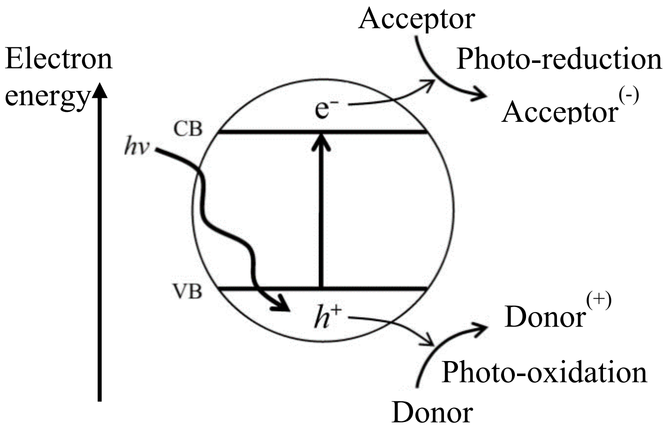


Fig. 1 Schematic representation of the primary steps in heterogeneous photocatalytic reactions on semiconductor surfaces

However, in practice, this simplistic model is not sufficient for the accurate prediction of the photocatalytic process on semiconductors. Most notably, charge traps can affect electron-hole recombination, and charge migration rates, thereby enhancing or reducing the overall photocatalytic activity [17]. Therefore, due to the significant impact of the lattice imperfections on the photocatalytic properties [18], any theoretical model for photocatalysts should be able to adequately account for them.

2.2 Titanium dioxide

Since the first studies of photosorption [19], photocatalytic oxidation [20], and later ultraviolet light-induced photoelectrochemical water splitting [21] on its surfaces, TiO_2 has been regarded as the most representative photocatalytic material and has been extensively investigated [22-24]. Titanium dioxide’s low cost, high resistance to photo-corrosion, non-toxicity, and the possibility of tailoring its properties for a specific application [25] have made it attractive for many environmental applications [26, 27] as well as in photovoltaics [28], sensors [29], etc. More in-depth information on titanium dioxide, its properties, and applications can be found in the reviews [30, 31].

2.2.1 Phases

Titanium dioxide generally occurs in nature in three crystalline forms: rutile (tetragonal), anatase (tetragonal), and brookite (rhombohedral). The latter is difficult to synthesize and therefore less studied [32]. All these structures consist of a titanium atom surrounded by six oxygen atoms

($[TiO_6]^{2-}$ octahedra), with different edge and corner sharing while keeping the stoichiometry fixed as TiO_2. In tetragonal structures, other than the lattice constants a and c, another extra parameter (internal parameter *u*) is needed to fully describe the structure and define the positions of the oxygen atoms. If a titanium atom is placed at the origin, then its nearest oxygen atoms are located at (0, 0, ±*uc*) in anatase, and at (±*ua*, ±*ua*, 0) in rutile. The structure parameters for anatase and rutile, obtained from the neutron diffraction method at room temperature are listed in Table 1. The unit cells of these phases are shown in Figure 2.

Table 1 Experimental structural data for anatase and rutile phases

Phase	***a* (Å)**	***c* (Å)**	***u***	**Ref.**
Anatase	3.785	9.514	0.20806	[33]
Rutile	4.594	2.959	0.30478	[33, 34]

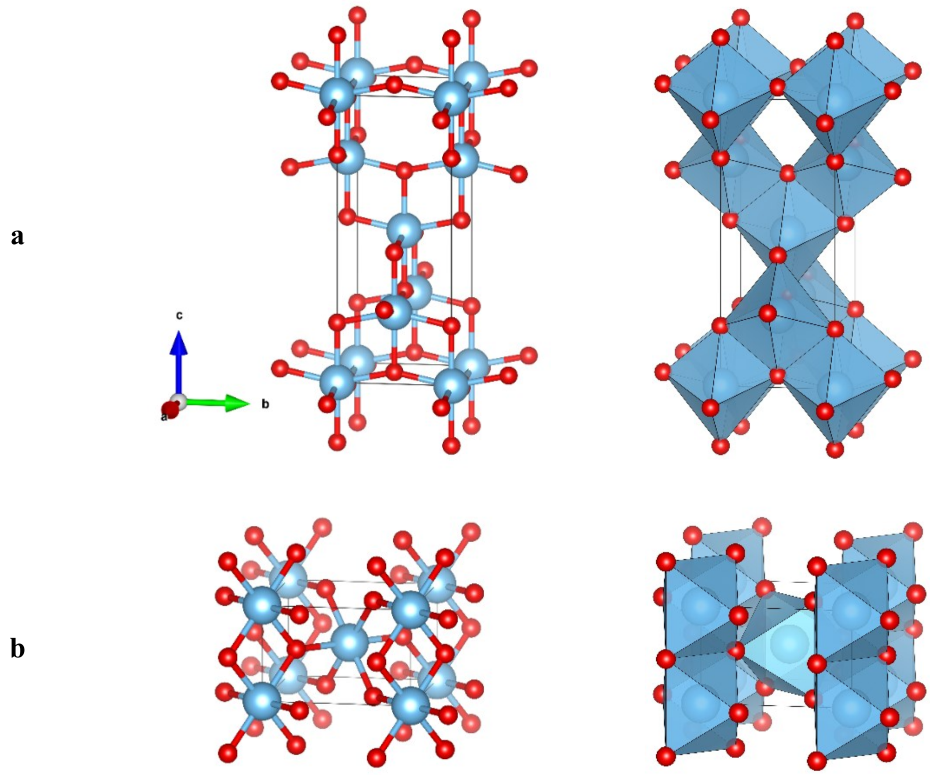


Fig. 2 Ball-and-stick and polyhedral representations of the a) anatase and b) rutile unit cells. Red spheres represent oxygen, and blue spheres represent titanium atoms (based on the structural data from [33])

Anatase and rutile both exhibit photocatalytic properties [35]. In bulk form, rutile is the thermodynamically most stable TiO_2 polymorph, but anatase is, in general, more efficient in photocatalytic and photovoltaic applications due to a lower recombination rate of electron-hole pairs [36]. Recent studies comparing single-crystalline anatase (101) and single-crystalline rutile (110) photocatalytic activities confirmed the superiority of the former [10]. Despite this fact, more surface chemistry studies have been conducted on rutile, which stems from the ease of preparing large synthetic rutile single crystals and their commercial availability [37, 38].

2.2.2 Anatase (101) surface

The properties of semiconductor surfaces are highly dependent on their crystallographic orientation [39]. Under typical preparation conditions for anatase, the (101) and (100)/(010) surfaces as well as some (001) surfaces are most commonly observed [40]. In general, both natural and synthetic anatase crystals are dominated by (101) facets, which correspond to the lowest surface energy in stoichiometric anatase [41, 42]. Anatase (101) surfaces exhibit a sawtooth structure with twofold-coordinated oxygen atoms at [010]-oriented ridges. The other oxygen atoms are threefold-coordinated as in the bulk [43]. The structure of the anatase (101) surface, with the various atom types is shown in Figure 3.

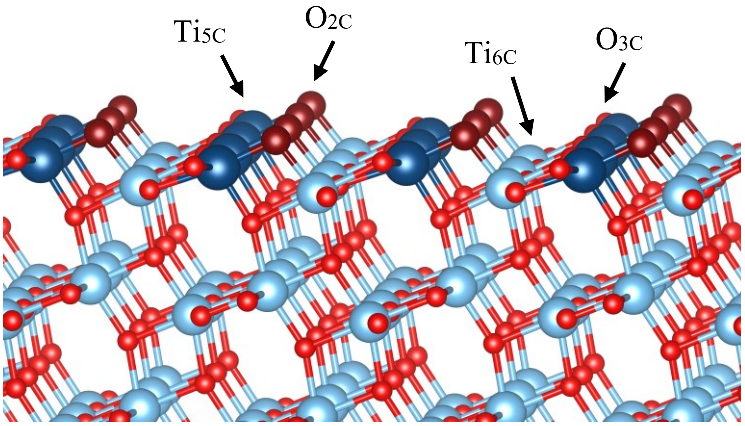


Fig. 3 Structure of the anatase (101) surface with under-coordinated atoms (fivefold-coordinated Ti and twofold-coordinated O) marked and visualized with darker colors

Studies show that the anatase (101) surface does not have a strong tendency to lose twofold-coordinated oxygen atoms upon annealing in ultrahigh vacuum [44]. Also, the bulk and subsurface oxygen vacancies are more stable than a vacancy on the anatase (101) surface [45]. Removal of a twofold-coordinated oxygen from the anatase (101) surface results in the formation of a fourfoldcoordinated and a fivefold-coordinated titanium atom on the surface. The highly undercoordinated titanium atoms are likely the reason for the lower stability of surface oxygen vacancies [44].

2.2.3 Polarons

Both photo-excited electrons and holes in TiO_2 have been found to be trapped in small polaron states both theoretically [46] and experimentally [47], while delocalized free carriers coexist [48]. Large polarons have also been experimentally observed [49] and theoretically studied [50]. Small electron-polarons are commonly observed in rutile [51, 52]. In bulk anatase, small electron-polarons are not stable with respect to the free carrier, while holes are trapped in deep hole-polaron states [53]. The situation on the TiO_2 surfaces is complex and depends on both the phase and its surface orientation [54]. As an example, while electron polarons are found to be stable on some anatase surfaces, they tend to go subsurface in rutile for most orientations [52, 55-58]. These polarons may diffuse to the surface and participate in reactions that require charge transfer from the TiO_2 surface to the adsorbed species [59].

2.3 Modelling surface reactions

In realistic physical systems, there are $>10^{20}$ atoms in each mm^3, which is beyond the reach of any *ab-initio* method with the current computational resources. Nevertheless, the ideal (infinite) periodic solids can be conveniently modeled with a relatively small number of atoms (supercells) repeating in all three dimensions. Even in the case of non-periodic systems, where using the model of a perfect crystal may seem inappropriate, it is possible to create a large enough periodic cell that, within its boundaries, resembles the local structural properties of a disordered system as closely as possible [60]. The translation invariance imposed by the supercell approximation leads to a simplification of the problem and permits the implementation of efficient computational algorithms. Cluster and slab models are the most popular models of the solid surface [61]. Cluster models represent the surface with a finite number of atoms, often treated with localized basis sets [62].

Slab models are supercell models with periodic boundary conditions applied to a layer of vacuum over a bulk-like region, resulting in alternating layers of infinitely large solid slabs and vacuum [62]. Figure 4 shows a 2D representation of the slab model generated from a supercell with vacuum, which results in alternating sheets of infinitely wide layers separated by the vacuum. The highlighted region represents the repeating unit.

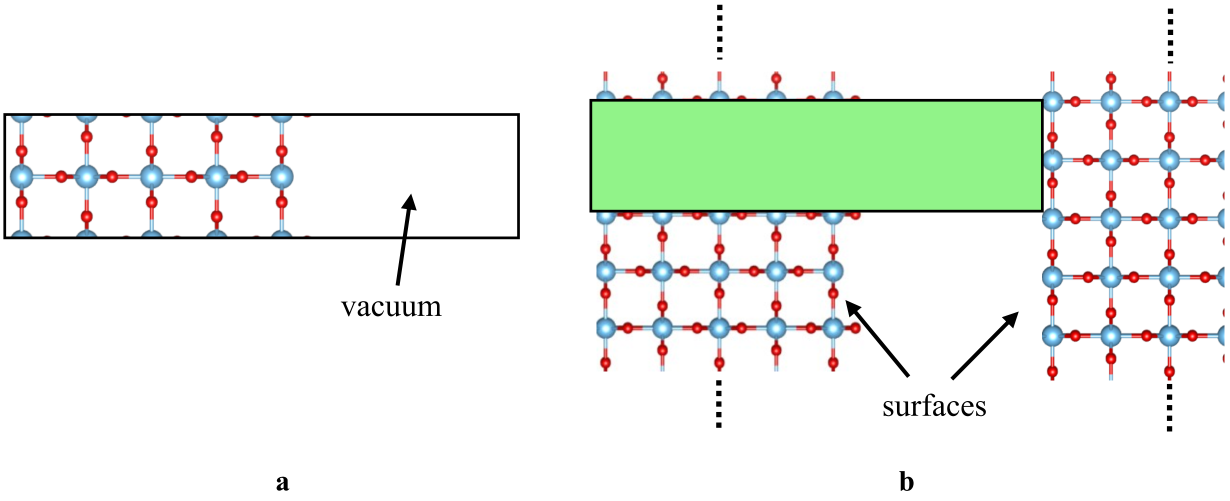


Fig. 4 a) supercell for slab model; b) supercell under periodic boundary conditions generating the infinite sheets separated by a vacuum region.

Slab models have the advantage of describing the band structure of the host crystal. However the periodic boundary conditions, used in slab models, will result in artificial interactions between the repeated images due to the finite size of supercells [63].

2.4 Charged supercell errors

Under periodic boundary conditions, the total energy of neutral systems with no dipole moment is well-defined [64]. However, the total energy of the charged systems depends on an arbitrary shift in the electrostatic potential [65]. Under the periodic boundary conditions within the pseudopotential momentum-space formalism, the otherwise divergent average electrostatic potential within a charged cell is conventionally set to zero [66]. This can be viewed as effective compensation for the net charge by a homogeneous background charge (jellium). However, to compare the energy of different models, the effect of the charge's interactions with its image and jellium must also be removed [67]. Moreover, artificial interactions between the periodic images due to the finite size of supercells in the slab model can cause significant errors in the calculation of the total energy of charged defects and directly influence formation energy, ionization energy, etc., making them supercell-size dependent [63].

Several *a posteriori* methods have been proposed for correcting the total energy due to the interaction of localized extra charge with its periodic images and the neutralizing background, most notably by Makov and Payne [64], Lany and Zunger [68], Freysoldt, Neugebauer and Van de Walle [69], and Komsa and Pasquarello [70]. Comparison of the common correction methods for charged defects can be found in [71, 72].

2.4.1 Charge correction methods

2.4.1.1 Potential alignment

In order to obtain physically meaningful results for the formation energy of the charged defects, one can correct for the undetermined offset in the energy of the models by ensuring that the offset in energy of the charged defect model and the neutral pristine host are consistent. This can be done by adding a potential alignment term to the formation energy [68].

The common approach for potential alignment is to calculate the potentials in charged and neutral systems by averaging over a number of points at distances "far" from the localized charge inside the supercell and applying this difference to the energy of the charged supercell [73]. However, in practice, due to the slow decay of the Coulomb potential in charged defects, using this method can be problematic [3].

2.4.1.2 Finite size scaling

Another conceptually simple approach to overcome charged error is to calculate the energies in supercells of various sizes and extrapolate the results to the limit of an infinitely large supercell [74]. Nevertheless, due to the long-range Coulomb interaction between the localized image charges, the resulting energies converge very slowly with respect to the supercell size [75, 76]. Moreover, depending on the origin of finite size errors and the employed scaling scheme, the exact form of the energy vs. inverse supercell size dependency can be different from what has been implicitly assumed [77]. Despite the reliability of the finite-size scaling method [75], due to its cost both in human and computational time, it is more convenient to estimate and eliminate these errors in smaller supercells.

2.4.1.3 Makov and Payne approach

Makov and Payne [64] proposed an image charge energy correction for a localized defect inside a cubic supercell without a net dipole in the form of:

$$\Delta E = \frac{q^2\alpha}{2L\varepsilon} + \frac{2\pi q Q}{3L^3\varepsilon} \tag{3}$$

where q is the charge of the system, α is the Madelung constant of the lattice, which can be determined by Ewald summation [78], L is the side length of the cubic supercell, ε is the dielectric constant, and Q is the second radial (quadrupole) moment of the defect charge distribution. The first term corresponds to the Madelung energy (point charge approximation), and the second term is the interaction of a localized charge distribution with the compensating background. Both terms are scaled by the dielectric constant of the medium. Calculating the quadrupole term Q is straightforward but computationally costly; therefore, the second term is often omitted. This method has also been extended to models with arbitrary supercell shapes [79] and anisotropic dielectric screening [80], but it is only valid for strongly localized charges.

2.4.1.4 Lany and Zunger charge correction

Lany and Zunger [68] proposed a practical approximation to Q in Makov and Payne's correction scheme by using the difference between the charged and neutral defect far from the defect as $\approx q/L^3(1-\varepsilon^{-1})$ and used it for calculating the quadrupole moment as:

$$Q = -\frac{q}{4}L^2\left(1-\frac{1}{\varepsilon}\right) \tag{4}$$

Therefore, the full charge correction can be approximated by a first-order correction. Lany and Zunger also showed the necessity of including a potential alignment in the energy correction, and proposed a correction scheme as:

$$\Delta E = \left(1 + c_{sh}(1-\varepsilon^{-1})\right)\Delta E_i^1 + q.\Delta V \tag{5}$$

where c_{sh} is the shape factor of the supercell (-0.369 for simple cubic), ΔE_i^1 is the first order image charge correction, equivalent to the first term in Makov-Payne's charge correction Eq. (3), and ΔV is the potential alignment, which is calculated using the average of the atomic-site electrostatic potentials, as reference, excluding the immediate vicinity of the defect. If the dielectric constant is sufficiently large and the supercell is approximately isotropic, the image charge correction term reduces to $\approx {}^2/_3\,\Delta E_i^1$.

2.4.1.5 Freysoldt, Neugebauer, and Van de Walle approach

Freysoldt, Neugebauer, and Van de Walle [69] proposed a more elaborate scheme for energy correction by modeling the extra charge with a Gaussian charge distribution embedded into a uniform isotropic dielectric medium. They decomposed the effects of the extra charge (q) into a short-range potential (V^{sr}), which accounts for the variations in the microscopic screening, and a long-range potential (V^{lr}), which is dominated by the macroscopically screened Coulomb potential. The total energy correction ΔE is obtained as follows:

$$\Delta E = \frac{1}{2}\int \left([q(r)+n]\left[\Delta V_q^{lr}(r)\right] + nV_q^{lr}(r)\right) d^3r + q\Delta V \tag{6}$$

$$V_q^{lr}(r) = \frac{1}{\varepsilon}\int \frac{q(r')}{|r-r'|} d^3r' \tag{7}$$

$$\Delta V = \int V_q^{sr}(r) d^3r \tag{8}$$

where n is the neutralizing background (jellium), ΔV^{lr} is the artificial long-range potential due to periodic repetition, and ε is the macroscopic dielectric constant, which can be computed from density-functional perturbation theory [81] or from the response to a finite sawtooth potential [82]. In practice, the potential alignment term ΔV relies on the short-range potential V^{sr} decaying to zero far from the charged defect, and therefore, the difference in the average potentials of the perfect and the charged defect models is used in its determination.

This method has been implemented in the S/PHI/nX package [83] and is also available as a stand-alone tool (sxdefectalign). It can be easily applied to supercells of various shapes and has been extended to model the extra charge in anisotropic mediums [65]. Usually, it uses a planar-averaged electrostatic potential in the alignment term, which cannot be easily calculated for defects with large atomic relaxations.

2.4.1.6 Komsa and Pasquarello approach

Komsa and Pasquarello proposed a correction scheme, considering the existence of surfaces and interfaces in the supercell, demonstrated its reasonable performance in slabs [70], and extended the method to calculate the charge correction for 2D materials [84, 85]. In this approach, the extra charge is also modeled with a Gaussian distribution, and the correction is calculated as:

$$\Delta E = E_{isolated} - E_{periodic} + q.\Delta V \tag{9}$$

where $E_{periodic}$ and $E_{isolated}$ are the energies of the charge distribution with, and without the periodic boundary conditions, respectively. $E_{periodic}$ includes the interactions with the periodic images and with the background charge. ΔV is the model potential alignment term, which can be calculated from the difference between the potential resulting from the model charge and the electrostatic potential obtained from the electronic-structure calculations far from the defect.

$E_{periodic}$ can be calculated as:

$$E_{\text{periodic}} = \frac{1}{2}\int V_{\text{model}}(r)\rho_{\text{model}}(r)\ \mathrm{dr} \tag{10}$$

where $V_{model}(r)$ is the electrostatic potential originating from charge $\rho_{model}(r)$ under periodic boundary conditions and can be obtained through the solution of the Poisson equation:

$$\nabla \cdot \left(\varepsilon(r)\nabla V_{\text{model}}(r)\right) = -\rho_{\text{model}}(r) \tag{11}$$

The energy of an isolated defect $E_{isolated}$ can be obtained analytically for simple cases [70] or through extrapolation of $E_{periodic}$ to an infinitely large size by uniformly scaling the system in all directions.

Chapter 3

Methods

3.1 Electronic structure calculations

Nowadays, electronic structure calculations play a key role in understanding the reaction mechanisms and predicting materials' properties, which makes it possible to design new catalysts and help the interpretation of the experimental data [86-88].

Within the Born-Oppenheimer approximation [89] (stationary nucleus), the nonrelativistic electronic Schrödinger equation $\hat{H}\psi = E\psi$ for N interacting electrons in a time-independent external potential $v_{ext}(r)$ can be described with a Hamiltonian operator (in atomic units) as:

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{N} \nabla_i^2 + \sum_{i=1}^{N} v_{ext}(r_i) + \sum_{i=1}^{N}\sum_{i<j}^{N} \frac{1}{|r_i - r_j|} = \hat{T} + \hat{V}_{ext} + \hat{V}_{ee} \qquad (12)$$

where T is the kinetic energy operator, V_{ext} is the potential energy operator for the interaction of the electrons with the external field, and V_{ee} is the potential energy operator for electron-electron interaction. In our cases, the external potential is simply the Coulomb potential of the nuclei. The solution of the Schrödinger equation with this Hamiltonian is a many-body wave function ψ that depends on all electronic degrees of freedom. Currently, there is no general method for the exact solution of this equation for realistic physical models that is feasible with the currently available computational resources [90].

Most of the calculations on heterogeneous catalytic systems nowadays use the DFT (density functional theory) approach [91, 92]. DFT is a first-principle-based method that also includes electron correlation at relatively low computational cost [93].

3.1.1 Density Functional Theory

Historically, the 1964 paper of Hohenberg and Kohn [94] marks the start of DFT as it is commonly known today. The 1st Hohenberg-Kohn theorem states that the external potential, to within a constant, is a unique functional of the charge density and charge density only. The 2nd theorem states that there is a universal functional for energy that depends only on the charge density and is valid for any external potential (hence, any atomic configuration). Using this functional, the total energy reaches its minimum value with respect to all possible charge densities if and only if the charge density is the true ground state. To sum up, the core concept of DFT is that the total electron density alone completely and exactly contains all the information about the ground state of a many-electron system. The ground state energy E is a variational functional of electron density ρ only, and the universal Hohenberg-Kohn functional $F[\rho]$ includes all the interactions, i.e. the kinetic energy functional $T[\rho]$, and the full electron-electron interactions functional $J[\rho]$.

$$E[\rho] = V_{ext}[\rho] + F[\rho] \tag{13}$$

$$F[\rho] = T[\rho] + J[\rho] \tag{14}$$

Hohenberg-Kohn presented finding a 3-dimensional charge density as an alternative to solving the 3N-dimensional many-body Schrödinger equation for calculating all the ground state properties. Although the proof was originally provided for the "v-representable"[1] densities with non-degenerate ground states and integer electron numbers, the theory was later extended to various degrees beyond these strict requirements [95-98].

3.1.2 Kohn-Sham approach

Eq. (14) has two unknown functionals: the kinetic energy of the electrons T and the non-classical part of the electron-electron interaction J. In 1965, Kohn and Sham [99] introduced the concept of using a non-interacting electronic system with the same density as the reference system but built from a set of one electron functions, usually referred to as Kohn-Sham orbitals, for which the kinetic energy can be easily calculated.

Within the Kohn-Sham formulation, the electronic energy of the ground state of a system comprising n electrons and N nuclei can be written as sum of the kinetic energy of the non-interacting electrons (T_0), interactions of electrons with external potential (nuclei with charge Z) (V_{ext}), Coulomb (classical) self-interaction of electron density, i.e. electron–electron repulsion (J_0) and exchange-correlation (XC) term (E^{XC}) which is the exchange and correlation energy of an

---

[1] A v-representable charge density is associated with an anti-symmetric ground-state wave function

interacting system and incorporates all the non-classical effects i.e. correction to the kinetic energy arising from the interacting nature of the electrons and non-classic corrections to the electron-electron repulsion energy.

$$E[\rho] = T_0[\rho] + V_{ext}[\rho] + J_0[\rho] + E^{XC}[\rho] \tag{15}$$

$$T_0[\rho] = -\frac{1}{2}\sum_{i=1}^{n} \int \psi_i^*(r)\, \nabla_i^2 \psi_i(r)\, dr \tag{16}$$

$$V_{ext}[\rho] = -\sum_{X=1}^{N} \int \frac{Z_X}{r_X} \rho(r)\, dr \tag{17}$$

$$J_0[\rho] = \frac{1}{2} \frac{\rho(r_1)\rho(r_2)}{r_{12}} dr_1 dr_2 \tag{18}$$

$$\rho(r) = \sum_{i=1}^{n} \int f_i |\psi_i(r)|^2\, dr \tag{19}$$

$$n = \sum_{i=1}^{n} f_i \tag{20}$$

$$\int |\psi_i(r)|^2 dr = 1 \tag{21}$$

where Ψ_i are the Kohn-Sham orbitals and f_i are the occupations of each one as $0 \leq f_i \leq 1$. The electronic ground state is determined by minimizing the total energy functional of Eq. (15) at a fixed ionic geometry. The charge density can be constructed from the Kohn-Sham orbitals and their occupation according to Eq. (19) subject to the constraints of Eqs. (20-21). The normalization of the occupied or partially-occupied Kohn-Sham orbitals is guaranteed by Eq. (21). These equations can be alternatively expressed as a nonlinear eigenvalue problem:

$$\left[-\frac{1}{2}\nabla^2 + v_{KS}(r)\right] \psi_i(r) = \varepsilon_i \psi_i(r) \tag{22}$$

$$v_{KS}(r) = v_{ext}(r) + v_H[\rho(r)] + v_{XC}[\rho(r)] \tag{23}$$

$$v_H[\rho(r)] = \int \frac{\rho(r')}{|r - r'|}\, dr \tag{24}$$

$$v_{XC}[\rho(r)] = \frac{\delta E^{XC}[\rho(r)]}{\delta \rho(r)} \tag{25}$$

where v_H is the Hartree potential and v_{XC} is XC potential. This leads to an effective one-electron formulation of our quantum many-body problem, with the XC energy functional/potential as our only unknown term.[2] Although the ground state charge densities can be obtained through the direct minimization of the total energy E with respect to the basis set coefficients [100-102], the objective function of the optimization depends on many parameters and is highly nonlinear with many local extrema.

The resulting Schrödinger-like equation of the Kohn-Sham formulation [Eq. (22)] is a non-linear eigenvalue equation with a Hamiltonian that implicitly depends on the wave function. This eigenvalue problem is commonly solved using the self-consistent field (SCF) method, which removes the nonlinearity of the Kohn-Sham eigenvalue equations by solving them for an approximate input Hamiltonian. The resulting Kohn-Sham orbitals will yield a charge density that, in general, is not consistent with the input Hamiltonian, but this density can be used to improve the approximate Hamiltonian.

The initial guess for the charge density can be chosen randomly, or based on other methods such as the superposition of the atomic densities or the extended Hückel method [103]. Then the Kohn-Sham potential is generated for this charge density. Evaluation of the Hartree potential in Eq. (24) requires solving a Poisson equation, which is commonly done using the fast Fourier transform [104], fast multipole method [105], or multigrid approach [106]. These algorithms can be efficiently implemented in parallel [107]. The constructed eigenvalue problem is commonly solved using the conjugate gradient method [108], residual vector minimization-direct inversion in the iterative subspace (RMM-DIIS) scheme [109], Chebyshev-filtered subspace iteration [110], or block Davidson algorithms [111] such as the locally optimal block preconditioned conjugate gradient (LOBPCG) [112]. Despite the existence of efficient parallel algorithms, this step is usually the most time-consuming part of the calculation for a large system [113].

After solving the Kohn-Sham equations, instead of using the obtained density from the previous SCF step directly for the next step, the iterative improvement of the charge density, i.e., mixing the new electron density with the result of the previous SCF steps, is often more efficient and also preferred [114]. The common mixing schemes include the Newton-Broyden methods [115],

---

[2] It should be noted that the Kohn-Sham method is an exact formulation of the Hohenberg-Kohn theorem.

Anderson method [116], and Pulay mixing [117] (also known as direct inversion on the iterative subspace, DIIS). The obtained self-consistent charge density can be used to calculate the ground state energy of the system. The changes in charge density or total energy are often used as the convergence criteria in the SCF cycle. It can be shown that by using a linear combination of the input and output densities in each step, this procedure is guaranteed to converge using the exact functional [118]. A simplified schematic flow chart of this process is depicted in Figure 5.

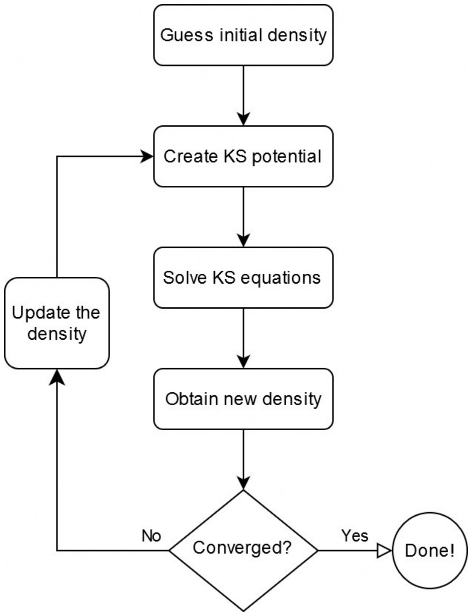


Fig. 5 A simplified flow-chart of self-consistent approach for solving the Kohn-Sham equations

The predictive power of DFT combined with the computational convenience of the Kohn-Sham formulation has led to the development of numerous software packages that provide the possibility of applying this method to the simulation of a wide variety of models. Some of the most popular packages for solid-state simulations include ABINIT [119], CASTEP [120], CP2K [121], CRYSTAL [122], FHI-aims [123], GPAW [124], QUANTUM ESPRESSO [125], S/PHI/nX [83], SIESTA [126], VASP [114, 127], and WIEN2k [128].

3.1.3 Brillouin zone sampling

Although the Kohn-Sham formulation of DFT makes it possible to find the ground state distribution for any number of electrons in an external potential, in reality explicitly performing these calculations for the models with practical size (on the order of 10^{23} atoms) using our current computational power and storage capacity is not feasible. To calculate the properties of bulk materials, one can use the translational symmetry of the [ideal] crystalline models to drastically decrease the amount of necessary computations.

Following Bloch's theorem, the wavefunction of an electron in an external periodic potential can change only by a phase factor upon translation by a lattice vector:

$$\psi(r+R) = e^{ikR}\psi(r) \tag{26}$$

where R is a lattice vector. This constraint on the wavefunction can also be stated as follows:

$$\psi_k(r) = e^{ik\cdot r}u_k(r) \tag{27}$$

where $u_k(r)$ is a function with the same periodicity as the external potential and k is the Bloch wave vector, which [for an infinite lattice] is a continuous quantum number with unique values only within a single unit cell of the reciprocal space. Therefore, in periodic systems, it is sufficient to know the wave functions within the primitive unit cell of the reciprocal lattice (first Brillouin zone). For example, the charge density can be obtained as follows:

$$\rho(r) = \sum_{i=1}^{n}\int_{BZ}\frac{1}{\Omega_{BZ}}f_{i,k}\left|u_{i,k}(r)\right|^2 d^3k \tag{28}$$

where Ω_{BZ} is the volume of the first Brillouin zone, and the occupation numbers are $f_{i,k} = 1$ for occupied states and 0 otherwise. Typically, the lattice-periodic factor of the wave function $u_k(r)$ is weakly dependent on k. Therefore, it is possible to estimate the Brillouin zone integral by sampling its value at a finite set of k points.

$$\rho(r) = \sum_{i=1}^{n}\sum_{k\in BZ}\frac{1}{\Omega_{BZ}}f_{i,k}\left|u_{i,k}(r)\right|^2 \tag{29}$$

In principle, it is possible to calculate this integral accurately by choosing the k-points on a very fine mesh. However, in practice, this approach requires a considerable amount of computational power. For sufficiently smooth functions, it is possible to obtain accurate results by using a carefully selected set of points (special points), in reciprocal space [129, 130].

The Monkhorst-Pack k-point set [131, 132] is a commonly used, regularly spaced k-point set for sampling the Brillouin zone. Although more efficient Brillouin zone sampling methods are also available [133, 134]. The Monkhorst-Pack set can be represented as:

$$\vec{k} = \sum_{i=1,2,3} \frac{2n_i - N_i - 1}{2N_i} \vec{b}_i$$

$$n_i = 1 \ldots N_i \tag{30}$$

where $\vec{b}_i$ are the reciprocal lattice vectors, and N_i are the number of subdivisions in each direction. The Monkhorst-Pack k-point set is usually specified by a product of three integer numbers, e.g., 8×8×8 which represents the number of reducible k-points in each direction. However, in practice, only the irreducible k-points are used in the calculations, which is a smaller set of k-points with the symmetrically equivalent ones removed. It is common to use the linear tetrahedron method for the Brillouin zone integration instead of simple histogram integration, which provides more accurate results with the same number of k-points [135, 136].

3.1.4 Plane wave basis

Finite basis approximation is an important step in converting the Kohn-Sham equation into an algebraic system of equations that can be solved numerically. Kohn-Sham orbitals Ψ_i are continuous functions of a continuous three-dimensional variable (r). These orbitals can be approximated by the basis functions as:

$$\psi_i(r) = \sum_{j=1}^{\infty} c_{ji}\phi_j(r) \approx \sum_{j=1}^{N_b} c_{ji}\phi_j(r) \tag{31}$$

where the φ_j functions form a complete basis. In principle, however, the expansion is truncated by some criterion.

Modern electronic structure methods, based on their choice for the form of basis functions, can be broadly categorized into two classes: the plane wave basis methods and the localized basis set methods.

Plane wave methods have several distinct advantages over localized basis sets, making them particularly attractive for the calculation of solid-state models. A few of the relevant ones for our discussion are:

- The plane wave basis set is independent of atomic positions and can give an unbiased description of the whole simulation model, including the bulk, vacancies, surfaces, and vacuum regions.

- Highly efficient fast Fourier transform algorithms can be used for real-space to reciprocal-space transformations.
- Although, in principle, an infinite basis set is required to expand the wave functions, in the case of the plane-wave basis, the expansion coefficients for the plane waves with smaller kinetic energy are more important. Therefore, the basis set can be truncated to only include plane waves with kinetic energies that are less than a specific cutoff energy. Unlike the localized basis sets, achieving basis set convergence is simple and systematically controlled by the cut-off energy in the plane wave basis set.
- The Hellmann-Feynman forces [137] can be easily calculated from the expectation value of the Hamiltonian with respect to the ionic coordinates without any Pulay forces [138].
- Basis-set superposition errors, which have to be carefully estimated in calculations with local basis sets are completely avoided.

On the other hand, there are some shortcomings in using the plane wave basis. Notably, the local atomic/bonding properties are not readily available, and a projection of the plane wave basis on a localized basis is usually required to retrieve this information [139]. Also, large cut-off energies are needed to accurately describe localized states using plane waves. This increases the computational cost for systems that include first-row atoms.

Moreover, in surface calculations with slab models, especially for investigating surface reactions, the distance between the slabs needs to be large enough to avoid artificial interactions. This makes the size of the supercell in the direction normal to the surface quite large and the corresponding reciprocal lattice vector quite short, resulting in a large number of plane waves within a given cutoff energy [140].

3.1.5 Linearized augmented-plane-wave

Rapid oscillations of the wavefunction in the region near the nucleus present a challenge for the plane wave expansion. In this region, the plane wave expansion converges very slowly, and a large number of plane waves is needed to represent the wavefunction accurately. The augmented plane-wave method [141] solves this issue by expanding the wavefunction with the spherical harmonics in the region surrounding the atoms but using the plane waves outside these spheres and joining these two continuously at the surface. However, there are several limitations inherent to this method, e.g., the formulation as an energy-dependent characteristic equation and the continuity problem when the spherical harmonics are zero at the surface, which makes it inefficient in practice. To avoid these problems, the linearized augmented plane-wave (LAPW) method was proposed [142], which adds the energy derivative to the original energy dependence of the radial basis-function. In the LAPW method, the derivatives are also kept continuous through the spherical region's boundary, which increases the number of plane waves required for the calculations.

3.1.6 Pseudopotentials

The core region near the nuclei is primarily composed of tightly bound core electrons that do not actively participate in chemical bonding. In most cases, we are solely interested in a relatively small change in the total energy of our model system due to the rearrangements of the valence electrons, and therefore the core electrons of the atoms can be safely ignored. This problem can be mitigated by adding pseudopotentials to the Kohn-Sham potential, which are carefully designed repulsive potentials and only affect a small region around the nuclei.

There are multiple ways to construct the pseudopotentials. Two commonly used ones include norm-conserving pseudopotentials [143], which conserve the core-region charge and are analytically straightforward to treat, and ultrasoft pseudopotentials [144], which generate a smoother wave function by including a compensating augmentation charge, allowing for a smaller basis set.

3.1.7 Projector-augmented wave method

By combining the ultrasoft pseudopotentials with the LAPW method, we can avoid all-electron calculations but achieve high accuracy with small numbers of plane waves (large grid spacing). This approach is known as the projector augmented wave (PAW) method [145] and can be efficiently implemented in the ultrasoft pseudopotential codes [127].

3.1.8 Exchange-correlation functionals

The only unknown term of the total energy expression for the Kohn-Sham system, Eq. (15) is the XC energy E^{XC}. Although DFT in the Kohn-Sham formulation is formally an exact theory for the ground-state electronic properties of a given system, it depends on an unknown ("very complicated and essentially uncomputable" [146]) universal XC functional (density functional) that has to be approximated. After the introduction of DFT, dozens of non-empirical and semi-empirical XC functionals have been developed, and finding accurate and transferable XC functionals is still an active field of research [147, 148].

3.1.8.1 Local Density Approximation

The local density approximation (LDA) is the simplest approach to representing the XC functional. LDA implicitly assumes that the XC energy at any point in space is a function of the electron density at that point in space only, i.e., the "local" value of ρ. In practice, the most widely used LDA functionals are those derived from an infinite homogeneous electron gas (HEG) model of the same density [99].

$$E_{LDA}^{XC} = \int \rho(r) e_{HEG}^{XC}(\rho) d^3r = \int \rho(r)[e_{HEG}^{X}(\rho) + e_{HEG}^{C}(\rho)]\, d^3r \tag{32}$$

where $e_{HEG}^{XC}(\rho)$ is the XC energy per electron of the HEG with a density of ρ. This approximation was later generalized to spin-polarized systems [149].

Within the LDA approximation, the XC functional is assumed to be decomposable into the exchange and correlation energy functionals. While the exchange functional $e_{HEG}^{X}(\rho)$ can be easily calculated analytically, the correlation functional $e_{HEG}^{C}(\rho)$ is only analytically known in extreme (very high and low density) limits, and is commonly estimated from fitting to accurate quantum Monte-Carlo simulations [150]. In general, LDA underestimates the equilibrium lattice constants of the solids by ~1% and may overestimate the binding energies by as much as by ~1 eV/atom [151].

3.1.8.2 Generalized gradient approximation

In real atomic/molecular/condensed systems, the electron density is a rather rapidly varying spatial function. To allow for these variations, a simple, intuitive approach is to include the gradient of the electron density in the XC functional [152]. Generalized gradient approximation (GGA) methods take the spatial inhomogeneity of the electron density into account by making the XC energies dependent not only on the density, but also on the "semi-local" gradient of the density $\nabla\rho$ [153]. In general, these functionals can be expressed as:

$$E_{GGA}^{X} = \int \rho(r) e_{HEG}^{X}(\rho) g_{GGA}^{X}(\rho, \nabla\rho) d^3r \tag{33}$$

where *g(ρ, ∇ρ)* is the local inhomogeneity correction factor. In comparison with LDA, GGA functionals usually give a better estimation of the atomization energies and the energy barriers [154].

3.1.8.2.1 PBE

The Perdew-Burkre-Ernzerhof (PBE) functional is one of the most widely used GGA-level XC functionals [154] and has been implemented in many DFT packages [92]. Although there have been several attempts to improve the PBE functional at the GGA level, namely revPBE [155], RPBE [156], and PBEsol [157], by modifying the form of the enhancement factor or the values of the employed constants, unfortunately these functionals do not universally improve upon the PBE results [158-161]. PBE is designed to satisfy a selected set of theoretical constraints on the properties of the universal functional. The exchange part of this functional includes a semi-local enhancement factor term *F*, which in turn depends on a dimensionless reduced density gradient *s*.

$$E_{PBE}^{X} = \int \rho(r) e_{HEG}^{X}(\rho) F_{PBE}^{X}(s) d^3 r \tag{34}$$

$$e_{HEG}^{X}(\rho) = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \rho(r)^{1/3} \tag{35}$$

$$F_{PBE}^{X}(s) = 1 + \kappa - \frac{\kappa}{1 + \mu s^2/\kappa} \tag{36}$$

$$s(r) = \frac{1}{2(3\pi^2)^{1/3}} \frac{|\nabla \rho(r)|}{\rho(r)^{4/3}} \tag{37}$$

where $\mu = 0.21951$, and $\kappa = 0.804$ are the constants that have been determined from the satisfaction of the correct behavior in the HEG limit, and the Lieb-Oxford bound [155] ($|E^{XC}| \leq 2.275|E_{LDA}^{X}|$) locally [156], respectively. Only the range of $0 < s < 3$ is important in most of the physical models [157].

3.1.9 Errors in common approximations of DFT

As mentioned earlier, although DFT is formally exact, in practice, DFT calculations rely on an approximation of the unknown XC functional. Considerable effort has been devoted to investigating the properties of the exact universal functional and using the results as a guide for the development of approximate functionals [158, 159]. However, a systematic approach for constructing or improving universally applicable functionals has not yet been realized [160-162].

3.1.9.1 Self-interaction error

For any system with one electron, T_0 and V are the exact kinetic and potential energies of the system, and there are no electron-electron interactions. This requires the sum of E^{XC} and J_0 to be zero [163]. Unfortunately, in LDA/GGA functionals, this sum is not zero, and these functionals suffer from one-electron self-interaction error (SIE). This issue also exists beyond the one electron limit [164].

SIE is commonly manifested as too much charge delocalization (delocalization error) [165]. This failure in predicting the stability of free vs. self-trapped charges can prevent the formation of small polarons [166]. SIE is also responsible for the underestimation of bandgaps by LDA/GGA functionals, which negatively impacts related quantities such as the formation energy of defects [167]. SIE is also the underlying reason for too low total energies, especially for systems with a non-integer number of electrons [168]. This is important in the calculation of the energy of transition-state complexes in chemical reactions.

3.1.9.1.1 Kohn-Sham Orbital energies

Although DFT itself assigns no formal interpretation to the Kohn-Sham orbitals and their energies, they have been proven to be more than just an artificial construct, and provide a useful approximation for some molecular properties [169], and a lot of effort have been invested in extracting chemical information from them [170]. For example, according to Janak's theorem [171], the Kohn-Sham orbital energies in DFT satisfy the following condition:

$$\frac{\partial E}{\partial f_i} = \varepsilon_i \tag{38}$$

where E is the total energy, and f_i and ε_i are the occupation and energy of the i-th Kohn-Sham orbital, respectively.

Also, a necessary non-self-interaction condition can be expressed in terms of the Kohn-Sham orbitals as [172]:

$$\left.\frac{\partial \varepsilon_i(f')}{\partial f'}\right|_{f'=f_i} = 0, \quad 0 \le f_i \le 1 \tag{39}$$

3.1.9.1.2 Non-integer electron number- linearity

Perdew, Parr, Levy, and Balduz [96] extended the Hohenberg-Kohn theorem to systems with fractional electron number and showed that for the exact functional, the total energy is a piecewise linear function of the total electron numbers, with a derivative discontinuity at any integer number. This requirement directly translates into a condition for the XC functional in Kohn-Sham DFT.

Using Janak's theorem, Eq. (38), the necessary non-self-interaction condition, Eq. (39) can be expressed as a condition on the total energy curvature:

$$\frac{\partial^2 E}{\partial f_i^2} = 0 \tag{40}$$

Unfortunately, the energies calculated by the LDA/GGA XC functionals or using the Hartree-Fock (HF)-type (non-local) exchange do not exhibit such behavior. Using Hartree-Fock-type exchange gives rise to a concave total energy function. In contrast, LDA- and GGA-based DFT functionals give rise to a convex total energy function. This behavior is schematically shown in Figure 6.

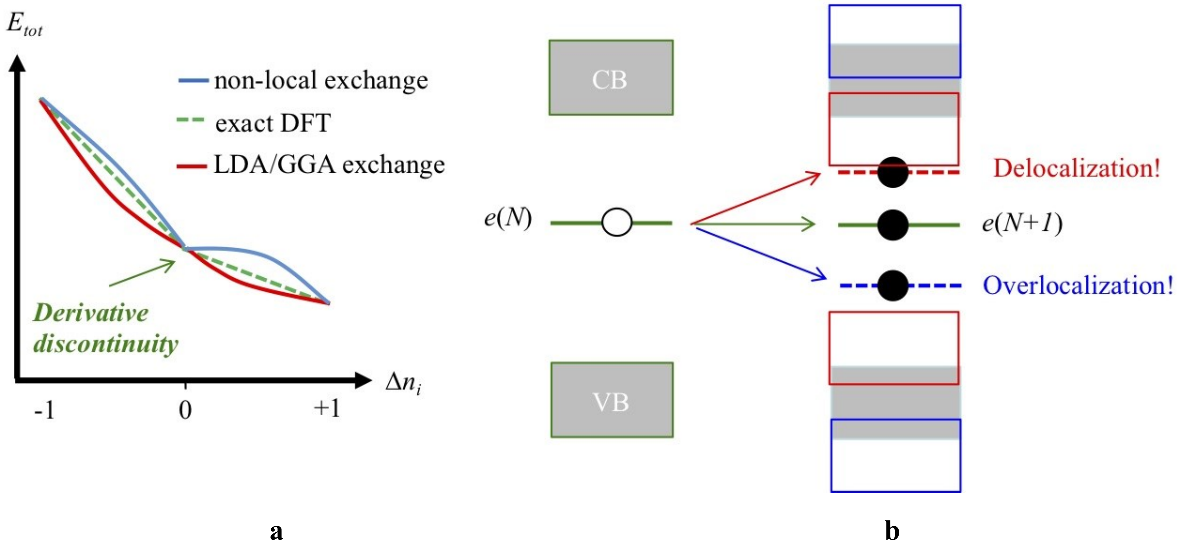


Fig. 6 Schematic presentation of change in a) total energy with fractional electron charges, and b) the Kohn-Sham orbital energies with the addition of an electron for different levels of approximations

As a manifestation of SIE [164], the under- or overestimated derivative discontinuity gives rise to a much too small or much too wide band gap, respectively [173-175], and the nonlinearity is shown to be related to the delocalization of Kohn-Sham states [176].

3.1.9.1.3 Koopmans' theorem

It has been shown that the original Koopmans' theorem [177], which was obtained using the frozen MO approximation for closed-shell systems in Hartree-Fock theory, is an exact relation in Kohn-Sham DFT for a fixed geometry of spin-polarized extended/finite systems [178].

The generalized Koopmans' theorem (gKT) in Kohn-Sham DFT states that by including the correlation and orbital relaxation, the first ionization potential (IP) of the N-electron system is equal to the negative of the highest occupied Kohn-Sham orbital energy (ε_{max}) [178-180]. Applying the gKT to the N+1 electron system will yield a similar relation for the electron affinity (EA) as in Eq. (38) or through Janak's theorem to Eq. (39) [175]. On the basis of the necessary non-self-interaction condition, this implies that EA should also be equal to the lowest unoccupied Kohn-Sham orbital energy (ε_{min}) of the N-electron system [181].

$$IP = E(N-1) - E(N) = -\varepsilon_{max}(N) \tag{41}$$

$$EA = -(E(N) - E(N+1)) = \varepsilon_{max}(N+1) \tag{42}$$

$$EA = \int_0^1 \varepsilon_{max}(N+f)df = \varepsilon_{max}(N+1) \tag{43}$$

$$EA = \varepsilon_{min}(N) \tag{44}$$

Since derivative discontinuity and piecewise linearity are very closely related, problems in both can be resolved by corrections based on the fulfillment of the generalized Koopmans' condition instead of explicit linearity evaluation [182]. Several approaches have been employed to mitigate the aforementioned problems with approximate functionals, among which hybrid functionals, and occupation-dependent nonlocal potentials are most commonly used.

3.1.9.2 Common mitigation methods

3.1.9.2.1 Hybrid Functionals

Hybrid density functionals add a fraction of the non-local exchange of Hartree-Fock theory to DFT XC functionals, thereby compensating the nonlinearity of LDA/GGA functionals with Hartree-Fock exchange. Although adiabatic connection methods [183] can be used for non-empirical estimation of the mixing ratio [184], it has been shown that when using a GGA functional, the mixing ratio cannot be a single constant [185]. Regardless, by using screened hybrid functionals like the one proposed by Heyd, Scuseria, and Ernzerhof (HSE) [186, 187], we can partially overcome the problems due to SIE [188] and achieve high accuracy [189]. However, in practice, for typical systems, hybrid functional calculations are one to two orders of magnitude more expensive than GGA calculations [190], and this increase in calculation time prevents hybrid functionals from being widely accessible, especially for investigating surface reactions using the slab models.

3.1.9.2.2 DFT+U

One common low-cost solution to the shortcomings of LDA/GGA is to use the DFT plus Hubbard U approach [191], which includes an on-site Coulomb repulsion potential U for the explicit treatment of electronic correlation with a Hubbard-like model [192]. The total energy functional of the system can be expressed as the sum of the original DFT energy E^{DFT}, an atomic orbital occupation-dependent energy correction for interaction within the (commonly d or f) shells E^{U}, and a double-counting correction term E^{dc} [193].

$$E^{DFT+U} = E^{DFT} + E^{U} - E^{dc} \tag{45}$$

$$E^{U} = \frac{1}{2}U\sum_{i\neq j} f_i f_j \tag{46}$$

$$V_i(r) = V^{DFT}(r) + U\left(\frac{1}{2} - f_i\right) \tag{47}$$

where V_i is the orbital-dependent potential. This potential separates the energy of the filled and empty Hubbard-corrected bands by the Coulomb parameter U, which can partially remedy the bandgap problem [194]. It is also repulsive for less than half-filled orbitals ($f_i < 0.5$) and attractive otherwise, which can oppose the delocalization error [195]. The double-counting term E^{dc} compensates for the electron-electron interaction contribution, which is already included in the E^{DFT}.

Extracting the precise XC energy of DFT for the correlated subspace is not trivial [196]. This term is commonly estimated from the fully localized limit [197] or from the mean-field method [191]. In practice, the DFT+U approach partially corrects the SIE and can introduce derivative discontinuities in the total energy [198]. The +U correction has been reported to be successful in many cases for bulk [199, 200] or surface [201, 202] modeling of TiO_2.

Although U values can be calculated from *ab initio* calculations [203, 204], they are usually obtained by fitting physical properties of the material such as bandgap, charge localization, and X-ray photoelectron spectrum to experimental or higher-level calculation results [205]. In most cases, it is not possible to determine a universal value for the U parameter to predict all properties of the material equally well [200]. Therefore, choosing a suitable U parameter is not trivial and depends on the employed density functional and desired property, and any chosen value may have undesired side effects [206-208].

3.1.9.2.3 Non-local external potentials for band-gap correction

Christensen proposed the method of adding an empirical external potential to atomic sites α in the self-consistent LDA calculation for band-gap correction [209]. Based on that, Lany and Zunger developed a method for band gap correction that makes the external potential also angular momentum l –dependent (nonlocal) for more flexibility in fitting experimental band structure properties [210]. The energy correction for this method can be simply expressed as:

$$E^{DFT+NLEP} = E^{DFT} + \sum V_{\alpha,l}^{NLEP} n_{\alpha,l} \tag{48}$$

where $n_{\alpha,l}$ is the occupation (partial charge) of the orbital l on the atom α. In contrast to the potential in LDA+U approach Eq. (47), NLEPs are simple constant external potentials that are treated as free parameters. They are not derived from a model for the electron-electron interaction, and they do not depend on the orbital occupations [68].

This method can be used to correct the band gap problem in materials where the upper valence bands and the lower conduction bands are made almost entirely of the orbitals of the anion and the cation, respectively.

3.1.9.2.4 External potentials for localization correction

To overcome the localization error, Lany and Zunger used non-empirical external potentials to minimize SIE by enforcing the generalized Koopmans' condition, ensuring linear behavior [210, 211]. The hole state (V_h) and electron state (V_e) potentials are in the form of:

$$V_h = \lambda_h \left(1 - \frac{n_m}{n_{host}}\right) \tag{49}$$

$$V_e = \lambda_e \left(\frac{1 - n_m}{1 - n_{host}} - 1\right) \tag{50}$$

where λ_h and λ_e are the potential strength parameters determined through the generalized Koopmans' theorem (gKT), n_m is the fractional occupancy of m^{th} sublevel, and n_{host} is the fractional occupancy of an unperturbed host.

V_h and V_e both vanish when the target orbital's occupancies are not changed with respect to the unperturbed host. Potential strength parameters λ are determined from Koopmans' condition. Therefore, this method does not rely on any empirical parameter for predicting the polaron energies. This method has been shown to be able to describe polaron formation in bulk anatase and rutile [212], but due to relying on the perfect bulk orbital occupancies, this method has not been designed for modeling surface and interfaces.

3.2 Charge correction

To correct the error in the total energy of models due to extra charge, I will use the correction scheme of Komsa and Pasquarello (as outlined in Section 2.5.1.6). A more in-depth discussion on this correction scheme can be found in their original paper [70].

In this work, I will extend this approach to have the possibility of handling more complicated cases, and I will also simplify the process of creating model charge distributions to have a convenient and numerically accurate solution. I will also show the cases where *a posteriori* corrections are not sufficient and a self-consistent charge correction method is necessary. A method for self-consistent potential correction for these models will also be presented.

3.3 Numerical methods

3.3.1 Geometry optimization

In constructing any atomic model for a physical system, the first step requires the optimization of the geometry to find characteristic stationary point structures (e.g., equilibrium geometry, transition state structure, etc.) on the potential energy surface (PES). The equilibrium geometry of a model is considered a minimum on the PES, while the transition state is a 1st-order saddle point. In general, there are two common approaches to the geometry optimization problem:

1. Direct optimization of the PES: In this approach, geometry optimization is performed by repositioning the atoms of the model and evaluating its energy at various points on the PES to estimate the characteristic point of interest. The objective function and the gradients for the optimization algorithms are given by the potential energy and forces, respectively. This approach can be easily applied to the optimization of equilibrium as well as transition states [213].

2. *ab initio* molecular mechanics: In each step, the forces on the atoms are calculated, and the atoms are allowed to evolve (relax) according to Newton's equations of motion under these forces. Although this geometry optimization scheme has been designed for equilibrium geometry optimization (simulated annealing), it can be modified to locate the transition states as well [214, 215].

Direct optimization on the PES is the most commonly used approach for geometry optimization. Some of the most widely used optimization algorithms include:

- Steepest descent: A series of energy evaluations is done in the negative gradient direction. Once the energy starts to increase, an approximate minimum can be found by interpolation between the calculated points on PES [216].
- Conjugate gradient: In each step, a new search direction is chosen such that it lowers the energy while remaining at or near the minimum in the previous search direction [217].
- Newton-Raphson: The PES is approximated by a local quadratic function and its minimum is found in the steepest descent direction. Although the PES is generally non-quadratic, minimization of energy on approximated PES must be done iteratively in order to converge to a minimum [218].
- Direct inversion of the iterative subspace: A new geometry is constructed as a linear combination of previous geometries, which minimizes the size of the Newton step (residuals) [219].

To find the equilibrium geometry, both the size/shape of the supercell and the positions of the atoms inside it must be optimized. A finite plane wave basis set is not complete with respect to changes in the volume, and the stress tensor calculated using the Hellmann-Feynman theorem differs from the 'true' stress. To overcome this problem, I use the Murnaghan equation of state [220], which relates the pressure P to the volume V and bulk modulus at zero pressure K_0:

$$P(V) = \frac{K_0}{K_0^{'}}\left[\left(\frac{V}{V_0}\right)^{-K_0^{'}} - 1\right] \tag{51}$$

K_0' is the derivative of the bulk modulus with respect to pressure at P=0.

In this way, the energy, stress, and pressure are calculated for different volumes using the same energy cutoff, and the results are fitted to the Murnaghan equation of state. Unless noted otherwise, 0.02 eV/Å has been used as the force criterion for structural relaxation in all calculations.

3.3.2 Locating transition states

The study of a reaction mechanism requires finding the minimum energy path (MEP) that connects the reactants to the products via suitable transition states on the PES. All points on the MEP are at an energy minimum in all directions perpendicular to the path. The projection of the MEP onto a coordinate space (reaction coordinate) is commonly called the reaction path [221].[3] To find the geometry of the transition states of a chemical reaction, we must find a geometry corresponding to a point on the PES that satisfies the following criteria [222]:

1. It must be a stationary point.
2. The force constant matrix at that point must have one and only one negative eigenvalue.
3. It must be the highest energy point on a continuous line connecting reactants and products.
4. It must be the lowest energy point that satisfies all three conditions above.

The calculation of equilibrium geometries is a relatively simple process. It only requires minimization of the energy with respect to the ionic degrees of freedom in the system, and the negative of the gradient always points towards the desired solution (local minima of the PES). In contrast, a transition structure optimization should follow the uphill path on the PES in one direction and the downhill path in all other orthogonal directions. More often, the desired uphill direction is not known in advance and must be determined during the optimization. Therefore, various optimization methods have been specifically designed to find the TS. These methods can be broadly classified as single-ended and double-ended [213].

Single-ended methods, e.g., the quasi-Newton method [223], start with an initial guess for the TS geometry and displace it towards the transition structure. Single-ended methods are much like the methods for finding minima of PES, except that they perform a constrained optimization and force the Hessian to contain a single negative eigenvalue. Double-ended methods, e.g., the string method

---

[3] It must be noted that the reaction path is a static path that neglects the kinetic energy of the system and is different from any “real” trajectory of the atoms during a chemical reaction.

[224], the growing string method [225], and the nudged elastic band [226], start with the reactants and products and work from both sides to find the reaction path. The TS can be determined later, after the reaction path has been established.

3.4.2.1 Plain elastic band method

The Plain elastic band (PEB) is a double-ended method of locating the MEP. PEB works by optimizing a number of intermediate geometries (images) along the reaction path. One can simply formulate the search for MEP as the minimization of an objective function S in the form of the energy of N-1 beads connected together with N springs of natural length zero [227]:

$$S(R_1,..,R_{N-1}) = \sum_{i=1}^{N-1} E(R_i) + \sum_{i=1}^{N} \frac{k}{2}(R_i - R_{i-1})^2 \tag{52}$$

where $[R_1, R_2....R_{N-1}]$ represent a string of images between the fixed end points (R_0, R_N) and k is the spring constant. With this minimization, each image finds the lowest energy possible while maintaining equal spacing with neighboring images. However, the elastic bands in this method tend to cut corners and get pulled off the MEP by the spring forces in regions where the MEP is curved, and slide down towards the fixed endpoints [228].

3.4.2.2 Nudged elastic band method

The non-zero component of the spring force perpendicular to the path in the PEB method leads to corner-cutting, while the parallel component of the true force (from the interaction between atoms of the system) leads to the down-sliding problem [226]. These problems can be solved by "nudging", which is projecting out the perpendicular component of the spring forces and the parallel component of the true forces. The nudged elastic band (NEB) was first introduced by Mills and Jónsson in a study of H_2 adsorption on the Cu(110) surface [229] and was later described in [226]. If we have an estimate of the unit tangent to the path at each image $\hat{\tau}_i$, we can redefine the force on each image as the parallel component of the spring force $F_i^{||}$ plus the perpendicular component of the true force $F_i^{\perp}$ as:

$$F_i = F_i^{\perp} + F_i^{||} \tag{53}$$

$$F_i^{\perp} = -\nabla E(R_i) + \nabla E(R_i).\hat{\tau}_i \tag{54}$$

$$F_i^{||} = [k_{i+1}(R_{i+1} - R_i) - k_i(R_i - R_{i-1})].\hat{\tau}_i \tag{55}$$

where $\nabla E(R_i)$ is the gradient of the energy with respect to the atomic coordinates (R) of the image i. The perpendicular component of the energy gradient is obtained by subtracting out its parallel component.

To estimate the tangent, the simplest method is to use the normalized line segment between the two images:

$$\hat{\tau}_i = \frac{R_{i+1} - R_{i-1}}{|R_{i+1} - R_{i-1}|} \tag{56}$$

Or it can be better estimated by considering the relative energy of each image and its neighbors in the string [230]. It has been shown that in the limit of many images, the NEB will always converge to a MEP [231], but its efficiency highly depends on the employed optimization algorithms [232], whose implementation in general is not trivial [231, 233].

3.4.2.3 Climbing image nudged elastic band method

The climbing image nudged elastic band (CI-NEB) [234] is a modification to the NEB method in which the highest energy image is driven up to the saddle point. This image does not feel the spring forces along the band ($F_i^{||}$) which are artificial forces used to maintain equal separations between the neighboring images. Instead, the true force in this image along the tangent is inverted. In this way, the image tries to maximize its energy along the band and minimize it in all other directions. When this image converges, it will be at the exact saddle point (transition state).

$$F_{i,max} = -\nabla E(R_{i,max}) + 2\nabla E(R_{i,max}).\,\hat{\tau}_{i,max} \tag{57}$$

To locate the transition states for a reaction, I use the NEB starting from a chain of images interpolated between the initial and final positions of the atoms in the reaction. The geometry of the start and end points of the reaction path is fixed, and the computational units are divided equally among the intermediate images to optimize their geometries and find the MEP. The CI-NEB method chooses the geometry with the maximum energy and tries to maximize its energy along the MEP. To verify that the identified transition geometries are first-order saddle points, I performed vibrational analysis using the finite differences method [235] to ensure the existence of one and only one imaginary frequency. Additionally, I optimized the ionic structure of the obtained transition states using the molecular dynamics method with maximum damping (which is effectively just a simple steepest descent method) to ensure that the located transition states are actually connecting the intended local minima on the PES.

3.4 Calculation parameters

For all my DFT calculations, I use the Vienna Ab initio Simulation Package (VASP) ver. 5.4.4, which is a very popular plane-wave based package for the electronic structure calculations in solid-state physics [236, 237] while being equally successful in chemical applications [238]. It is currently proprietary software but is distributed as source code.

VASP has been written in Fortran and internally parallelized using the MPI (message passing interface) scheme. An OpenMP multithreaded version is also currently under development [239]. VASP offers multiple levels of data parallelization, including the parallelization over the k-points, Kohn-Sham orbitals at each k-point, and the basis set coefficients [240]. Additionally, it supports the use of GPU accelerators [241, 242] and has the flexibility to be tweaked for high-performance computations on various hardware/cluster architectures [243].

I used the PAW potentials supplied with the VASP package [127] for my calculations. The Ti3p states were explicitly included in the calculations. The default cutoff value (ENMAX), and number of valence electrons for the potentials used in my calculations are listed in Table 2. H.66 and H1.33 are pseudo-hydrogens used to passivate the dangling bonds of the bottom surface atoms and to keep the states at the bottom surface localized [244].

Table 2 Default plane-wave cutoff energy and the number of valence electrons for the PAW potentials used in this study

potential	default cutoff value (eV)	valence electrons
Ti_pv	222	10
O	400	6
H.66	250	2/3
H1.33	250	4/3
C	400	4

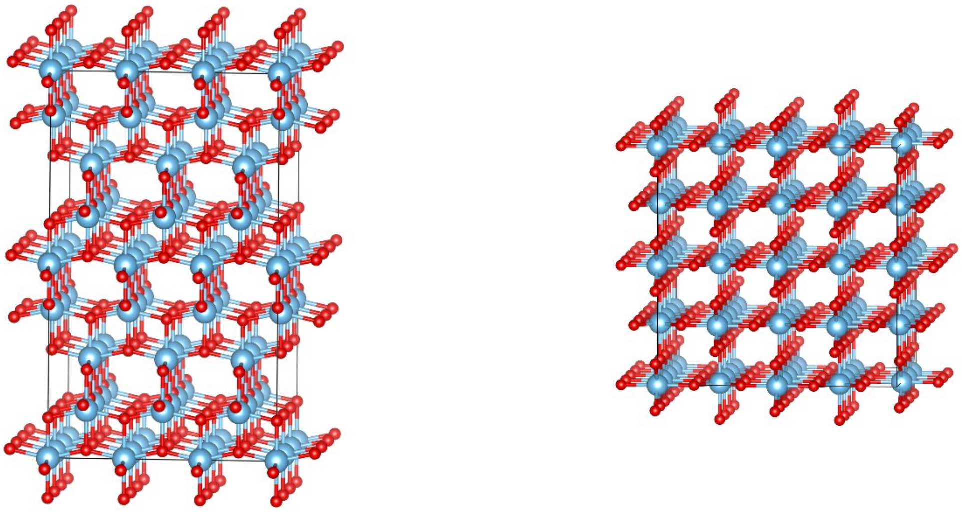
Unless otherwise stated, all the calculations have been performed using a cutoff energy for the planewave basis set (ENCUT) of 420 eV, and a cutoff energy for the plane wave representation of the augmentation charges (ENAUG) of 840 eV.

I used an 8×8×8 Γ-centered Monkhorst-Pack grid for Brillouin zone integration in unit cell calculations and only the Γ point for slab model calculations. No constraint has been put on the charge density symmetry, and all calculations were spin-polarized. Three-dimensional periodic boundary conditions were automatically applied to all the models. I used a patch provided by G. Kresse to modify the VASP code and increase its numerical accuracy in dealing with the charge of pseudo-hydrogen atoms (2/3 and 4/3) to 7 digits after the decimal separator.

I used the PBE functional, which is a non-empirical GGA functional, in all my calculations [245]. As mentioned earlier, GGA-based functionals suffer from bandgap underestimation and charge delocalization errors. To overcome these limitations of the PBE functional, I used Lany and Zunger's gap correction method [210] and polaron correction [246] for the Ti and O atoms in my calculations.

I modified the VASP code to include support for the atom type- and angular momentum-dependent semi-empirical NLEPs, as described in [210], with a patch provided by S. Lany. This patch modifies the LDA+U routines in the VASP [247] to include both gap and localization corrections.

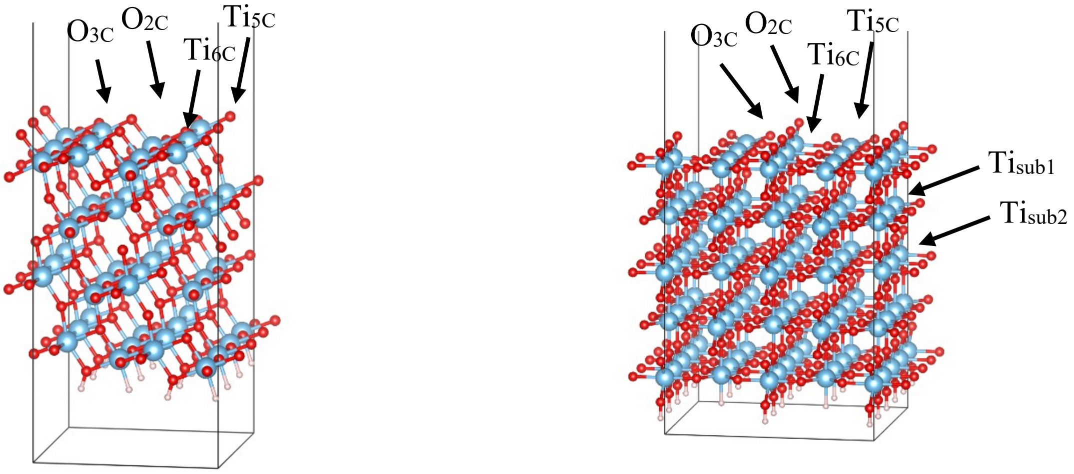
To model the bulk materials, I created supercells consisting of 3×3×2 Bravais cells for anatase (216 atoms) and $2\sqrt{2}\times2\sqrt{2}\times4$ for rutile (192 atoms). These models are shown in Figure 7.



a **b**

Fig. 7 Ball-and-stick model representation of the bulk models for a) anatase, and b) rutile phase

The anatase (101) surface was modeled by a slab of 4 double layers, with a (3 × 1) surface unit (144 atoms), and the rutile (110) surface was modeled by a slab of 5 layers with a (4 × 2) surface unit (240 atoms). In both slab models, the two bottom layers were fixed to the optimized bulk positions, and the fivefold-coordinated titanium atoms (Ti_{5C}) and the twofold-coordinated oxygen atoms (O_{2C}) at the bottom were saturated by $H^{1.33}$ and $H^{0.67}$ pseudo-hydrogen atoms, respectively. A vacuum layer of 10 Å was added to the surface. The relaxed slab models are shown in Figure 8.



a **b**

Fig. 8 The structures of a) the anatase (101), and b) the rutile (110) slab models. Five- and sixfold-coordinated titanium (blue), and two- and threefold-coordinated oxygen surface atoms (red) are labeled as Ti_{5C}, Ti_{6C}, O_{2C}, *and* O_{3C}, *respectively. Pseudo-hydrogen atoms are white.*

It should be noted that these slabs have a small dipole moment orthogonal to the surface, 2 Debye for anatase and 7 Debye for rutile. Correcting for the error arising from the periodic repetition shifts the total energy of the perfect slabs by -0.005 eV and -0.051 eV, respectively. Since we cannot presently calculate the dipole correction for charged slabs in VASP, I neglect dipole corrections altogether.

For locating the transition state of reactions, I used the CI-NEB method as implemented in the VASP Transition State Theory (VTST) tools patch [234] ver. 3.1.[4] This patch also contains an improved tangent definition [230] as well as several optimizers that are solely based on the force (and not the energy), including a modified conjugate-gradient, limited-memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) [248], and Fast Inertial Relaxation Engine (FIRE) [249]. VTST also replaces the damped molecular dynamics optimization method in VASP with its own optimizer routines.

---

[4] http://theory.cm.utexas.edu/vtsttools

Chapter 4

Results

4.1 *A posteriori* charge correction

Before my work, there was no user-friendly implementation available of the correction method by Komsa and Pasquarello. I have identified the following issues in their scheme that needed to be addressed to obtain consistent and reproducible results:

The model of the extra charge in the Komsa-Pasquarello (KP) method uses a single Gaussian charge distribution, which may not be a suitable choice for modeling extra charge localization on multiple sites. As an example, Figure 9 shows the localization of the extra positive charge in an anatase (101) surface model with a subsurface oxygen vacancy.

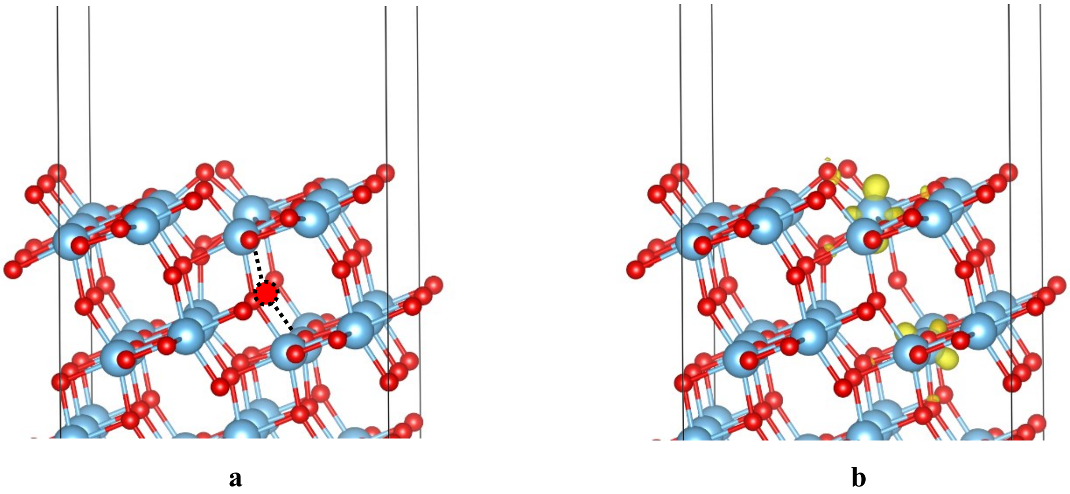


Fig. 9 a) Position of the missing subsurface oxygen atom and b) localization of the extra positive charge on the titanium atoms next to a subsurface oxygen vacancy in the anatase (101) surface model

Additionally, to construct a model that resembles the extra charge distribution and its environment in the electronic structure calculations, several parameters need to be defined, i.e., the position of the charge, the width of the Gaussian distributions, and the position of the interfaces. The choice of these parameters can heavily influence the accuracy of the calculated energy correction values. The process of obtaining the model parameters is usually quite involved, and in the case of large slab models, it can be very time-consuming. Judging the quality of the constructed model is also rather complicated.

To have the possibility of calculating energy corrections for more complex models and routinely integrating it into my workflow, I extended the KP method and implemented it in a standalone code called SLABCC (slab charge correction) [250]. The details of SLABCC's implementation are available in Appendix 1.

4.1.1 Modeling the charge distribution

First, we need to construct a model charge that generates a potential closely mimicking the potential due to the extra charge in the electronic structure calculation. Assuming our model charge distribution is described as the sum of Gaussian charges, we will have:

$$\rho_{\text{model}}(r) = \sum_i \frac{q_i}{\sigma_i^3 (2\pi)^{3/2}} \exp\left(-\frac{(r-r_i)^2}{2\sigma_i^2}\right) \tag{58}$$

$$Q_{\text{model}} = \sum_i q_i \tag{59}$$

where q_i is the charge localized in each Gaussian, σ_i is the width of each Gaussian charge, and r_i is the position of the Gaussian charge centers.

Additionally, in the calculation of the charge correction for slab models, the dielectric profile of our model will influence the potential generated by our model charge distribution. We can assume a dielectric profile in the form of Eq. (60) in the direction normal to the slab:

$$\varepsilon(z) = \frac{\varepsilon_2 + \varepsilon_1}{2} + \frac{\varepsilon_2 - \varepsilon_1}{2} \operatorname{erf}(\frac{z - z_0}{\beta}) \tag{60}$$

where z_0 is the interface position in the z-direction, ε_1 and ε_2 are dielectric tensors on either side of the interface, and β controls the smoothness of the transition.

In the KP method, the potential created by the model charge (V_{model}) is calculated by solving the Poisson equation, Eq. (11) for the model charge embedded in the model dielectric medium. In materials with large anisotropy in their dielectric tensor, using a single ε value may not be adequate to construct a proper model. SLABCC solves this problem by explicitly including the dielectric tensor as:

$$\varepsilon = \begin{vmatrix} \varepsilon_{11} & 0 & 0 \\ 0 & \varepsilon_{22} & 0 \\ 0 & 0 & \varepsilon_{33} \end{vmatrix} \tag{61}$$

In the case of slab models where the dielectric profile varies only in one direction (z), Eq. (11) can be written as:

$$\frac{\partial}{\partial z}\varepsilon_{33}(z)\frac{\partial}{\partial z}V_{\text{model}}(r) + \varepsilon(z)\nabla^2 V_{\text{model}}(r) = -\rho_{\text{model}}(r) \tag{62}$$

This discrete Poisson equation can be efficiently solved in Fourier space. I discretize the analytical model charge distribution on the same spatial grid as that given for the charge distribution by the electronic structure calculation.

4.1.2 Choosing optimal model parameters

In general, the mentioned model parameters, i.e., the positions of the center of Gaussian charges, the fraction of total charge in each Gaussian, the width of the Gaussian distributions, and the position of the interfaces in the model, are all expected to be manually defined by the user. The process of obtaining these parameters is usually quite involved, and evaluating the quality of the constructed model is rather time-consuming. These drawbacks can severely limit the practicality of a total energy correction code. To overcome these limitations, in SLABCC, I try to estimate model parameters by automatically constructing a model that generates a potential mimicking the potential generated by the extra charge of our electronic structure calculations (V_{calc}). I use the mean squared error (MSE) of my model potential as an index for the goodness-of-fit of the model:

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left(V_{\text{calc},\,i} - V_{\text{model},\,i}\right)^2 \tag{64}$$

where N is the total number of grid points and V_i is the value of the potential on each grid point.

Ideally, our model charge distribution should produce the same potential as the electronic structure calculation at each grid point. This corresponds to MSE=0. Therefore, our goal is to minimize the MSE by changing the model parameters. The analytical calculation of the MSE derivative with respect to each of the parameters is not trivial. Also, numerical approximation of derivatives using the finite difference method, given the computational cost of MSE evaluation, is far too expensive. Consequently, using a derivative-free (direct) search optimization method to find the minimum of the MSE from the values calculated at discrete points in n-dimensional space can provide an efficient solution. An extensive review of the commonly used derivative-free optimization algorithms can be found in [251]. Technical details of the model parameter optimization in SLABCC are available in Appendix 1.

4.1.3 Energy calculations for periodic and isolated model charge

After constructing our model charge, we can easily calculate the electrostatic energy for our charge distribution $\rho_{model}(r)$ using its electrostatic potential V_{model} from Eq. (10). For calculating the energy of our model charge distribution in the isolated limit, I uniformly scale the model charge medium (slab) to larger sizes, embedding the model charge in larger slabs, and calculate the energy of this new model under periodic boundary conditions. Finally, I extrapolate the obtained energies to the limit of the model charge embedded in an infinitely large medium (see, for example, Figure 12).

4.1.3.1 Bulk models

Although SLABCC was originally designed for calculating the energy correction for slab models, it can be easily adapted to calculate the energy correction for bulk models as well [252]. Defining the same value for the dielectric tensor of the regions inside and outside of the “slab” generates a charge model embedded in a uniform (bulk) medium. When generating charge models in bulk, SLABCC automatically turns off the optimization of the interface positions. This leads to faster convergence due to the smaller number of model parameters involved in the optimization process. Considering that the uniform scaling of the model charge does not change the ratio of the extra charge in the region of space with different dielectric constants, the extrapolation scheme used for the slab models can also be used for calculating the $E_{isolated}$ of the bulk models.

4.1.3.2 Monolayer and 2D models

I have also added the possibility of modeling the extra charge with trivariate Gaussian charge distributions as follows:

$$\rho_{\text{model}}(r) = \sum_i \frac{q_i}{\sigma_{i,x}\sigma_{i,y}\sigma_{i,z}(2\pi)^{3/2}} \exp\left(-\frac{\left(r_x - r_{i,x}\right)^2}{2\sigma_{i,x}^2} - \frac{\left(r_y - r_{i,y}\right)^2}{2\sigma_{i,y}^2} - \frac{\left(r_z - r_{i,z}\right)^2}{2\sigma_{i,z}^2}\right) \quad (65)$$

where $\sigma_{i,x/y/z}$ refer to the Gaussian width in each Cartesian direction, and $r_{i,x/y/z}$ are elements of the charge center position vectors in the x/y/z directions. This provides us with the possibility of modeling more complex charge distributions, which is especially important when the extra charge is localized near the surface or for models of the extra charge in layered or 2D models [253].

The linear interpolation of the energy of a model charge distribution in the isolated limit in the case of uniformly scaled models may not be suitable for monolayer/2D models since the asymptotic form is generally unknown [84], and when scaling to large sizes, this relation may not be trivial [254]. Noh et al. have proposed using a fifth-order polynomial for fitting [255] for these cases. However, in SLABCC, I have implemented an energy extrapolation method using a second-order polynomial with an exponential term, as proposed in [85]:

$$E = c_0 + c_1 x + c_2 x^2 + d e^{-c_3 x} \quad (66)$$

$$d = \frac{c_1 - \frac{\partial E_M}{\partial x}}{c_3} \quad (67)$$

where c_i are the fitting parameters, E_M is the Madelung energy, and x is the scaling factor ($x = \frac{1}{\alpha}$; α being the extrapolation ratio). Inclusion of the *d* term in this fitting method guarantees that the energy gradient is similar to the Madelung energy at the limit of isolated charge $\alpha \to 0$.

4.1.4 Potential alignment term

The alignment term ΔV in Eq. (9) must be calculated at the position least affected by the model charge. In the case of a single localized Gaussian charge in an isotropic dielectric medium, this point is the farthest point from the charge center. As a more general solution, to calculate the ΔV I use the extremum point of V_{model} in the 3D Cartesian space to also account for the effects of the dielectric medium anisotropy. In most cases, if the model charge is describing the potential well enough, this term should be negligible.

4.1.5 Validation

For validating my implementation of the KP method, I calculated the energy correction values for the formation energies and EA of a Cl vacancy in the +1 charge state on the surface of a 3×3×2 NaCl slab (16.98×16.98×11.32 Å) with various vacuum thicknesses along the third direction. This is one of the models used by Komsa and Pasquarello in their original paper [70]. The localization of the extra (positive) charge around the surface vacancy in a model with a vacuum thickness of 6×lateral size is shown in Figure 10.

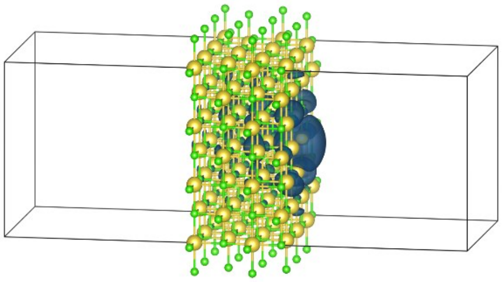
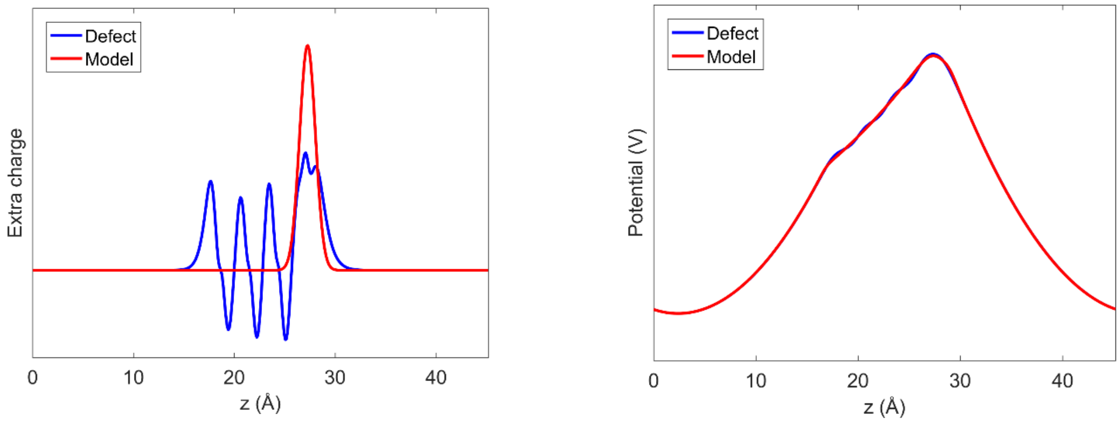


Fig. 10 Localization of the extra charge on the surface of NaCl slab model

I calculated the total energy correction using the SLABCC code, assuming a dielectric constant (ε=2.45). The planar average of the extra charge distribution and its potential for the VASP calculations and the SLABCC's model are shown in Figure 11. As can be seen, the potential of the charge model closely follows the potential of the extra charge in the VASP calculations.



a **b**

Fig. 11 Planar average of the a) charge distribution and b) resulting potential of the charged defect in VASP calculations and the SLABCC model

For calculating the isolated energy of the charged defect (when the defect is embedded in an infinitely large medium), I use the extrapolation method. In each step, the Gaussian charge is embedded in a larger medium, and the total energy of the resulting model under periodic boundary conditions is calculated. In the limit of an infinitely large cell, the energy of the model will correspond to the energy of the isolated charge in the dielectric environment. Figure 12 shows the extrapolation scheme and the estimated $E_{isolated}$ for this model. α is the scaling factor, which is defined as the ratio of the extended supercell's side to that of the original supercell.

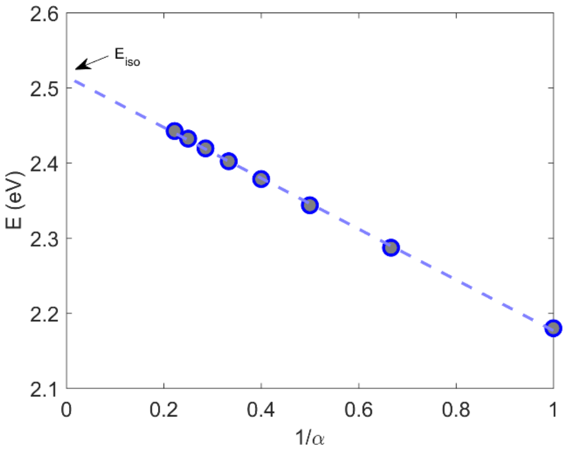


Fig. 12 Electrostatic energies of the model charges as obtained by uniformly scaling all dimensions in the supercell and their extrapolation to an infinitely large supercell.

Figure 13 shows the formation energy and EA of NaCl defect slabs with various vacuum thicknesses. As can be seen, by adding the energy correction values estimated by the SLABCC to the total energies, the resulting values are independent of the vacuum size of the model.

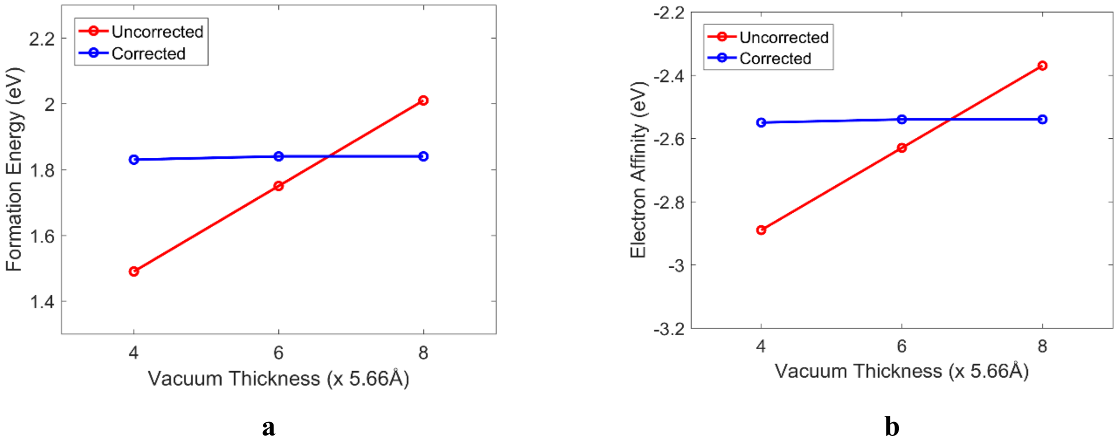
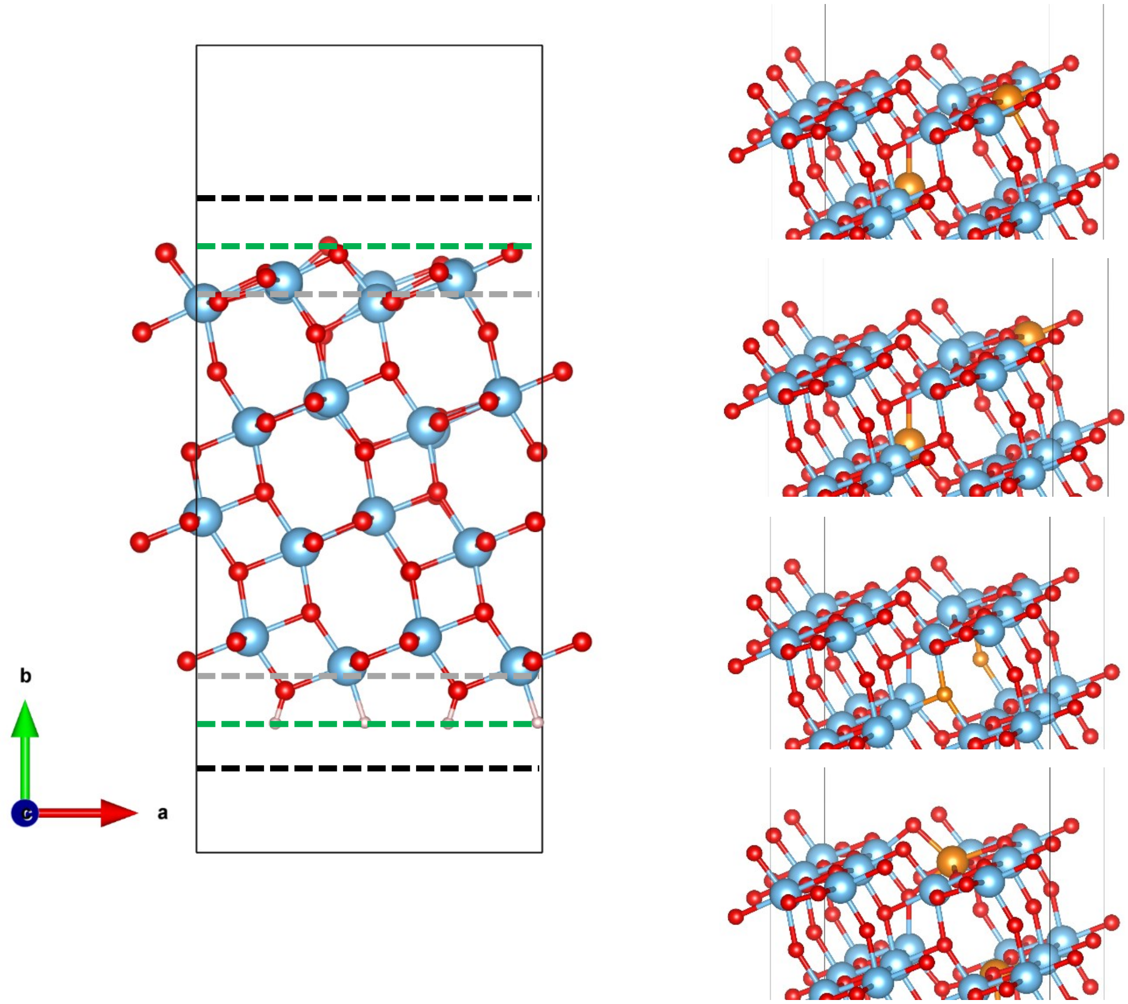


Fig. 13 Corrected and uncorrected a) formation energy and b) EA, of positively charged Cl vacancies on the NaCl surface for models with a vacuum thickness of 4-8 × lateral size

For additional verification of my implementation, I used SLABCC in the calculation of the vertical ionization energy for a sub-surface oxygen vacancy in an anatase (101) surface model. A slab of 4 double layers, with a (3 × 1) surface unit (144 atoms) and various vacuum sizes (10-30 Å) was used to model the surface. The location of the missing oxygen atom on the surface and the difference in the total charge distribution of the neutral and positively charged models are shown in Figure 9. The dielectric constant of anatase, required for the total energy correction, was obtained from density functional perturbation theory [256] as 4.79. The details of my slab model and calculations have been explained in Section 3.2.3.

Initial guesses of the SLABCC model charge parameters are graphically presented in Figure 14. I used three different sets of interface positions (shown by colored dashed lines) and four different sets of Gaussian charge positions (marked by the orange color) as the initial guesses for our charge models. The dielectric tensor is assumed to be isotropic inside the slab; therefore, only a single value (the geometric average of the tensor elements) is provided.



a **b**

Fig. 14 Initial guesses for a) the position of the interfaces in the direction normal to the surface, and b) the position of the center of Gaussian model charges

The final model charges generated by the SLABCC starting from these initial parameters and the calculated correction energies are all identical. The final positions of the Gaussian model charges and the interfaces are shown in Figure 15.

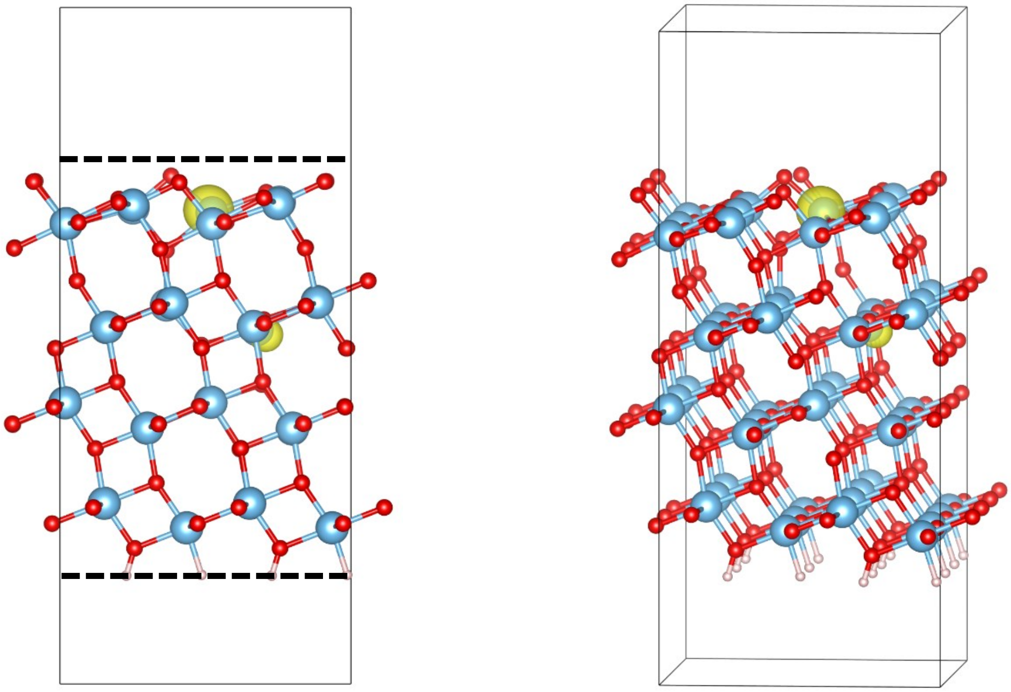


Fig. 15 The localization of the constructed model charge and the position of the model slab interfaces

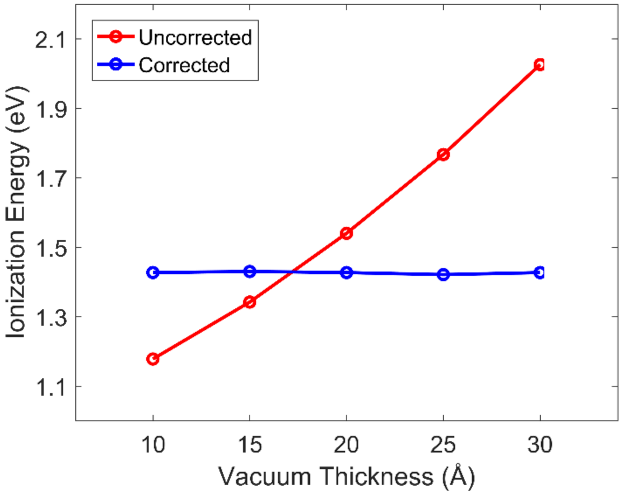


Fig. 16 Vertical ionization energies of the slab models with various vacuum thicknesses with, and without the energy correction for the extra charge

Figure 16 shows the variations of the vertical ionization energies (the energy of the neutral system subtracted from the energy of the positively charged system) for the models with various vacuum thicknesses. As can be seen, adding the correction obtained from the SLABCC to the total energy of the charged models makes them independent of their size. The contents of a sample SLABCC input file for the energy correction calculation of this model are shown in Appendix 1.

For verification of my SLABCC implementation for monolayer models, I calculated the vertical ionization energy of a negatively charged h-BN model with a carbon substitution on the nitrogen site. A 128-atom model with ideal h-BN geometry (a = 2.49Å) was used as the reference for the model. I used the HSE06 hybrid functional [257] in these electronic structure calculations. I used 6.28 for the in-plane dielectric constant and 1.83 for the dielectric constant in the direction normal to the surface. The localization of the extra charge is shown in Figure 17.

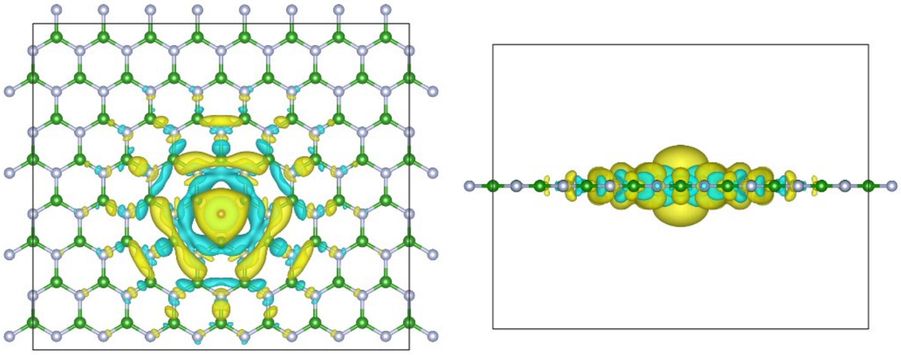


Fig. 17 Localization of extra negative charge around the substitution carbon in the h-BN monolayer model

As mentioned earlier, linear interpolation of the energy of uniformly scaled models is not suitable for the calculation of isolated model charge energy. As it can be seen in Figure 18, the energy of model charges deviates from the linear trend, especially in large scaling factors. The energy of the model charge distribution in the isolated limit for these 2D models has been estimated using a second-order polynomial with an exponential term [Eqs. (66,67)].

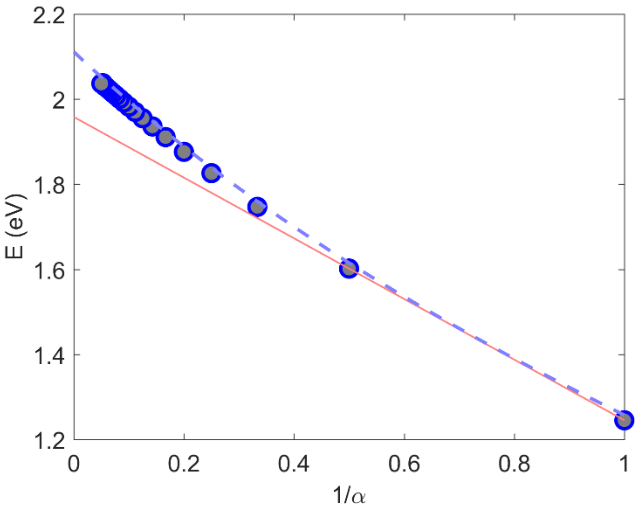


Fig. 18 Electrostatic energies of the model charges as obtained by uniformly scaling all dimensions in the supercell and their extrapolation to an infinitely large supercell. The red line represents a linear fitting and the dashed line shows the results of fitting data to Eq. (66)

I also calculated the variations of the vertical ionization energies for h-BN monolayer models with various vacuum thicknesses. As can be seen in Figure 19, adding the correction obtained from the SLABCC to the total energy of the charged models makes them independent of their size. As the sign of energy correction is positive for all of these cases, the crossing of the corrected and uncorrected energies occurs at a very large vacuum thickness (~60 Å).

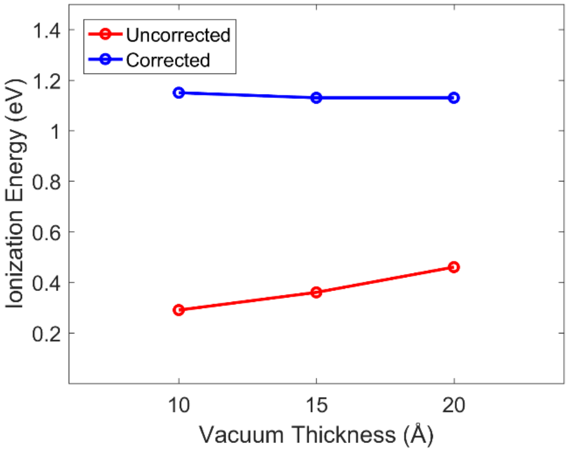


Fig. 19 Vertical ionization energies of the monolayer models with various vacuum thicknesses with, and without the energy correction for the extra charge

It should be emphasized that the energy values for these tests are only meant to test the reliability of the charge correction scheme as implemented in SLABCC. It has been shown that the HSE-type hybrid functional applied to monolayer h-BN models cannot be Koopmans-compliant while reproducing the band gap accurately [253]. More examples, as well as a detailed description of all the SLABCC input parameters can be found in its manual.

4.1.6 Limitations and future outlook

Currently, the execution of the SLABCC is limited to a single node, where all the processes have access to the same data in memory. To leverage the massively parallel computation environments, the code needs to be parallelized using MPI or a similar method to run on modern clusters. In order to fully benefit from the large number of computational cores, we need to use a parallel optimization algorithm in the SLABCC to evaluate the MSE of several models in parallel. This feature is currently under development in the NLOPT library.

As the optimizers currently used in the SLABCC are all local optimizers, the initial values provided by the user need to be reasonable for convergence to the correct solution. By using a global optimizer algorithm instead, we can eliminate the need for an initial guess and make the process more user-friendly.

Only models with orthogonal cells are currently supported by SLABCC. Additionally, in the current version, only the VASP data formats are supported, but this can be easily extended to include more file formats.

Finally, it should be noted that *a posteriori* charge correction methods in general can only compensate for the error in the total energy of the model due to the presence of a localized extra charge under periodic boundary conditions. However, the charge also affects the localized one-electron levels, and the corresponding artifacts may have a significant impact on the results [258]. A self-consistent potential correction should be used instead to eliminate the effects of the extra charge in these cases [259].

4.1.7 Summary

I implemented the method proposed by Komsa and Pasquarello for *a posteriori* charge correction of slab models under PBC in a standalone code called SLABCC. I extended the method to model the extra charge of QM calculations with a sum of Gaussian charges, which is necessary in models where the extra charge is localized on multiple atomic sites. I have also included the anisotropy of the dielectric tensor in the calculations, which is crucial for charge correction in monolayer and 2D models. To have an accurate representation of the extra charge, all the model parameters provided by the user needed to be manually adjusted and verified. To make the code more practical and user-friendly, I have added an optimization routine to SLABCC that automatically finds the best model parameters.

The code is freely accessible under a permissive license. The details of SLABCC and its features are published in the Computer Physics Communications journal [250]. We have also used SLABCC and verified its functionality with bulk, slab, and 2D models in our publications [252, 253, 259].

4.2 Self-consistent charge correction

A posteriori charge correction methods can only compensate for the error in the total energy of the model due to the presence of a localized extra charge under PBC. However, the extra charge and its compensating jellium also introduce artifacts in the wavefunction due to incorrect electrostatic potential during the SCF iterations. This also affects the localized one-electron levels, and depending on the model, the corresponding artifacts may have a significant impact on the obtained results [258]. A common symptom of this issue is the appearance of unphysical "ghost states", which are additional states at energies below or close to the physical valence states. In the case of slab models, these states can be fully or partially localized in the vacuum region. I encountered this during my investigation of reactions on the anatase surface (see Figure 22). In such cases, a self-consistent correction scheme is critically needed to eliminate the spurious effects of the extra charge in models with PBC. To address this issue, I contributed to the development of a self-consistent potential correction (SCPC) method for periodic charged models and verified its functionality based on the *a posteriori* corrections with SLABCC [259].

The SCPC method adds a corrective potential term to the KS potential Eq. (23) that is defined as the difference between the electrostatic potential due to the extra charge under PBC and the potential for the same model charge in isolation. As this correction needs to be updated in each iteration during the SCF cycle, it is crucial to keep the required computational cost as low as possible. SCPC can be added to VASP 5.4.4 using a patch, and it is also part of the official VASP package from version 6.2 onwards[5].

4.2.1 Charge model construction

Unlike SLABCC, in which the extra charge is constructed from a sum of Gaussian functions, SCPC uses the difference between the total electronic density of the charged system (ρ_{charged}) and the neutralized defect model ($\rho_{\text{reference}}$) as its model charge instead. This procedure omits the need for model fitting in each iteration but may introduce some errors in the final results.

$$\rho_{\text{model}}(r) = \rho_{\text{charged}}(r) - \rho_{\text{reference}}(r) \qquad (68)$$

---

[5] https://github.com/aradi/SCPC-Method

Technically, the neutralized defect model with the same geometry as the charged one should be used as the reference. However, in practice, it is more convenient to use a pristine model as the reference system instead. By using the unperturbed supercell, obtaining the reference does not require additional calculations, and it does not need to be updated at each geometry optimization step.

4.2.2 Potential correction

In SCPC, the periodic electrostatic potential (V_{per}) corresponding to the model charge is obtained from the constructed model of extra charge by solving the Poisson equation. This equation is solved using the DL_MG library [260]. The difference between the periodic electrostatic potential of the charged defect model ($V_{charged}$) and that of the reference system ($V_{reference}$) on the supercell edges can be used to define the Dirichlet boundary conditions. The dielectric profile of the material (ε) is set to a homogeneous constant for bulk models and a smoothed boxcar function for slab models. The positions of the slab surfaces should be provided by the users.

The potential of an isolated charge, V_{iso} is calculated using the self-consistent iterative procedure of Fisicaro *et al.* [261], which incorporates the macroscopic dielectric profile of the material in the calculation with open-boundary conditions:

$$\nabla^2 V_{iso}(\boldsymbol{r}) = -\left[\frac{\delta\rho(\boldsymbol{r})}{\varepsilon(\boldsymbol{r})} + \rho^{iter}(\boldsymbol{r})\right] \tag{69}$$

$$\rho^{iter}(\boldsymbol{r}) = \nabla ln\varepsilon(\boldsymbol{r}) \cdot \nabla V_{iso}(\boldsymbol{r}) \tag{70}$$

Here, ρ_{iter} results from the spatial variation of the dielectric function. The numerical solution of V_{iso} is obtained using the PSPFFT library [262] in SCPC.

Finally, from V_{per} and V_{iso} we can calculate the corrective potential V_{cor}, which is then added to the total electronic potential.

$$V_{cor} = V_{iso} - V_{per} \tag{71}$$

Alternatively, one can obtain V_{cor} by solving the Poisson equation for the compensating jellium background, using the difference between V_{iso} and V_{per} at the edges of the supercell to determine the Dirichlet boundary conditions. Using this approach, the V_{cor} can be calculated on a coarser grid in comparison to the iterative solution to Eqs. (69,70).

4.2.3 Validation

For testing the SCPC method, we checked the variation of the formation energy with supercell size for a positive chlorine vacancy, V_{Cl}^{+} on the surface of a NaCl (001) slab. For this calculation, the slab model was constructed from a 3 × 3 × 3 (× 5.64 Å) solid model with various vacuum thicknesses added to it. A kinetic energy cutoff of 262.5 (356.2) eV was used for expanding the wave functions (charge density). Spin-polarized calculations were carried out using the Γ-point approximation. The value of 2.45 was used for the macroscopic dielectric profile of the solid region.

Figure 20 shows the change in formation energy of a positively charged Cl vacancy on the NaCl surface for models with different vacuum thicknesses, with and without corrections. As can be seen, despite SLABCC and SCPC being conceptually different, they are both capable of correcting the formation energies and making them model-size independent.

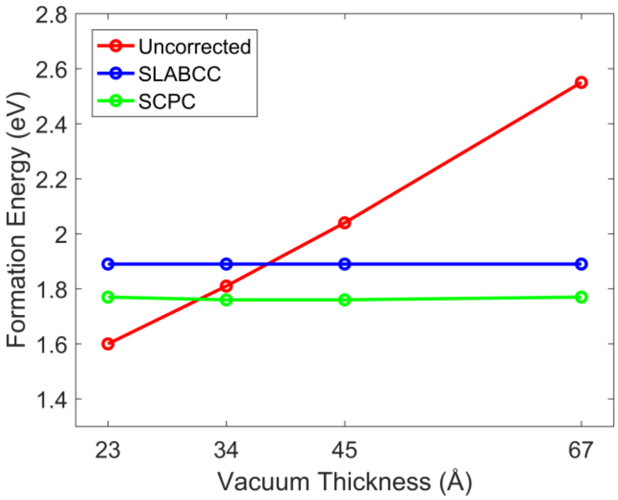


Fig. 20 Formation energy of positively charged Cl vacancies on the NaCl surface for models with different vacuum thicknesses, without correction, with SLABCC, and with SCPC.

Additionally, I checked the effect of SCPC on “ghost states” due to the countercharge in the vacuum part of a negatively charged anatase (101) slab model with an adsorbed O_2 molecule on the surface. In this model, the solid part consisted of three double-layers with a lateral size of 1×3, using the experimental lattice parameters. A kinetic energy cutoff of 420 eV (840) eV was applied to expand the wave functions (charge density). Spin-polarized calculations were carried out using the Γ-point approximation. For the macroscopic dielectric constant, an average value of 6.68 was used. I used the optimized position of slab surfaces from SLABCC calculations to generate the dielectric profile for these models.

Figure 21 shows the planar average of electrostatic potential (Hartree and ionic contribution) in the direction perpendicular to slab surfaces with ~10/30/50 Å of vacuum between the slabs without SCPC. As can be seen, there is a spurious potential well in the vacuum between the slabs, which attracts charge spill-out. This potential may give rise to "ghost states" (two-dimensional Rydberg states) in the middle of the vacuum, and in a plane-wave calculation, these states may get erroneously occupied. To minimize the effects of periodically repeated slabs in the study of surface reactions, the distance between them must be sufficiently large. Nonetheless, as can be seen, increasing the vacuum thickness between the slabs increases the induced dipole moment and deepens this potential well.

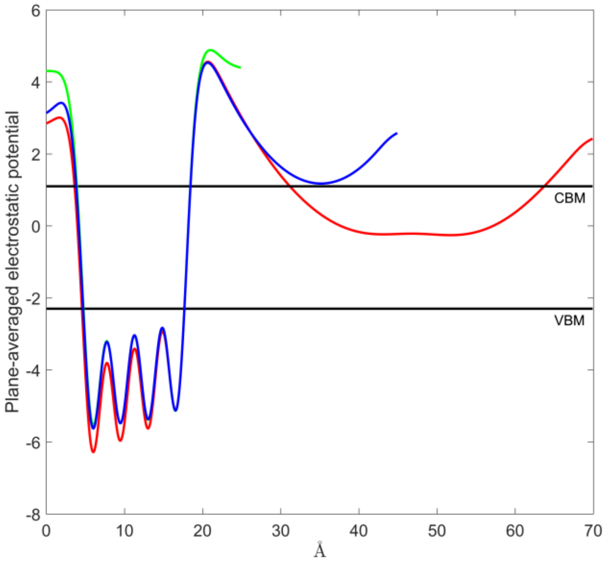


Fig. 21 Variations of the plane-averaged electrostatic potential in the direction perpendicular to the surface for negatively charged anatase (101) slab models with an O_2 molecule on the surface

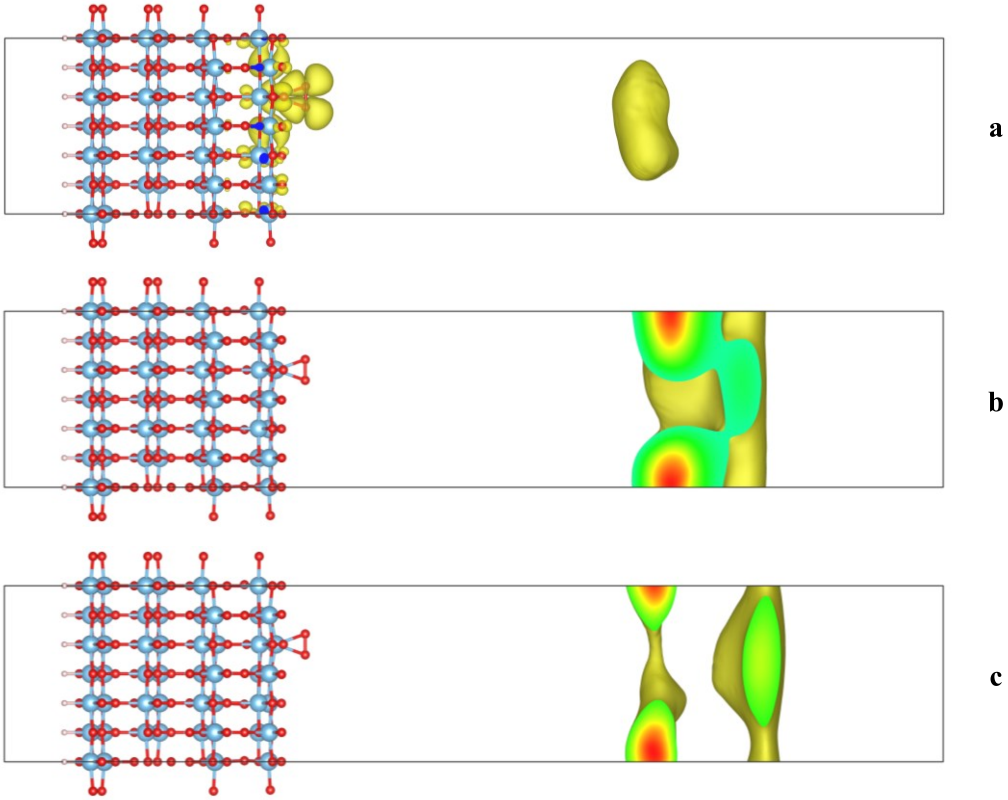


Fig. 22 Localization of a) the highest occupied and b,c) lowest unoccupied states, in [TiO$_2$:O$_2$]$^-$ model without SCPC correction (isosurface level 0.0002)

Figures 22,23 show the spatial localization of the highest occupied state and the two lowest unoccupied ones without, and with SCPC correction, respectively. As can be seen, without the self-consistent correction, all these states are either partially or fully localized in the vacuum between the slabs. By including SCPC in the calculations, these ghost states disappear, and the occupied state is also fully localized on the slab itself.

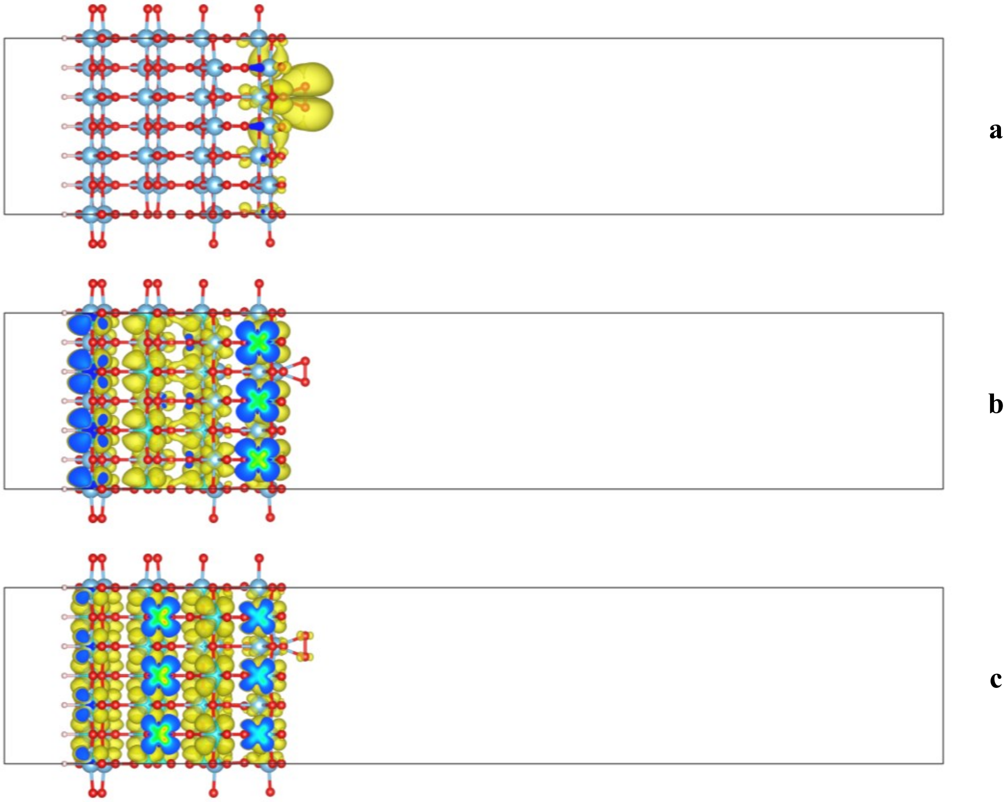
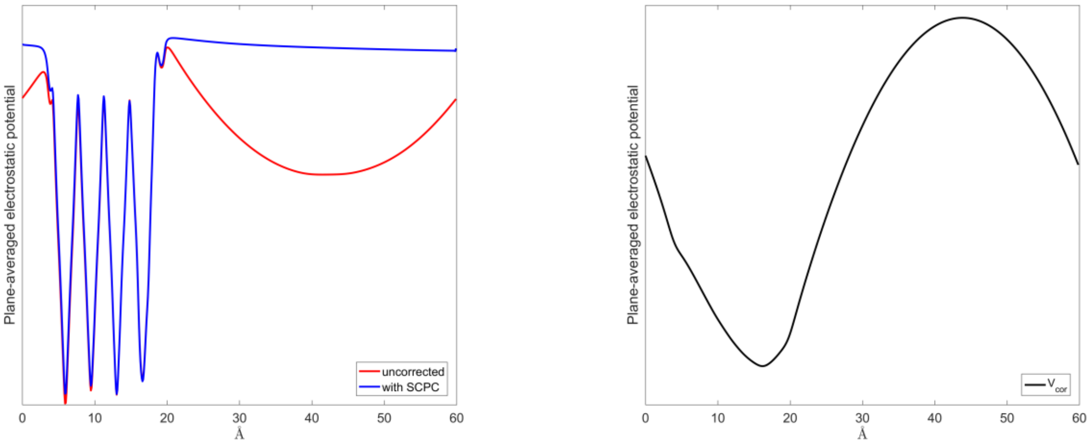


Fig. 23 Localization of a) the highest occupied and b,c) lowest unoccupied states, in $[TiO_2{:}O_2]^-$ *model with SCPC correction (isosurface level 0.0002)*

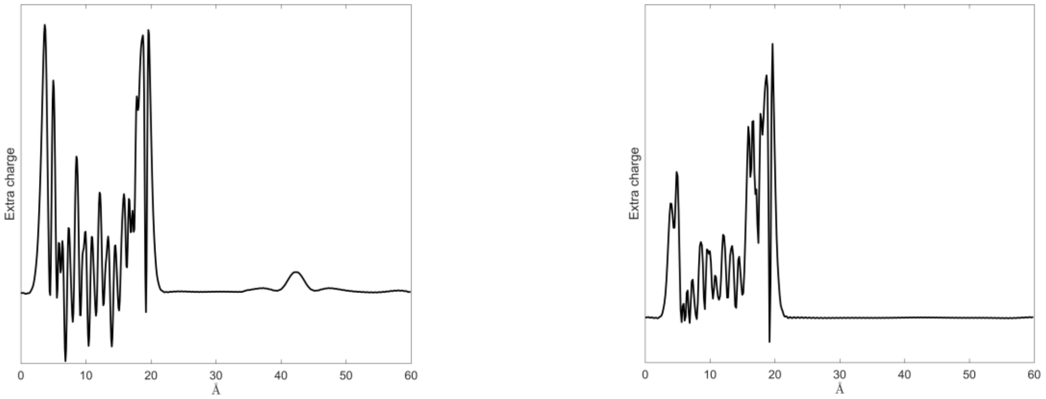
Figure 24a shows the variations of the plane-averaged electrostatic potential of the model in the direction perpendicular to the surface with and without SCPC. A planar average of the corrective potential is shown in Figure 24b. As can be seen, by applying the corrective potential, the spurious potential well in the vacuum between the slabs is eliminated.

The planar average of extra charge in the direction perpendicular to the surface with and without self-consistent potential correction is shown in Figure 25. As expected, applying SCPC mitigates the charge spilling issue and makes the extra charge fully localized on the slab, while *a posteriori* energy correction methods such as SLABCC cannot deal with such cases at all.



a **b**

Fig. 24 a) Variations of the plane-averaged electrostatic potential of models with and without SCPC and b) SCPC corrective potential; in the direction perpendicular to the surface for negatively charged anatase (101) slab models with O_2 molecules on the surface



a **b**

Fig. 25 Planar average of the extra charge a) without, and b) with SCPC correction

Since the corrective potential has a cusp at the supercell boundaries, a small charge accumulation occurs there, which in some cases can cause numerical fluctuations in the total energy. That, in turn, slows down the convergence of the SCF procedure. To minimize the impact of cell boundaries, SCPC uses a damping region in the region near the cell boundary to minimize these fluctuations and speed up the convergence in SCF iterations.

4.2.4 Limitations and future outlook

The spurious potential well in the vacuum between the charged slabs is more relevant to models with a large vacuum between the slabs. The large number of grid points in these models makes them computationally expensive. Additionally, the extra calculations required for SCPC and the limited parallel scalability of VASP can make these simulations impractical on shared computational resources. To overcome this issue, the parallel performance of SCPC and the Poisson solvers used by it must be improved. It is also possible to calculate a loosely converged initial guess for the charged model without SCPC and then use that guess as the starting point of SCPC simulations for faster turnaround.

For modeling the slab, SCPC requires user-defined slab surface positions to generate the dielectric profile. Unlike SLABCC, SCPC does not optimize or verify these parameters, and improper inputs by the user may lead to incorrect results. To ensure the correctness of calculations, the effectiveness of SCPC in eliminating ghost states and the variation of calculated energies in models with different vacuum thicknesses must be individually verified in each case.

The current implementation of SCPC only supports orthogonal cells with slabs perpendicular to the third dimension, and the contribution of the corrective potential has not yet been added to other quantities such as forces, which are required for geometry optimization. However, it is possible to rotate the slab model to be perpendicular to the third direction. Also, SCPC does not produce accurate results for models with localized charges near the numerical grid boundaries. However, this issue can be solved by shifting the center of the extra charge to the center of the simulation cell.

4.2.5 Summary

I showed that in charged slab models with a large vacuum region between the slabs*, a posteriori* charge correction is not sufficient and a self-consistent charge correction method is necessary. I contributed to the development of a new self-consistent potential correction (SCPC) method for charged models under PBC and its verification with *a posteriori* charge correction using SLABCC. In SCPC, the correction term for the KS potential is derived from the difference in electrostatic potential due to the extra charge under PBC and the potential for the same model charge in isolation. I showed that, similar to SLABCC, SCPC is also capable of making the calculated energies model-size independent. Additionally, SCPC can eliminate the ghost states due to the spurious potential well in the vacuum between the slabs and also prevent charge spill-out.

Our method is included in the latest version of VASP and can be added to the old versions as a patch. The details of SCPC and its features are published in Physical Review Letters [259].

4.3 Modelling titanium dioxide surfaces

To investigate photocatalytic reactions on the surface of titanium dioxide, we need a method that is capable of reproducing the bandgap and equilibrium geometry of titanium dioxide and can also describe the charge localization both in the bulk and on the surface of titanium dioxide models. As has already been discussed in Section 3.1.9, errors in the LDA and GGA functionals prevent them from properly describing such quantities. In this section, I will present a computationally low-cost method for correcting these shortcomings.

4.3.1 Bandgap and geometry

I used Lany and Zunger's published gap correction NLEP values for anatase and rutile [212], which are $\Delta V_{O,s} = +3.0$ eV, $\Delta V_{O,p} = -1.0$ eV, $\Delta V_{Ti,p} = -2.2$ eV, and $\Delta V_{Ti,d} = +1.7$ eV. These values are chosen to reproduce the experimental results for the low-temperature optical gaps [263, 264] and the equilibrium geometry.

The unit cell parameters and the calculated band gap values are listed in Table 3. Rutile has a direct gap at Γ while anatase has an indirect gap between X and Γ. As can be seen, by applying band gap correction to the PBE functional, the resulting structural parameters are minimally affected, but it greatly compensates for the typical underestimation of the bandgap by GGA functionals. The resulting bandgaps are in good agreement with the experimental values for the low-temperature optical gap.

Table 3 Unit cell parameters and the band gap for anatase and rutile structures

phase		***a*** **(Å)**	***c*** **(Å)**	***c/a***	***u***	**bandgap (eV)**
anatase	PBE + gap correction (this work)	3.741	9.526	2.546	0.201	3.4
	PBE [265]	3.802	9.774	2.571	0.205	2.08
	PBE [266]	3.759	9.585	2.550	-	-
	PBE [267]	3.784	9.531	2.519	-	2.12
	Experimental [263]	-	-	-	-	3.42
	Experimental [33]	3.785	9.514	2.514	0.208	-
rutile	PBE + gap correction (this work)	4.571	2.960	0.648	0.304	3.0
	PBE [265]	4.647	2.974	0.640	0.305	1.69
	PBE [266]	4.593	2.935	0.639	-	-
	PBE [267]	4.592	2.954	0.643	-	1.84
	Experimental [264]	-	-	-	-	3.06
	Experimental [33, 34]	4.594	2.959	0.644	0.305	-

4.3.2 Charge localization in bulk

A small polaron is the consequence of electron-phonon interactions, which within the Born-Oppenheimer approximation are not taken into account when solving the electronic Schrödinger equation. Therefore, a polaronic state does not arise spontaneously. Small polarons can be modeled by inducing a local displacement in the perfect geometry of the bulk around a specific atom when an extra electron or hole is present. To obtain an initial guess for the local strain around the center of the small polarons, we can substitute one of the elements in the lattice with another element that has ±1 valence electrons. For example, the initial guess for the geometry of the electron polarons in the TiO_2 lattice can be obtained by replacing one Ti with Nb/Al atom in the perfect lattice and subsequently relaxing the whole model. After relaxation, the host atom can be resubstituted with a negative/positive charge, and the system can be relaxed again.

To apply the Lany-Zunger polaron correction, we need to calculate the orbital occupations n_{host} for the Ti3d, O2p for Eqs. (49,50) in the calculations with the band gap correction. These values depend on the employed pseudopotential. I calculated these values from the average of the occupancy matrix for all the levels in that orbital. The resulting occupations are listed in Table 5.

The published parameters for the polaron correction in bulk anatase are λ(Ti3d) = 4.2 eV and λ(O2p) = 4.8 eV [212]. I calculated the total energy correction for the charged models using SLABCC. The required values of the high-frequency dielectric constants were obtained from density functional perturbation theory as 4.79 and 5.56 for anatase and rutile, respectively.

The resulting Kohn-Sham level energies, with respect to the band edges, for the polarons formed by adding an extra charge to the N-electron system in my models, are shown in Table 4. ΔKS(N+1) refers to the energy level of the electron polaron with respect to the conduction band edge, and ΔKS(N-1) to that of the hole polaron with respect to the valence band edge. Removing the extra electron or hole, while keeping the geometry fixed, leads to a neutralized system with N electrons, and the resulting gap state is denoted ΔKS(N). These are compared to ΔSCF, which is the difference between the total energies of the (N±1) and the (N)-electron systems, calculated self-consistently at the equilibrium geometry of the former, and related also to the band edges. As can be seen, the published λ-values satisfy the gKT in anatase to a good approximation, but not for the electron-polaron in rutile.

Table 4 Vertical transitions (in eV) between the defect level and the band edge as calculated from the position of the Kohn-Sham levels (ΔKS) and from total energy differences (ΔSCF), using the λ parameters published in Ref. [212].

phase	polaron type	ΔKS(N±1)	ΔKS(N)	ΔSCF
anatase	electron-polaron	-0.23	-0.24	-0.27
	hole-polaron	+1.25	+1.28	+1.35
rutile	electron-polaron	-1.01	-0.65	-0.91

This implies that the potential strength $\lambda_e^R\left(Ti3d\right)$, which restores the linear behavior of $E(n)$ must be different for this phase. These results appear to be reasonable considering that the *ab initio U*-values also differ in the two modifications [58]. It should be noted that the success of HSE-type hybrid functionals is connected to the correct description of electronic screening [268], which is also different in anatase and rutile.

It has been shown [246] that, if λ exceeds a certain critical value, the geometry of the polaron is rather insensitive to the correction strength, but the charge addition energy changes continuously. Using this observation, the optimal λ parameter can be determined by varying it at a fixed geometry and checking the fulfillment of the gKT. Figure 26 shows the change in the energy eigenvalues and electron addition energies for hole polarons in bulk rutile with changes in correction strength. Based on this procedure, I have selected the optimum value for electron polarons in rutile as $\lambda_e^R(Ti3d)$ = 3.7 eV.

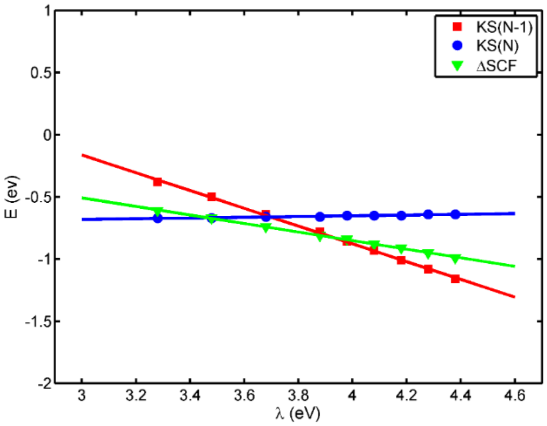


Fig. 26 Variations of energy eigenvalues and electron addition energies for hole polarons in bulk rutile with changes in correction strength λ

The stability of a polaron can be quantified by using its self-trapping energy, which is the energy difference between the polaron and a delocalized charge at the band edge state. The latter is calculated by adding the extra charge and relaxing the system with no initial distortion.

Using the optimized λ parameters, I calculated the vertical ionization energy of electron polarons with respect to the conduction band edge in bulk anatase and rutile (at the optimized geometry) to be -0.2 eV and -0.6 eV, respectively, and the vertical ionization energy of the hole polaron with respect to the valence band edge in anatase to be +1.2 eV. These results are in agreement with HSE06 calculations [269], which found it to be -0.5 eV for electron polarons in rutile and +1.3 eV for hole polarons in anatase. My calculations also reproduce the self-trapping energies reported by Lany [212]: +0.1 eV and -0.3 eV for the electron-polaron in anatase and rutile, respectively, and -0.6 eV for the hole-polaron in anatase.

4.3.3 Charge localization on surface

For surface calculations, I have determined the n_{host} values for the slabs first, using only the NLEPs for gap correction. For fully coordinated surface and subsurface atoms, the change with respect to the bulk calculations was negligible; however, the deviation was significant for the undercoordinated surface atoms. The occupation values are listed in Table 5. I adjusted my correction according to that for undercoordinated atoms while using the optimal λ values from the bulk calculations.

Table 5 Average orbital occupation for the surface atoms of the anatase (101) and rutile (110) models

	bulk		**surface**	
	O(2p)	Ti(3d)	O_{2C} (2p)	Ti_{5C} (3d)
anatase	0.615	0.170	0.598	0.179
rutile	0.622	0.169	0.600	0.173

Next, I localized electron and hole polarons at various atomic sites in my slab models and evaluated the fulfillment of the gKT. As shown in Tables 6 and 7, the gKT remains fulfilled at atoms with bulk-like environments within 0.1 eV.

Table 6 Kohn-Sham level positions and charge addition energies with respect to the conduction band edge (in eV), for electron-polarons localized at different positions in the rutile (110) slab. Values are first given for uniform λ values optimized for bulk. The result after coordination-dependent adjustments is given in parenthesis.

polaron position in the rutile (110) slab	**ΔKS(N)**	**ΔSCF**
Ti_{5C} on surface	-1.25 (-1.23)	-1.13 (-1.23)
Ti_{6C} subsurface, layer 1	-1.17 (-1.17)	-1.09 (-1.09)
Ti_{6C} subsurface, layer 2	-0.83 (-0.83)	-0.79 (-0.79)

However, the deviation between ΔKS(N) and ΔSCF increases above 0.1 eV for the undercoordinated surface atoms (except for Ti_{5C} on anatase (101)). I attribute this to the change in the electronic screening near the surface, and have reoptimized λ values of undercoordinated atoms, to satisfy the gKT also for polarons localized on them. The ΔKS and ΔSCF values using reoptimized the λ are shown in parentheses in Tables 7. This procedure did not have an effect on the fulfillment of the gKT at other positions. The coordination-dependent λ parameters are listed in Table 8.[6]

---

[6] I have not changed the parameters for the hole correction in rutile, since I was not able to localize hole polarons there.

Table 7 Kohn-Sham level positions and charge addition energies with respect to the conduction band edge for the electron-polaron, and to the valence band edge for the hole-polaron (in eV), with the polaron localized on different atoms in the anatase (101) slab. Values are first given for uniform λ values [212]. The result after coordination-dependent adjustments is given in parenthesis.

polaron position in the anatase (101) slab		ΔKS(N)	ΔSCF
electron	**Ti$_{5C}$ on surface**	-1.59 (-1.59)	-1.57 (-1.57)
	Ti$_{6C}$ on surface & Ti$_{6C}$ subsurface	-1.21 (-1.21)	-1.20 (-1.21)
hole	**O$_{2C}$ on surface**	+2.65 (+2.65)	+2.54 (+2.65)
	O$_{3C}$ subsurface	+2.15 (+2.15)	+2.21 (+2.21)

Table 8 Optimized, coordination-dependent λ-parameters of the polaron correction for the rutile (110) and anatase (101) slabs.

polaron	atom-coordination (orbital)	rutile (110)	anatase (101)
electron	Ti$_{5C}$(3d)	4.0	4.2
	Ti$_{6C}$(3d)	3.7	4.2
hole	O$_{2C}$(2p)	4.8	5.1
	O$_{3C}$(2p)	4.8	4.8

Next, I tested the surface-adapted Lany-Zunger polaron correction method. Table 9 lists the self-trapping energies calculated by my optimized parameters. The results comply with the experimentally known situation [52, 58]: electron polarons can be localized in rutile at almost any Ti atom, but are slightly preferred energetically at a Ti$_{6C}$ site immediately below a Ti$_{5C}$ site. In the anatase (101) slab, electron-polarons could only be localized on the surface, but even there, they are less strongly bound than in rutile.

Table 9 Self-trapping energy of polarons at the surface of the rutile (110) and anatase (101) slabs.

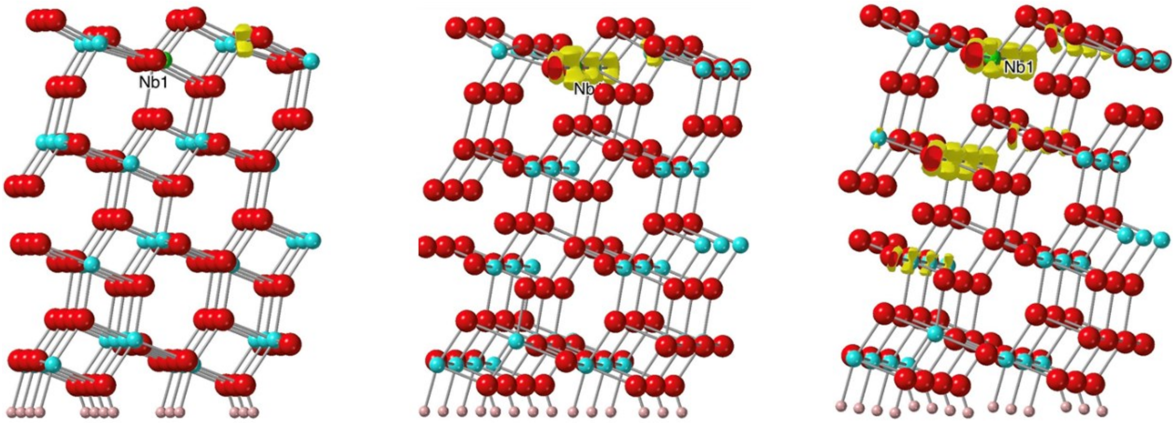
polaron	position	rutile (110)	anatase (101)
electron	Ti$_{5C}$ on surface	-0.3	-0.3
	Ti$_{6C}$ on surface	–	-0.1
	Ti$_{6C}$ subsurface, layer 1	-0.4	–
hole	O$_{2C}$ on surface	–	-1.1
	O$_{3C}$ on surface	–	-0.7

The vertical ionization energies of the surface electron-polarons can be deduced from the STS (scanning tunneling spectroscopy) measurements of Ref. [58], which provide 0.7±0.1 eV on the rutile (110) surface and 1.0±0.1 eV on the anatase (101) surface. In the case of rutile, the STS peak also appears if measured away from a surface oxygen vacancy, so it can be assumed that it arises due to free electron polarons. My value for the most stable position of the free polaron (see Table 6, Ti_{6C} subsurface 1) with the optimized parameters for rutile (in parentheses) is 1.1 eV, i.e., about 0.4 eV higher than the experimental value. In the case of anatase, the STS peak only appears if measured above an oxygen vacancy.

I have, therefore, also calculated the vertical transition energies ($\Delta KS = \varepsilon_{KS} - \varepsilon_{CB}$) of a surface oxygen vacancy in both models, removing an O_{2C} atom. On the rutile (110) surface, I find the two electrons of the vacancy to be the most stable in two polaron states at Ti_{6C} subsurface 1 sites (Figure 8), symmetric to the oxygen vacancy, in agreement with HSE calculations [270]. The one-electron level, computed in my five-layer model is 1.47 eV below the conduction band edge, somewhat deeper than the 1.22 eV obtained in a four-layer calculation by HSE [270]. On the anatase (101) surface, the electrons are retained by the surface vacancy, giving rise to two levels in the four double-layer slab at 1.76 and 2.11 eV below the conduction band edge, again deeper than the HSE values from a three double-layer calculation, 1.20 and 1.39 eV, respectively [57]. (The much smaller splitting between the two states in the latter indicates that a GGA calculation cannot mimic the non-local exchange even with the applied corrections.) Also, similar to the HSE calculation [57], I find that a subsurface oxygen vacancy loses one electron into a polaron state on a surface Ti_{5C} site, giving rise to a level at 1.48 eV energy. It appears likely that this less bound polaron gives rise to the observed STS peak at 1.0 eV on the anatase (101) surface. The deviation from the experiment is then the same as in the case of the rutile (110) surface. This analysis shows that the applied corrections allow for a qualitatively correct picture of the polarons in both modifications of TiO_2 but the quantitative accuracy is somewhat limited.

While the Lany-Zunger polaron correction method can successfully describe the charge localization on pure titanium dioxide models, adapting it for models with impurity atoms (dopants) where more than one type of atomic orbital contributes to the band state is not trivial. In TiO_2, very often the DFT+U method or hybrid functionals are used for polaron studies, irrespective of their fulfillment of the gKT. Such functionals provide a higher degree of localization than plain GGA, but not necessarily the correct amount. The standard HSE06 functional (0.25,0.20) satisfies the gKT and reproduces the 0K single-particle gap for anatase [271]. Therefore, we will use it for the study of charge localization in the presence of niobium dopants.

Starting from a model with lattice parameters corresponding to the HSE06 result for an ideal bulk crystal (a = 3.755 Å, c = 9.561 Å), if we introduce an extra electron by substituting Nb on the sixfold-coordinated surface Ti_{6C} site, (Nb is more stable on Ti_{6C} than on a Ti_{5C} site by 0.3 eV [272]), we obtain three different localized solutions. In the first case, similar to the pure TiO_2 case, the electron is trapped in the dangling bond of a surface Ti_{5C} site (Figure 27a). In the second case, the electron stays on Nb forming essentially an effective mass-like (EMT) donor state [53] (Figure 27b). These two cases are similar in energy. However, the third case where an electron is spread out in a (001) plane, around a subsurface Ti_{6C} atom (Figure 27c) is about 0.3 eV lower in energy than the first two. Since this state is not centered on the Nb atom but on an adjacent Ti, it is a polaron bound by the ionized donor. These results show that unlike in bulk, an electron can be self-trapped on the anatase (101) surface in various forms, with the 2D-localized state being the most favorable.



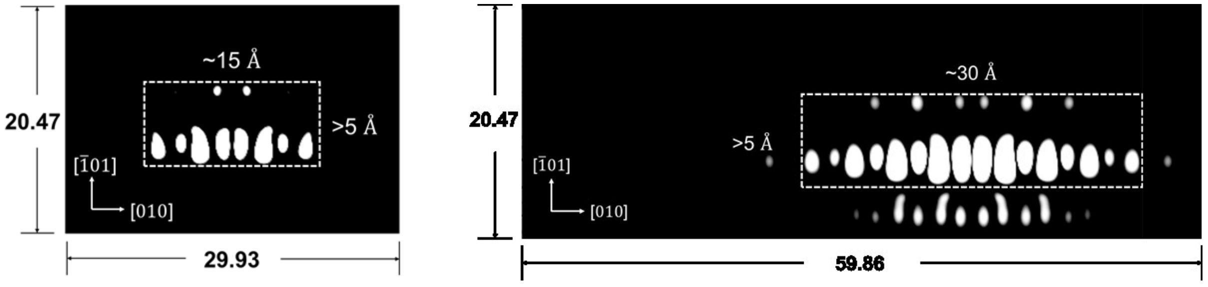
a **b** **c**

Fig. 27 Localization of the extra electron of a Nb donor (at a surface Ti_{6C} site) a) trapped at a surface Ti_{5C} atom, b) centered on the dopant, and c) in a 2D polaron state

The model sizes allowed by a HSE calculation are insufficient to assess the size of the 2D-localized state. Therefore, we have used the PBE+U functional (U =3.9 eV has proven to be capable of capturing the proper localization in this case [58, 273]) on two larger 960-atom and 1920-atom anatase (101) slabs, to show that these localized states resemble the scanning tunneling microscopy (STM) observations [58]. 960-atom and 1920-atom models were constructed from 8×2 and 16×2 surface unit cells with 5 double-layers using unit cell parameters listed in Table 3. No pseudo-hydrogen termination was used in these models. In the 960-atom model, the bottom 2 double-layers were kept frozen, but in 1920-atom model, all atoms were allowed to relax.

Substitution of a Nb donor on a surface Ti_{6C} site in the 960-atom slab gives rise to a 2D-localized state in the (001) plane, centered on a subsurface Ti atom. The state extends to the bulk region of the plane. A similar state was found by Selçuk and Selloni [273] in a study where hydrogen absorption onto O_{2C} provided the excess electron. Using first-principles molecular dynamics simulations, they also observed the motion of this polaron between different (001) planes at 400K within picoseconds, so it is quite mobile. The simulated STM image (at -1.0 V bias and 1.05 Å tip distance) of the 960-atom slab with the medium-size 2D polaron state is shown in Figure 28. As can be seen, on the one hand, the STM view does not show the full size of the polaron, since it is subsurface, and on the other, at its greatest extent, the state extends beyond the cell boundary. Therefore, we have created a similar 2D polaron in the large slab, placing the Nb atom at the bottom surface, and then cleaved the upper half of the slab for STM simulation. The STM simulation was performed in constant-height mode (-1.0 V bias and 1.27 Å tip distance) using P4VASP[7], which is based on the simplification of the Tersoff-Hamann approach [274, 275] by Selloni et al. [276].

Our simulated STM images (Figure 28) and the observed sizes in the two slabs are in excellent agreement with the those reported by Setvin et al. [58]. Since the 2D polaron state is localized entirely in the large (1920-atom) slab, we could calculate the position of its electronic level. After potential alignment to the pristine system, we obtain it at 33 meV below the conduction band.



a **b**

Fig. 28 Simulated STM images of 2D polarons in a cleaved a) medium-sized and b) large slab model.

The good agreement with the measured 40 meV for the "large polaron" observed by ARPES on the (001) surface after creating vacancies [49], suggests that the 2D state we find may exist

---

[7] https://www.vasp.at/py4vasp/

independent of the Nb dopant or of the (101) surface. Indeed, our HSE06 calculations in the 96-atom bulk supercell (2√2×2√2×1 multiple of the Bravais cell, with 2×2×2 Monkhorst-Pack Brillouin-zone sampling) confirm this. With the help of an Nb dopant, a 2D state can be created in the (001) plane, which persists even after the Nb is replaced by Ti and the system is allowed to relax again.

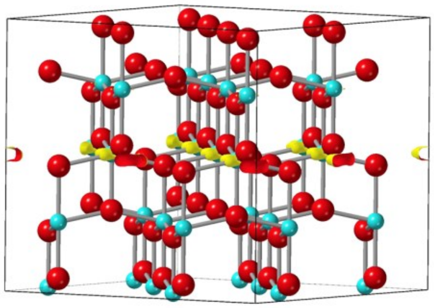


Fig. 29 Localization of an excess electron in a 2D polaron state in a bulk anatase model.

4.3.4 Summary

I applied the band-gap and polaron correction methods of Lany and Zunger to slab models and showed that the original method is not accurate on the surface due to the different screening environment. I showed that by making the atom- and angular-momentum-dependent parameters of the Lany–Zunger polaron correction also coordination-dependent, it is possible to correctly describe charge trapping in small polaron states on the anatase (101) and rutile (110) surfaces while keeping the total computational cost low. Our results also show that the "large polaron" in anatase is actually of medium size and can be obtained in a static calculation. These states are two-dimensional in the (001) plane and can get trapped at the (101) surface.

4.4 Photocatalytic CO oxidation on the anatase (101) surface

I have used my developed scheme [277] to investigate the possibility of having a fully cyclic photo-assisted CO oxidation reaction on the anatase (101) surface. The idea of a hole-mediated CO oxidation was suggested on the basis of calculations with an uncorrected GGA functional [13], so I first repeated them using my improved functional. In addition, I considered the restoration of the starting surface by O_2 molecules, which act as electron scavengers after the photo-excitation. I will show that the whole reaction sequence is energetically feasible, so a photocatalytic CO oxidation cycle on the anatase (101) surface in the presence of molecular oxygen is possible.

The process of heterogeneous photocatalysis can be decomposed into a series of chemical reactions on the surface of the photocatalyst, following the electron-hole pair generation by photoexcitation. I will not consider here the mechanism of separating the electron and the hole, and I will decouple the reactions involving them by considering independent models with a negative and a positive charge, respectively:

$$2[TiO_2] + e^- + h^+ \rightarrow [TiO_2]^+ + [TiO_2]^- \qquad (72)$$

I will investigate first the interaction of CO in the gas phase with the positively charged anatase (101) surface $[TiO_2]^+$, then the electron scavenging reaction of O_2 with the negatively charged anatase (101) surface $[TiO_2]^-$, and finally the restoration of the neutral surface.

4.4.1 Adsorption of CO on the positively charged surface

$$[TiO_2]^+ + CO \rightarrow [TiO_2{:}CO]^+ \qquad (73)$$

The interaction of the CO molecule with the anatase (101) surface has been studied experimentally using infrared reflection and absorption spectroscopy [10, 278], and theoretically with DFT using the PBE functional [13, 278]. For the positively charged surface, I found the most stable adsorption configuration to be the case when the carbon atom of CO binds to an undercoordinated surface titanium atom (Ti_{5C}…C-O as shown in Figure 31b), in agreement with the experimental data. The calculated adsorption energy of -0.5 eV is in good agreement with the value obtained by an uncorrected PBE functional, -0.45 eV [13].

4.4.2 Oxidation of CO on the positively charged surface

$$[TiO_2{:}CO]^+ \rightarrow [TiO_2{:}V_O{+}CO_2]^+ \qquad (74)$$

The next step is the reaction of the adsorbed CO with a neighboring undercoordinated oxygen atom on the surface (O_{2C}). As found in Ref. [13], the bond of this O_{2C} with its Ti_{6C} neighbor may break up, and a CO_2 unit can be formed, bonded to a Ti_{6C} atom (Figure 31c). This configuration can be considered as chemically adsorbed CO_2 on the surface of anatase (101), next to a surface oxygen vacancy. I found the energy barrier for this reaction to be +0.5 eV, with a net gain in the total energy at the end of the reaction to be -0.6 eV. In comparison, the corresponding energies reported with the uncorrected PBE functional were +0.4 eV and -1.6 eV, respectively [13]. While the barriers are comparable, the net energy gain is much lower after the correction, due to the delocalization error of the PBE functional which is especially prominent in the charge distribution of the positively charged models which include an oxygen vacancy. For example, the localization of the positive charge (difference between the charge distribution of the positively charged model and the neutralized model with the same geometry) for the $[TiO_2{:}V_O{+}CO_2]^+$ model with PBE and PBE + band-gap, and polaron corrections has been shown in Figure 30.

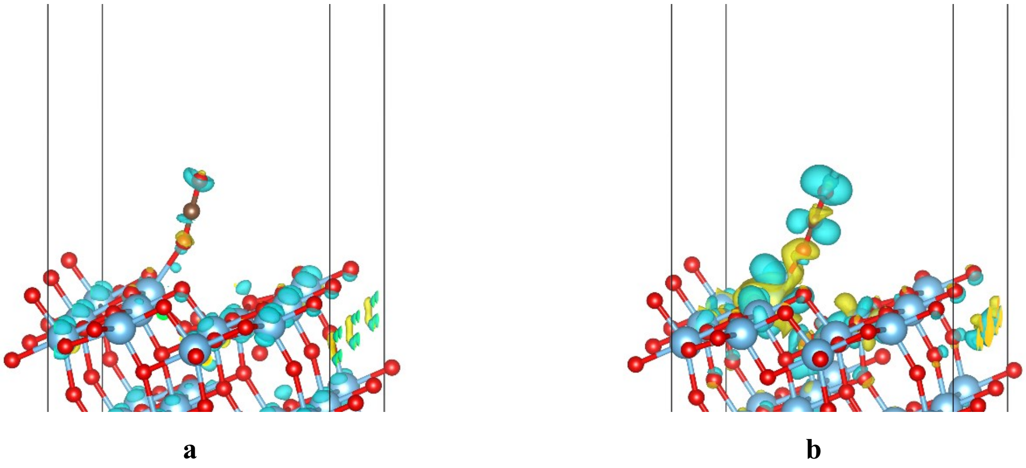


Fig. 30 Localization of the positive charge in $[TiO_2{:}VO{+}CO_2]^+$ models with a) PBE and b) PBE + band-gap, and polaron corrections (isosurface level 0.001)

4.4.3 Desorption of CO_2 from the positively charged surface

$$[TiO_2{:}V_O{+}CO_2]^+ \rightarrow [TiO_2{:}V_O]^+ + CO_2 \qquad (75)$$

After the formation of CO_2 on the anatase surface, the desorption process leaves an oxygen vacancy at the O_{2C} site. I found the required energy for desorption to be +0.7 eV. (The calculated energy with PBE functional was reported to be +0.4 eV [13]) NEB calculations (as well as earlier molecular dynamics simulations [13]) show a negligible barrier for dissociation. The ball-and-stick model representation of the reaction steps involved in the oxidation of CO on the surface and desorption of CO_2 is shown in Figure 31.

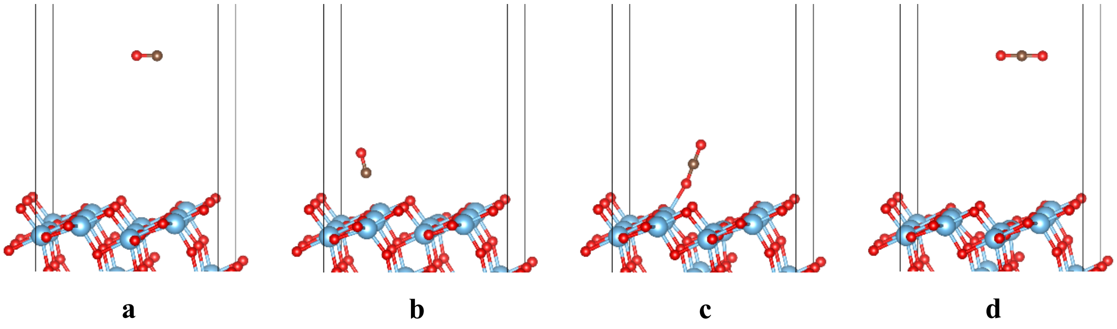


Fig. 31 Ball-and-stick model representation of the stable atomic configurations in CO oxidation over the positive anatase (101) surface. From left to right (a-d): $[TiO_2]^+ + CO$, $[TiO_2{:}CO]^+$, $[TiO_2{:}V_O{:}CO_2]^+$, $[TiO_2{:}V_O]^+ + CO_2$

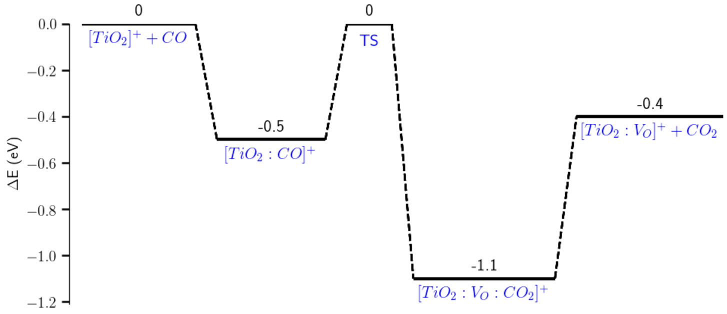


Fig. 32 Schematic energy diagram for the oxidation of CO on the positively charged anatase (101) surface

The schematic energy diagram for the reaction of CO with the positively charged anatase (101) surface Eqs. (73-75) is shown in Figure 32. As can be seen, the total energy balance is negative (the process is exothermic), and the energy released by the absorption can surmount the barrier for forming $[TiO_2{:}V_O{+}CO_2]^+$. The reaction heat released during that amply covers the dissociation energy of CO_2. Although some of the absorption and reaction energy may be dissipated, the whole reaction sequence can be expected to occur with a high likelihood at room temperature. Next, we consider electron scavenging by O_2 molecules.

4.4.4 Adsorption of O_2 on the negatively charged surface

$$[TiO_2]^- + O_2 \rightarrow [TiO_2{:}O_2]^- \qquad (76)$$

Using charge scavengers can prevent electron-hole recombination, increase the carrier lifetime, and enhance the photocatalytic activity. Molecular oxygen is a widely used electron scavenger for photocatalytic oxidation reactions, which, on the other hand, generate active oxygen species [279]. It has also been shown that the presence of excess electrons is essential to O_2 adsorption on TiO_2 [280], and the presence of O_2 gas is necessary for photo-assisted CO oxidation on the anatase (101) surface [10].

In agreement with the previous studies [281], I found that in the most stable configuration of adsorbed O_2 on the negatively charged anatase (101) surface, both oxygen atoms are bonded to a Ti_{5C}. This configuration is shown in Figure 33. The adsorption energy is 1.2 eV, in good agreement with the 1.0 eV reported based on calculations with the PBE0 hybrid functional [281].

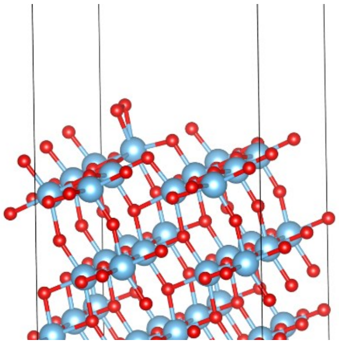


Fig. 33 Ball-and-stick model representation of the most stable configuration for the O_2 molecule adsorbed on the negatively charged anatase (101) surface model

4.4.5 O_2^- surface diffusion and annihilation of V_O^+

It can be expected that the negatively charged O_2 molecule will diffuse along the MEP on the surface to annihilate the positively charged surface vacancy.

$$[TiO_2:V_O]^+ + [TiO_2:O_2]^- \rightarrow [TiO_2:O] \quad (77)$$

I found the activation energy for the O_2 diffusion on the negatively charged anatase surface to be 0.4 eV. This transition state directly connects the two most stable configurations (Ti_{5C}..O-O) adjacent to each other on the anatase (101) surface.

The reaction of O_2 with the neighboring V_O results in an oxygen adatom, bonded to a twofold-coordinated surface oxygen (O_{2C}…O). This reaction was also found to be barrierless. This process for the surface reconstruction in conjunction with the charge neutralization releases ~2.3 eV in total. This energy was calculated from the difference in energy of finite-sized models corresponding to the following reaction:

$$[TiO_2{:}V_O]^+ + [TiO_2{:}O_2]^- \rightarrow [TiO_2{:}O] + [TiO_2] \qquad (78)$$

As the NLEPs are implemented in VASP based on the simplified DFT+U approach introduced by Dudarev *et al.* [282], applying the correction to atoms does change the total energy of the model. The models of the reactants and products used for the energy calculation of Eq. (78) (as a result of surface reconstruction) have different numbers of O_{2C} atoms, and therefore the difference in their energy due to this inconsistency must be accounted for. Hence, I used the energy of a pristine surface with NLEP correction removed from one of its surface O_{2C} atoms as the reference for the second product of Eq. (78).

4.4.6 O surface diffusion and O_2 formation on the neutral surface

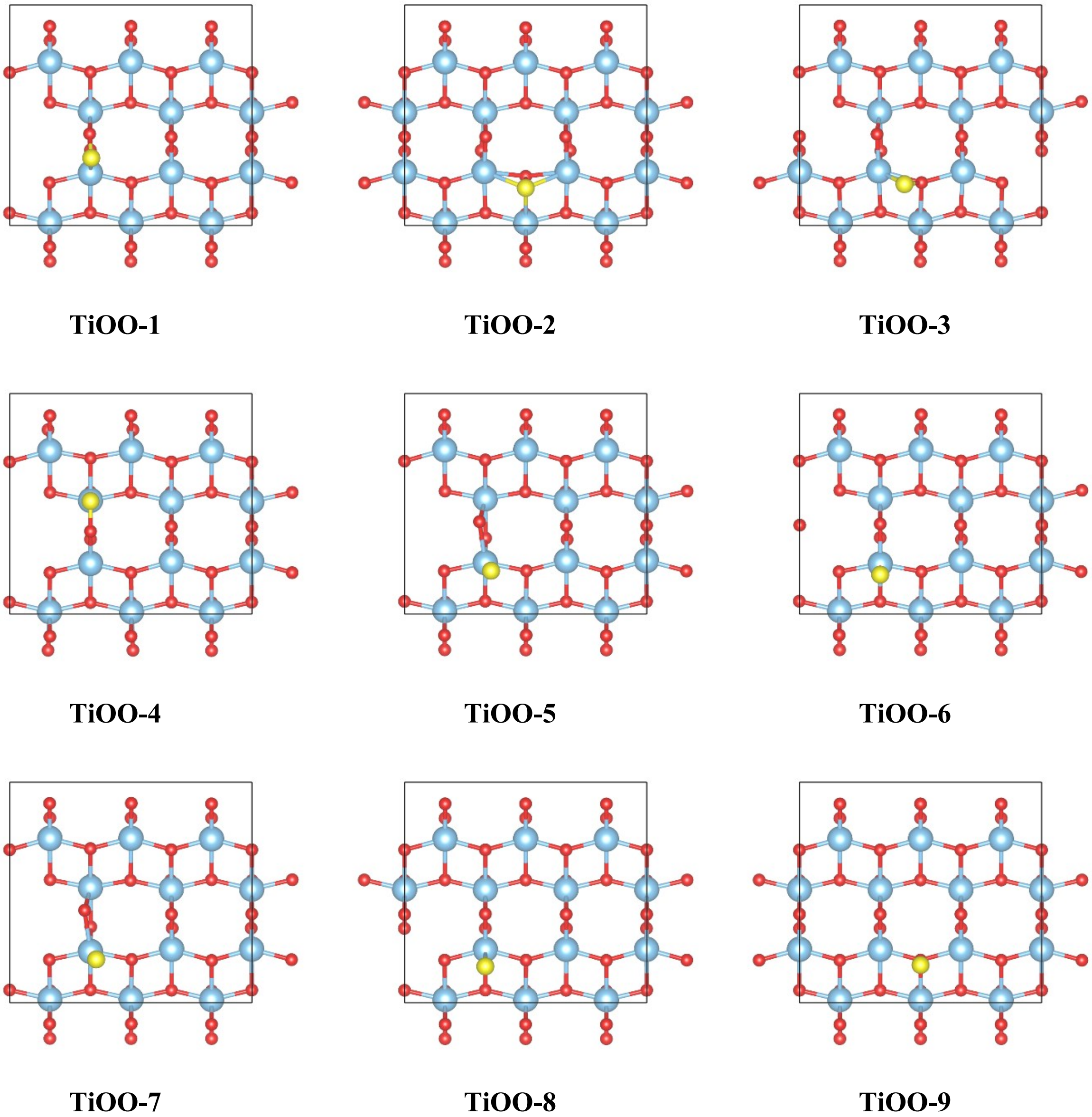
The neutral oxygen adatom, created in the last step [Eq. (77)], can also diffuse on the surface and react with other similar oxygen adatoms, forming an adsorbed oxygen molecule.

$$2[TiO_2{:}O] \rightarrow [TiO_2{:}O_2] \qquad (79)$$

There are multiple locally stable configurations to be considered for the diffusion path of an oxygen atom on the neutral anatase (101) surface. To ensure the stability and uniqueness of the obtained geometries, I have optimized the geometries using a much tighter global break condition for the electronic SC-loop (10^{-7} eV) and for the ionic relaxation loop (10^{-4} eV/Å). I could identify 9 stationary points for the oxygen adatom on the neutral slab surface. The energies of these configurations relative to the most stable one (TiOO-1) are listed in Table 10, and the corresponding geometries are shown in Figure 34.

Table 10 Relative energy of various configurations for an oxygen adatom on anatase (101) surface with respect to the most stable geometry

Configuration	Relative energy (eV)
TiOO-1	0
TiOO-2	+0.39
TiOO-3	+0.47
TiOO-4	+0.65
TiOO-5	+0.74
TiOO-6	+0.78
TiOO-7	+0.83
TiOO-8	+1.22
TiOO-9	+1.75



TiOO-1 **TiOO-2** **TiOO-3**

TiOO-4 **TiOO-5** **TiOO-6**

TiOO-7 **TiOO-8** **TiOO-9**

Fig. 34 Ball-and-stick model representation of the stable configurations of the single oxygen adatom on the anatase (101) surface from top view. The oxygen adatom is marked yellow.

I used the CI-NEB method to locate the possible transition states and identify the MEPs connecting these configurations. I propose an MEP for the diffusion of an oxygen adatom from the TiOO-1 to an equivalent position on the adjacent atom on the surface following the TiOO-1 → TiOO-7 → TiOO-3 → TiOO-2 path, as shown in Figure 35. Due to the symmetric nature of the surface, the rest of the diffusion path follows the same configurations on the equivalent neighboring surface atoms but in the reverse order.

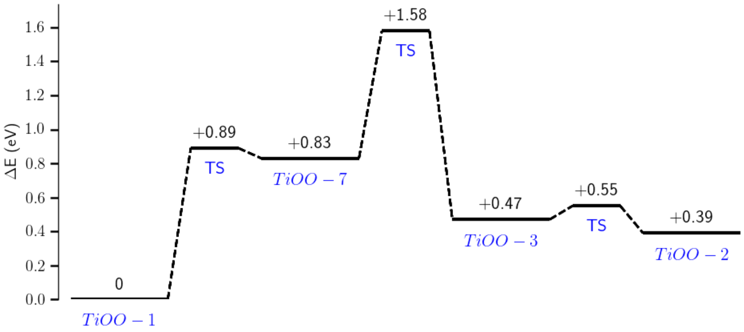


Fig. 35 Minimum energy path for diffusion of a single oxygen atom on the anatase (101) surface. All energies are relative to TiOO-1 geometry.

Assuming perfectly adiabatic conditions, diffusion of an atom from TiOO-1 to an equivalent position on the adjacent atom on the surface requires no extra energy. Following this path, the maximum energy barrier for each of the diffusion steps is calculated to be ~0.9 eV (TiOO-1 → TiOO-7). There are other steps that release energy during the diffusion process along this path, and the formation of an O_2 molecule from two oxygen adatoms at the adjacent local minima (TiOO-1) releases ~0.9 eV in total, which may compensate for the required energy for oxygen diffusion on the surface.

An exhaustive search for the transition states of a reaction involving multiple chemisorbed atoms on the surface is a complicated task. The PES for these reactions usually has a high degree of freedom, and numerous configurations may have close energy or near-zero energy derivatives with respect to the geometry changes. In the case of O_2 formation on the neutral anatase (101) surface, it is possible to encounter surface restoration and desorption while searching for the MEP. The CI-NEB method uses force projection in order to direct the optimizers towards MEP and TS. In situations where the calculated forces are not very accurate (far from or close to the minimum), the optimizers may not always be successful. I found an energy barrier for the formation of O_2 molecules on a neutral anatase (101) surface with +1.3 eV energy relative to the isolated O adatoms on adjacent TiOO-1 sites. My findings set an upper bound on the required energy for this reaction, but other undiscovered paths with different (lower energy) transition states may also be present, and a systematic search is needed to ensure this TS corresponds to the MEP. Figure 36 shows the geometry of this TS (Figure 36b) and two other configurations along the reaction path (Figure 36a,c) for the O_2 formation on the neutral anatase (101) surface. The reaction path following this

TS involves breaking a bond between an oxygen adatom and an O_{3C} atom on the surface, followed by bond formation between two oxygen adatoms.

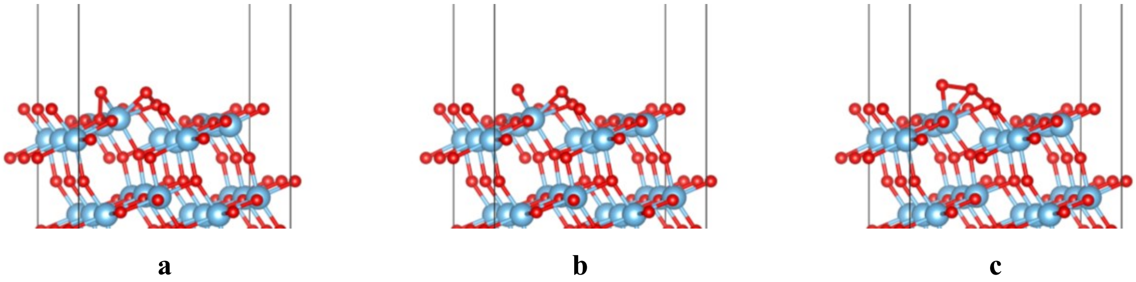


Fig. 36 Ball-and-stick model representation of configurations along the reaction path for O_2 formation on the natural anatase (101) surface: a) along the path towards separate O adatoms, b) TS, and c) along the path towards O_2

The geometry of the most stable configuration for an O_2 molecule on the neutral anatase (101) surface is shown in Figure 37. The orientation of an O_2 molecule in this geometry differs from that of most stable configuration for an O_2 molecule on a negatively charged anatase (101) surface (Figure 33).

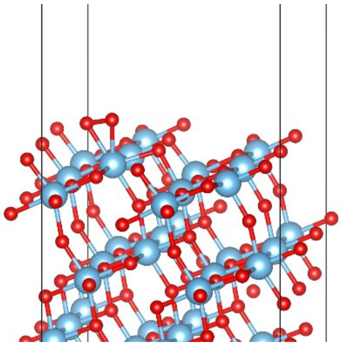


Fig. 37 Ball-and-stick model representation of the most stable configuration for the O_2 molecule adsorbed on the neutral anatase (101) surface model

4.4.7 Desorption of O_2 from the neutral surface

Finally, the resulting O_2 from Eq. (79) can desorb from the surface:

$$[TiO_2{:}O_2] \rightarrow [TiO_2] + O_2 \qquad (80)$$

which requires 0.4 eV of energy, and leaves a pristine surface behind. The schematic energy diagram for the reaction of oxygen atoms on the uncharged anatase (101) surface and its desorption, Eqs. (79,80) is shown in Figure 38.

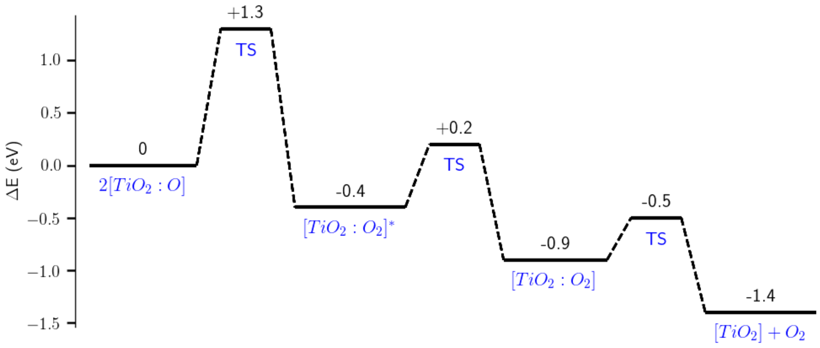


Fig. 38 Schematic energy diagram for the formation of the O_2 molecule on the uncharged anatase (101) surface and its subsequent desorption

4.4.8 Overall reaction

Combining all the steps above in Eqs. (72-80), we will have:

$$[TiO_2] + CO + O_2 + e^- + h^+ \rightarrow [TiO_2] + CO_2 + \tfrac{1}{2}O_2 \qquad (81)$$

This reaction is compatible with the experimental observations that have found the existence of both ultraviolet radiation and oxygen gas to be essential for the oxidation of CO on anatase (101) surfaces. The surface of the catalyst is reconstructed at the end of this reaction, and considering all the proposed reaction steps, the energy required for each fundamental step was found to be sufficiently small, implying a plausible mechanism for photocatalytic oxidation of CO over the anatase (101) surface.

4.4.9 Summary

Using the developed methods for modeling anatase (101) surfaces and corrections for charged models under periodic boundary conditions, I investigated the possibility of a complete catalytic cycle for the oxidation of CO over the anatase (101) surface. In addition to already published reports that have shown that this reaction creates an oxygen vacancy on the photocatalyst surface, I proposed a mechanism to eliminate the surface oxygen vacancies and the positive charge by including electron-scavenging oxygen molecules in the gas phase. I investigated the diffusion path for oxygen adatoms on the anatase (101) surfaces, the formation of oxygen molecules, and their subsequent desorption.

The calculated energies are consistent with the available theoretical reports using hybrid functionals. Additionally, the required energy for each of the fundamental steps in the proposed reaction path is sufficiently small, which can be compensated for by the energy gain during the surface vacancy annihilation, O_2 formation, and subsequent desorption. This implies that it is a plausible mechanism for photocatalytic oxidation of CO over the anatase (101) surfaces.

Chapter 5 Summary and outlook

My implementation of the Komsa-Pasquarello charge correction scheme simplifies the process of calculating the energy correction [250]. I extended the method to handle mediums with anisotropic dielectric tensors as well as cases where the extra charge is localized at multiple sites [252, 253]. I discovered ghost states while investigating surface reactions using negatively charged slab models and contributed to the development of a self-consistent correction method for these models [259].

I showed that it is possible to mitigate common shortcomings of GGA-based functionals and fulfill the gKT for the charges localized both on the surface and in the bulk models of TiO_2 by making the atom- and angular-momentum-dependent parameters of the Lany-Zunger polaron-correction also coordination-dependent [277]. This method can provide a qualitatively correct picture of the polarons in both anatase and rutile. I also encountered a 2D electron polaron in anatase, which could explain the experimental results [283].

I used the developed methods to investigate the photocatalytic CO oxidation on the anatase (101) surface. The obtained results are in good agreement with the available experimental data as well as the results obtained from more advanced (and computationally much more expensive) approximations. I proposed a mechanism for eliminating the surface oxygen vacancies by including electron-scavenging oxygen molecules in the gas phase, which makes it possible to have a complete catalytic cycle for the oxidation of CO over the anatase (101) surface.

In future works, I plan to improve the performance of my charge correction code SLABCC, simplify the process, and reduce the required user intervention. I will also systematically investigate the possible reaction paths for the formation of O_2 on the anatase (101) surface. Our current implementation of the self-consistent potential correction is computationally too expensive for large-scale studies. The parallel performance of SCPC must be improved for use in these studies. The contribution of the corrective potential should be added to other quantities such as forces, which are required for geometry optimization.

Appendix 1 SLABCC

A1.1 Implementation

To calculate the total energy correction due to the extra charge under periodic boundary conditions, I implemented the charge correction scheme proposed by Komsa and Pasquarello [70] in a standalone code named SLABCC [250]. This code is written in C++ using the Armadillo linear algebra library [284]. Armadillo is a template-based library designed to have a syntax similar to MATLAB/Octave and can automatically use OpenMP multi-threading to speed up matrix operations through integration with Intel MKL or OpenBLAS.

A1.2 Flowchart

Figure 39 shows a simplified flowchart of the SLABCC code. Various stages of the calculations have been marked and the relevant steps in the calculation of correction terms in Eq. (9) are highlighted.

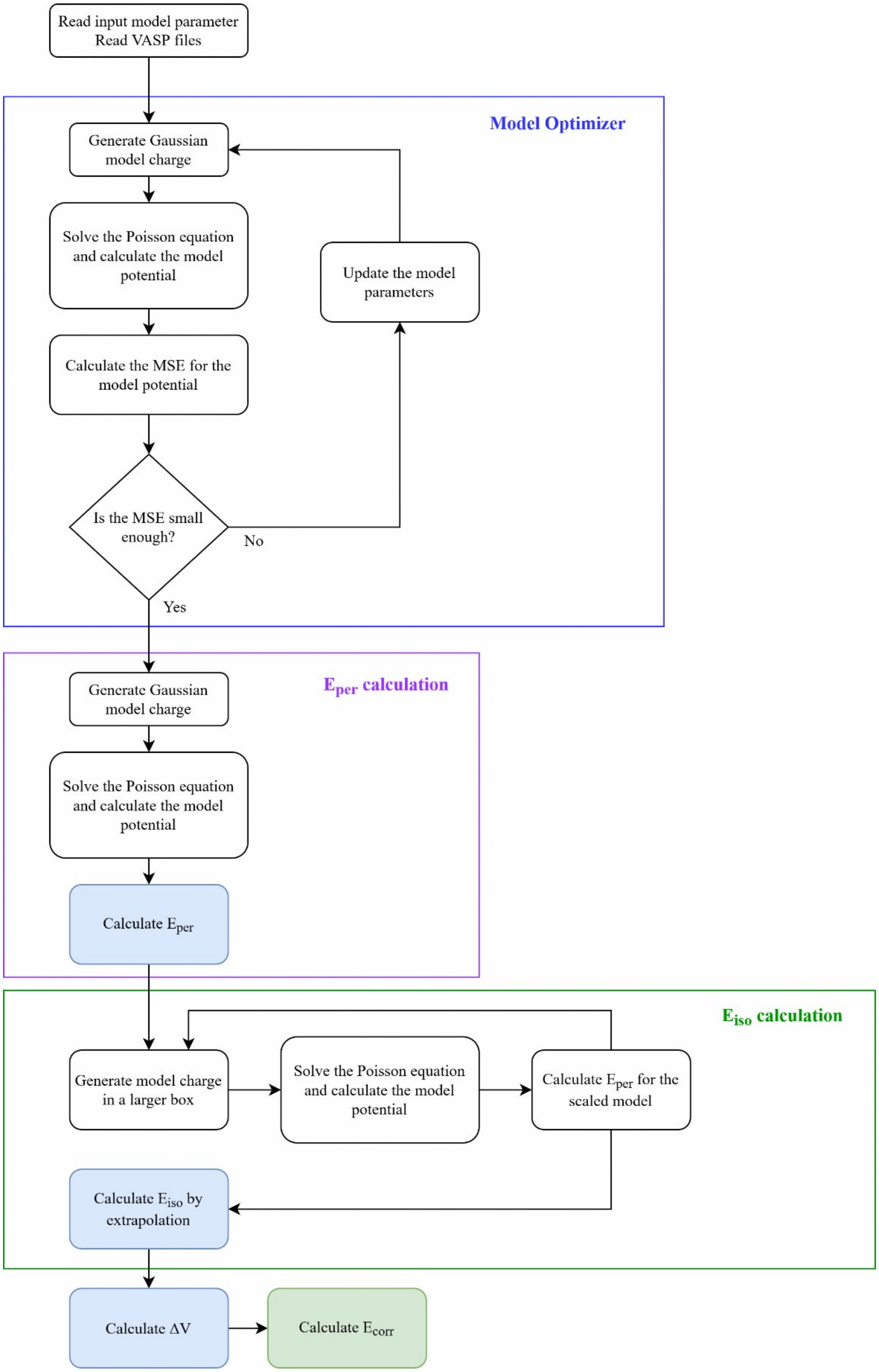


Fig. 39 Simplified flowchart of the SLABCC code.

A1.3 Optimization of the model parameters

In the case of the models with multiple charges, we also need to define the fraction of the charge in each Gaussian, q_i in Eq. (58), while applying the constraint on the total charge of the model Eq. (59). Therefore, we are left with $5n_c+1$ (n_c=number of Gaussian charges in the model) parameters to optimize. In some cases, to reduce the number of parameters and improve the performance of optimization, it is desirable to have the possibility of optimizing only a subset of the model parameters while keeping the rest fixed.

To have the flexibility to efficiently tackle a wide range of models with various numbers of variables, I integrated the Constrained Optimization BY Linear Approximations (COBYLA) [285], Bound Optimization BY Quadratic Approximation (BOBYQA) [286], and S.G. Johnson's implementation of the Subplex (subspace-searching simplex) (SBPLX) [287] algorithms as implemented in the NLOPT library [288] for minimizing the MSE in the SLABCC. COBYLA is an alternative to the well-known simplex[8] reflection method of Nelder and Mead [289] but additionally, it uses the trust region framework [290], and supports arbitrary nonlinear inequality, and equality constraints. COBYLA employs the function values at the vertices of a simplex to approximate the objective function with a linear multivariate interpolation. In contrast to the common methods based on two-level factorial designs, the simplex contains the minimum number of points required to approximate the first-order variations, which decreases the overall computational cost. COBYLA also ensures that the simplex does not collapse into a lower-dimensional hyperplane. Details of the algorithm can be found in [291]. Unlike COBYLA, the BOBYQA algorithm constructs quadratic models to solve the optimization problem. SBPLX is a variant of Nelder-Mead that uses Nelder-Mead on a sequence of subspaces.

All three algorithms support bound constraints to define the minimum and maximum values that a model variable can take. However, not all support nonlinear inequality constraints. To enforce total charge conservation in the case of generating a model with multiple Gaussian charges, I have added a quadratic penalty to the MSE based on the difference between the total charge of the model and the target value from the input files.

---

[8] A simplex in n-dimensions is a convex hull of n+1 points.

Figure 40 shows the convergence behavior of various optimization algorithms in two different models. Figure 40a shows the convergence of MSE (and equivalently model parameters) in constructing a model charge for a slab with a single Gaussian. The convergence of the optimizers for a bulk model with two Gaussian charges is shown in Figure 40b. In each plot, all three optimizers start from the same initial conditions and converge to the same MSE value. However, for the sake of clarity, the lines are plotted with a vertical offset. As can be seen from these two examples, the performance of the optimizers, i.e., the number of iterations before convergence, depends on the problem they are working on. In each optimization step, the charge distribution and the dielectric profile need to be recalculated with the new parameters, and the resulting Poisson equation needs to be solved for the evaluation of the model's MSE. Therefore, the performance of the optimizers directly influences the total runtime of SLABCC. By default, SLABCC will use the BOBYQA algorithm for the optimization process, but users can change the optimization algorithm in the input file.

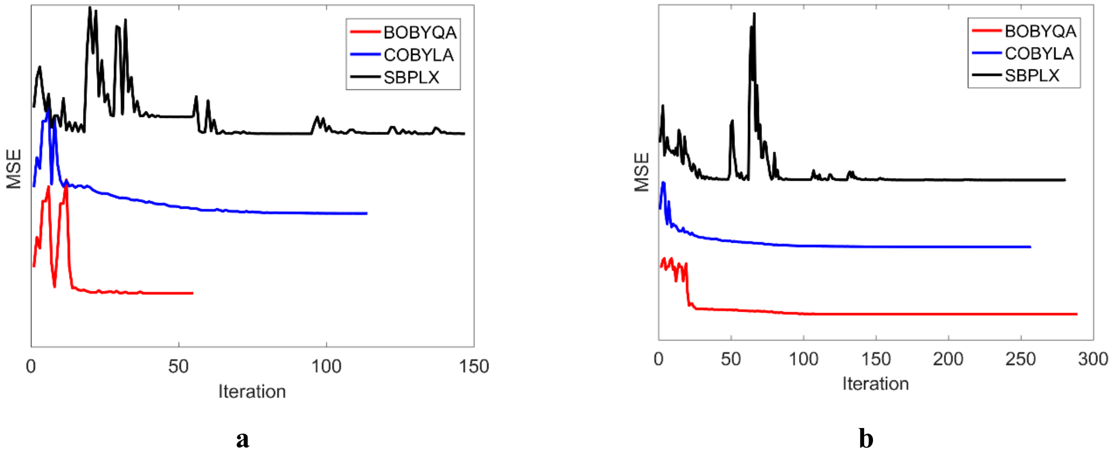


Fig. 40 Change in the MSE of the model in each iteration for a) a single Gaussian charge on a slab model and b) two Gaussian charges in a bulk model

A1.4 Performance optimizations and internal checks

Depending on the number of model parameters, the initial guess provided by the user, the optimization algorithm, and its convergence tolerance, the process of finding optimal model parameters may require tens or hundreds of iterations. To speed up the process, it is possible in SLABCC to optimize the model parameters on a coarser grid than the original input files.

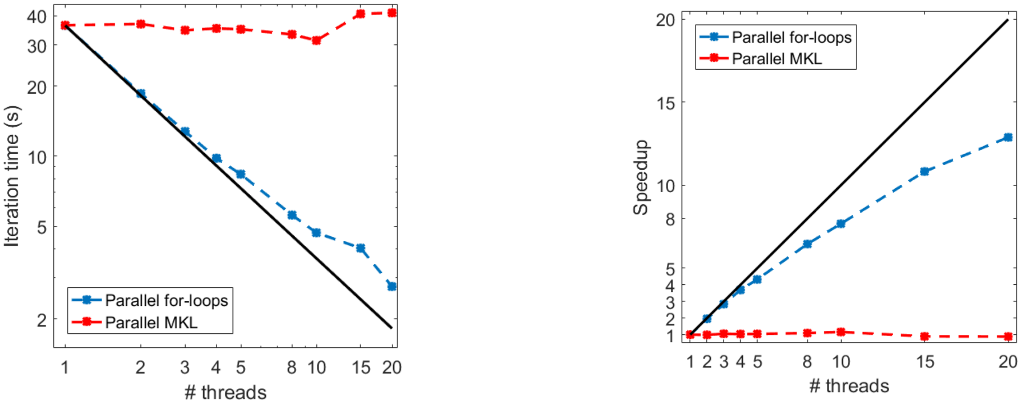
However, to obtain accurate results, all the energy calculations are performed on the same (finer) grid size.

In the case of highly localized charges and/or interpolation to very large model sizes, to limit the effect of discretization error on the calculated energies, I have implemented a grid refining scheme that checks the total charge of the model on the grid and, if the error is larger than a predefined threshold, regenerates the same model charge on a finer grid. In this case, I also interpolate the reference potential onto the fine grid for MSE calculations.

I have also implemented several internal checks in SLABCC to ensure the validity of the results. I check for large changes in the model parameters after the optimization, the significant non-linearity of the model energy during the $E_{isolated}$ extrapolation process, and the significant delocalization of the extra charge. The code can also output several diagnostic files from each part of its calculation, e.g. differences in the charge distribution and potential of input files, model charge distribution in CHGCAR format, potential generated by the model charge in LOCPOT format, generated dielectric profile along the normal axis to the surface, as well as the planar average of these files along each axis. These files can be used for manual checking of the correctness of the SLABCC and its development.

A1.5 Parallel scaling

Parallel scaling of SLABCC ver. 0.3.3 constructing a model with a 170×170×240 grid size, for a slab normal to the 3rd dimension on a 2-socket Intel Xeon E5-2660 v3 (2×10 cores in total) machine (Haswell architecture) is shown in Figure 41. To prevent non-local data access, I ensured that in parallel regions of the code, all data is allocated and accessed by the threads pinned to the same NUMA domain. The code is linked to the Intel MKL for BLAS operations.



a **b**

Fig. 41 a) iteration time and b) speedup of the SLABCC in strong scaling on a single node across two NUMA[9] domains with scatter binding. The solid black line corresponds to ideal scaling.

Most of the time in my SLABCC code is spent in the Poisson solver. I tried two parallelization schemes in the 3D Poisson solver to utilize all the cores of the machine. The red line in Figure 41b shows the speedup when using the threaded MKL for solving each of the 240×240 matrix equations in the direction normal to the slab surface. The blue line shows the speedup when using OpenMP for solving multiple equations in parallel, each on a single thread. As the matrices are fairly small, using more threads for solving them does not improve the performance, and using multiple threads to solve multiple equations in parallel is a better approach with lower computational overhead. Due to a large part of my code being memory-bound and also a possible load imbalance rooted in Armadillo's adaptive solver implementation [292], there is a deviation from the ideal speedup (black line).

A1.6 Distribution

I have published the full source code of SLABCC, including build configurations (makefile) and a detailed manual, under a permissive license (BSD 2-Clause) on GitHub.[10] To increase the accessibility of SLABCC, the codebase is also mirrored on Codeberg[11], which is built on a fully open-source collaborative version control system (Gitea/Forgejo). I use automatic builds and

---

[9] Non-uniform memory access

[10] https://github.com/MFTabriz/slabcc

[11] https://codeberg.org/meisam/slabcc

integration tests on Continuous Integration (CI) for quality assurance. An extensive test set and expected results are also provided at Zenodo.[12]

A1.7 Sample input

Sample SLABCC input file for energy correction calculation using a model with two localized Gaussian charges. Lines starting with # are treated as comments.

```
# PATH TO THE VASP FILES
CHGCAR_neutral = ../01-neutral/CHGCAR
LOCPOT_neutral = ../01-neutral/LOCPOT
CHGCAR_charged = ../02-positive/CHGCAR
LOCPOT_charged = ../02-positive/LOCPOT
# CENTER OF THE LOCALIZED GAUSSIAN CHARGES
charge_position = 0.22573  0.70587  0.33316; 0.54513  0.66851  0.33321
# DIELECTRIC CONSTANT INSIDE THE SLAB
diel_in = 4.79
# DIRECTION NORMAL TO THE SURFACE
normal_direction = b
# POSITION OF THE SLAB INTERFACES
interfaces = 0.16  0.71
```

---

[12] https://doi.org/10.5281/zenodo.1323559

Bibliography

1. Dionysiou, D.D., G. Li Puma, J. Ye, J. Schneider, and D. Bahnemann, *Photocatalysis, Applications*. RSC Energy and Environment Series No. 15, ed. H. Frei, R. Rinaldi, and T.S. Zhao. 2016, Cambridge, UK: The Royal Society of Chemistry.
2. Schirmer, O.F., *O−bound small polarons in oxide materials.* Journal of Physics: Condensed Matter, 2006. **18**(43): p. R667.
3. Komsa, H.-P., T.T. Rantala, and A. Pasquarello, *Finite-size supercell correction schemes for charged defect calculations.* Physical Review B, 2012. **86**(4): p. 045112.
4. Green, I.X., W. Tang, M. Neurock, and J.T. Yates, *Spectroscopic Observation of Dual Catalytic Sites During Oxidation of CO on a Au/TiO_2 Catalyst.* Science, 2011. **333**(6043): p. 736-739.
5. Vannice, M.A., *Titania-supported metals as CO hydrogenation catalysts.* Journal of Catalysis, 1982. **74**(1): p. 199-202.
6. Sastre, F., M. Oteri, A. Corma, and H. Garcia, *Photocatalytic water gas shift using visible or simulated solar light for the efficient, room-temperature hydrogen generation.* Energy & Environmental Science, 2013. **6**(7): p. 2211-2215.
7. Boccuzzi, F., E. Guglielminotti, G. Martra, and G. Cerrato, *Nitric Oxide Reduction by CO on Cu/TiO_2 Catalysts.* Journal of Catalysis, 1994. **146**(2): p. 449-459.
8. Valden, M., X. Lai, and D.W. Goodman, *Onset of Catalytic Activity of Gold Clusters on Titania with the Appearance of Nonmetallic Properties.* Science, 1998. **281**(5383): p. 1647-1650.
9. Comotti, M., W.-C. Li, B. Spliethoff, and F. Schüth, *Support Effect in High Activity Gold Catalysts for CO Oxidation.* Journal of the American Chemical Society, 2006. **128**(3): p. 917-924.
10. Xu, M., Y. Gao, E.M. Moreno, M. Kunst, M. Muhler, Y. Wang, H. Idriss, and C. Wöll, *Photocatalytic Activity of Bulk TiO_2 Anatase and Rutile Single Crystals Using Infrared Absorption Spectroscopy.* Physical Review Letters, 2011. **106**(13): p. 138302.

11. Wu, X., A. Selloni, and S.K. Nayak, *First principles study of CO oxidation on TiO_2(110): The role of surface oxygen vacancies.* The Journal of Chemical Physics, 2004. **120**(9): p. 4512-4516.
12. Petrik, N.G. and G.A. Kimmel, *Off-Normal CO_2 Desorption from the Photooxidation of CO on Reduced TiO_2(110).* The Journal of Physical Chemistry Letters, 2010. **1**(17): p. 2508-2513.
13. Wanbayor, R., P. Deák, T. Frauenheim, and V. Ruangpornvisuti, *First principles theoretical study of the hole-assisted conversion of CO to CO_2 on the anatase TiO_2(101) surface.* The Journal of Chemical Physics, 2011. **134**(10): p. 104701.
14. Zhang, Z. and J.T. Yates, *Electron-Mediated CO Oxidation on the TiO_2(110) Surface during Electronic Excitation.* Journal of the American Chemical Society, 2010. **132**(37): p. 12804-12807.
15. Suppan, P., *Chemistry and Light*. 1994: The Royal Society of Chemistry. p5.
16. Herrmann, J.M., *Photocatalysis*, in *Kirk-Othmer Encyclopedia of Chemical Technology*. 2017, John Wiley & Sons, Inc.
17. Ohtani, B., *Titania Photocatalysis beyond Recombination: A Critical Review.* Catalysts, 2013. **3**(4): p. 942.
18. Bak, T., J. Nowotny, N.J. Sucher, and E. Wachsman, *Effect of Crystal Imperfections on Reactivity and Photoreactivity of TiO_2 (Rutile) with Oxygen, Water, and Bacteria.* The Journal of Physical Chemistry C, 2011. **115**(32): p. 15711-15738.
19. Kennedy, D.R., M. Ritchie, and J. Mackenzie, *The photosorption of oxygen and nitric oxide on titanium dioxide.* Transactions of the Faraday Society, 1958. **54**(0): p. 119-129.
20. Kato, S.-i. and F. Mashio, *Titanium Dioxide-Photocatalyzed Liquid Phase Oxidation of Tetralin.* The Journal of the Society of Chemical Industry, Japan, 1964. **67**(8): p. 1136-1140.
21. Fujishima, A. and K. Honda, *Electrochemical Photolysis of Water at a Semiconductor Electrode.* Nature, 1972. **238**(5358): p. 37-38.
22. Schneider, J., M. Matsuoka, M. Takeuchi, J. Zhang, Y. Horiuchi, M. Anpo, and D.W. Bahnemann, *Understanding TiO_2 Photocatalysis: Mechanisms and Materials.* Chemical Reviews, 2014. **114**(19): p. 9919-9986.

23. Carp, O., C.L. Huisman, and A. Reller, *Photoinduced reactivity of titanium dioxide.* Progress in Solid State Chemistry, 2004. **32**(1): p. 33-177.

24. Linsebigler, A.L., G. Lu, and J.T. Yates, *Photocatalysis on TiO_2 Surfaces: Principles, Mechanisms, and Selected Results.* Chemical Reviews, 1995. **95**(3): p. 735-758.

25. Nursam, N.M., X. Wang, and R.A. Caruso, *High-Throughput Synthesis and Screening of Titania-Based Photocatalysts.* ACS Combinatorial Science, 2015. **17**(10): p. 548-569.

26. Hoffmann, M.R., S.T. Martin, W. Choi, and D.W. Bahnemann, *Environmental Applications of Semiconductor Photocatalysis.* Chemical Reviews, 1995. **95**(1): p. 69-96.

27. Fagan, R., D.E. McCormack, D.D. Dionysiou, and S.C. Pillai, *A review of solar and visible light active TiO_2 photocatalysis for treating bacteria, cyanotoxins and contaminants of emerging concern.* Materials Science in Semiconductor Processing, 2016. **42**: p. 2-14.

28. Bai, Y., I. Mora-Seró, F. De Angelis, J. Bisquert, and P. Wang, *Titanium Dioxide Nanomaterials for Photovoltaic Applications.* Chemical Reviews, 2014. **114**(19): p. 10095-10130.

29. Bai, J. and B. Zhou, *Titanium Dioxide Nanomaterials for Sensor Applications.* Chemical Reviews, 2014. **114**(19): p. 10131-10176.

30. Fujishima, A., T.N. Rao, and D.A. Tryk, *Titanium dioxide photocatalysis.* Journal of Photochemistry and Photobiology C: Photochemistry Reviews, 2000. **1**(1): p. 1-21.

31. Chen, X. and S.S. Mao, *Titanium Dioxide Nanomaterials: Synthesis, Properties, Modifications, and Applications.* Chemical Reviews, 2007. **107**(7): p. 2891-2959.

32. Beltrán, A., L. Gracia, and J. Andrés, *Density Functional Theory Study of the Brookite Surfaces and Phase Transitions between Natural Titania Polymorphs.* The Journal of Physical Chemistry B, 2006. **110**(46): p. 23417-23423.

33. Howard, C.J., T.M. Sabine, and F. Dickson, *Structural and thermal parameters for rutile and anatase.* Acta Crystallographica Section B, 1991. **47**(4): p. 462-468.

34. Abrahams, S.C. and J.L. Bernstein, *Rutile: Normal Probability Plot Analysis and Accurate Measurement of Crystal Structure.* The Journal of Chemical Physics, 1971. **55**(7): p. 3206-3211.

35. Luttrell, T., S. Halpegamage, J. Tao, A. Kramer, E. Sutter, and M. Batzill, *Why is anatase a better photocatalyst than rutile? - Model studies on epitaxial TiO_2 films.* Scientific Reports, 2014. **4**: p. 4043.

36. Zhang, J., P. Zhou, J. Liu, and J. Yu, *New understanding of the difference of photocatalytic activity among anatase, rutile and brookite TiO_2.* Physical Chemistry Chemical Physics, 2014. **16**(38): p. 20382-20386.

37. Ahmed, A.Y., T.A. Kandiel, T. Oekermann, and D. Bahnemann, *Photocatalytic Activities of Different Well-defined Single Crystal TiO_2 Surfaces: Anatase versus Rutile.* The Journal of Physical Chemistry Letters, 2011. **2**(19): p. 2461-2465.

38. Diebold, U., *The surface science of titanium dioxide.* Surface Science Reports, 2003. **48**(5): p. 53-229.

39. Huang, M.H., G. Naresh, and H.-S. Chen, *Facet-Dependent Electrical, Photocatalytic, and Optical Properties of Semiconductor Crystals and Their Implications for Applications.* ACS Applied Materials & Interfaces, 2018. **10**(1): p. 4-15.

40. Vittadini, A., A. Selloni, F.P. Rotzinger, and M. Grätzel, *Structure and Energetics of Water Adsorbed at TiO_2 Anatase (101) and (001) Surfaces.* Physical Review Letters, 1998. **81**(14): p. 2954-2957.

41. Lazzeri, M., A. Vittadini, and A. Selloni, *Structure and energetics of stoichiometric TiO_2 anatase surfaces.* Physical Review B, 2001. **63**(15): p. 155409.

42. Lazzeri, M., A. Vittadini, and A. Selloni, *Erratum: Structure and energetics of stoichiometric TiO_2 anatase surfaces [Phys. Rev. B 63 , 155409 (2001)].* Physical Review B, 2002. **65**(11): p. 119901.

43. Hebenstreit, W., N. Ruzycki, G.S. Herman, Y. Gao, and U. Diebold, *Scanning tunneling microscopy investigation of the TiO_2 anatase (101) surface.* Physical Review B, 2000. **62**(24): p. R16334-R16336.

44. Diebold, U., N. Ruzycki, G.S. Herman, and A. Selloni, *One step towards bridging the materials gap: surface studies of TiO_2 anatase.* Catalysis Today, 2003. **85**(2): p. 93-100.

45. He, Y., O. Dulub, H. Cheng, A. Selloni, and U. Diebold, *Evidence for the Predominance of Subsurface Defects on Reduced Anatase TiO_2 (101).* Physical Review Letters, 2009. **102**(10): p. 106105.

46. Schelling, P.K. and J.W. Halley, *Localization of polarons: A calculation in the adiabatic approximation.* Physical Review B, 2000. **62**(5): p. 3241-3245.

47. Nowotny, J., M. Radecka, and M. Rekas, *Semiconducting Properties of Undoped TiO_2.* Journal of Physics and Chemistry of Solids, 1997. **58**(6): p. 927-937.

48. Janotti, A., C. Franchini, J.B. Varley, G. Kresse, and C.G. Van de Walle, *Dual behavior of excess electrons in rutile TiO_2.* physica status solidi (RRL) – Rapid Research Letters, 2013. **7**(3): p. 199-203.

49. Moser, S., L. Moreschini, J. Jaćimović, O.S. Barišić, H. Berger, A. Magrez, Y.J. Chang, K.S. Kim, A. Bostwick, E. Rotenberg, L. Forró, and M. Grioni, *Tunable Polaronic Conduction in Anatase TiO_2.* Physical Review Letters, 2013. **110**(19): p. 196403.

50. Spreafico, C. and J. VandeVondele, *The nature of excess electrons in anatase and rutile from hybrid DFT and RPA.* Physical Chemistry Chemical Physics, 2014. **16**(47): p. 26144-26152.

51. Yang, S., A.T. Brant, N.C. Giles, and L.E. Halliburton, *Intrinsic small polarons in rutile TiO_2.* Physical Review B, 2013. **87**(12): p. 125201.

52. Yim, C.M., M.B. Watkins, M.J. Wolf, C.L. Pang, K. Hermansson, and G. Thornton, *Engineering Polarons at a Metal Oxide Surface.* Physical Review Letters, 2016. **117**(11): p. 116402.

53. Deák, P., B. Aradi, and T. Frauenheim, *Polaronic effects in TiO_2 calculated by the HSE06 hybrid functional: Dopant passivation by carrier self-trapping.* Physical Review B, 2011. **83**(15): p. 155207.

54. Wallace, S.K. and K.P. McKenna, *Facet-Dependent Electron Trapping in TiO_2 Nanocrystals.* The Journal of Physical Chemistry C, 2015. **119**(4): p. 1913-1920.

55. Deskins, N.A., R. Rousseau, and M. Dupuis, *Defining the Role of Excess Electrons in the Surface Chemistry of TiO_2.* The Journal of Physical Chemistry C, 2010. **114**(13): p. 5891-5897.

56. Deskins, N.A. and M. Dupuis, *Electron transport via polaron hopping in bulk TiO_2: A density functional theory characterization.* Physical Review B, 2007. **75**(19): p. 195212.

57. Deák, P., J. Kullgren, and T. Frauenheim, *Polarons and oxygen vacancies at the surface of anatase TiO_2.* physica status solidi (RRL) – Rapid Research Letters, 2014. **8**(6): p. 583-586.

58. Setvin, M., C. Franchini, X. Hao, M. Schmid, A. Janotti, M. Kaltak, C.G. Van de Walle, G. Kresse, and U. Diebold, *Direct View at Excess Electrons in TiO_2 Rutile and Anatase.* Physical Review Letters, 2014. **113**(8): p. 086402.

59. Di Valentin, C. and A. Selloni, *Bulk and Surface Polarons in Photoexcited Anatase TiO_2.* The Journal of Physical Chemistry Letters, 2011. **2**(17): p. 2223-2228.

60. Freysoldt, C., B. Grabowski, T. Hickel, J. Neugebauer, G. Kresse, A. Janotti, and C.G. Van de Walle, *First-principles calculations for point defects in solids.* Reviews of Modern Physics, 2014. **86**(1): p. 253-305.

61. Evarestov, R.A., T. Bredow, and K. Jug, *Connection between slab and cluster models for crystalline surfaces.* Physics of the Solid State, 2001. **43**(9): p. 1774-1782.

62. Deák, P., *Choosing Models for Solids.* physica status solidi (b), 2000. **217**(1): p. 9-21.

63. Hine, N.D.M., K. Frensch, W.M.C. Foulkes, and M.W. Finnis, *Supercell size scaling of density functional theory formation energies of charged defects.* Physical Review B, 2009. **79**(2): p. 024112.

64. Makov, G. and M.C. Payne, *Periodic boundary conditions in ab initio calculations.* Physical Review B, 1995. **51**(7): p. 4014-4022.

65. Kumagai, Y. and F. Oba, *Electrostatics-based finite-size corrections for first-principles point defect calculations.* Physical Review B, 2014. **89**(19): p. 195205.

66. Ihm, J., A. Zunger, and M.L. Cohen, *Momentum-space formalism for the total energy of solids.* Journal of Physics C: Solid State Physics, 1979. **12**(21): p. 4409.

67. Freysoldt, C., J. Neugebauer, and C.G. Van de Walle, *Electrostatic interactions between charged defects in supercells.* physica status solidi (b), 2011. **248**(5): p. 1067-1076.

68. Lany, S. and A. Zunger, *Accurate prediction of defect properties in density functional supercell calculations.* Modelling and Simulation in Materials Science and Engineering, 2009. **17**(8): p. 084002.

69. Freysoldt, C., J. Neugebauer, and C.G. Van de Walle, *Fully Ab Initio Finite-Size Corrections for Charged-Defect Supercell Calculations.* Physical Review Letters, 2009. **102**(1): p. 016402.

70. Komsa, H.-P. and A. Pasquarello, *Finite-Size Supercell Correction for Charged Defects at Surfaces and Interfaces.* Physical Review Letters, 2013. **110**(9): p. 095505.

71. Komsa, H.-P., T. Rantala, and A. Pasquarello, *Comparison between various finite-size supercell correction schemes for charged defect calculations.* Physica B: Condensed Matter, 2012. **407**(15): p. 3063-3067.

72. Walsh, A., *Correcting the corrections for charged defects in crystals.* npj Computational Materials, 2021. **7**(1): p. 72.

73. Lany, S. and A. Zunger, *Assessment of correction methods for the band-gap problem and for finite-size effects in supercell defect calculations: Case studies for ZnO and GaAs.* Physical Review B, 2008. **78**(23): p. 235104.

74. Castleton, C.W.M. and S. Mirbt, *Ab initio study of neutral vacancies in InP using supercells and finite size scaling.* Physica B: Condensed Matter, 2003. **340–342**: p. 407-411.

75. Castleton, C.W.M., A. Höglund, and S. Mirbt, *Managing the supercell approximation for charged defects in semiconductors: Finite-size scaling, charge correction factors, the band-gap problem, and the ab initio dielectric constant.* Physical Review B, 2006. **73**(3): p. 035215.

76. Lento, J., J.L. Mozos, and R.M. Nieminen, *Charged point defects in semiconductors and the supercell approximation.* Journal of Physics: Condensed Matter, 2002. **14**(10): p. 2637.

77. Freysoldt, C., J. Neugebauer, A.M.Z. Tan, and R.G. Hennig, *Limitations of empirical supercell extrapolation for calculations of point defects in bulk, at surfaces, and in two-dimensional materials.* Physical Review B, 2022. **105**(1): p. 014103.

78. Ewald, P.P., *Die Berechnung optischer und elektrostatischer Gitterpotentiale.* Annalen der Physik, 1921. **369**(3): p. 253-287.

79. Kantorovich, L.N., *Elimination of the long-range dipole interaction in calculations with periodic boundary conditions.* Physical Review B, 1999. **60**(23): p. 15476-15479.

80. Murphy, S.T. and N.D.M. Hine, *Anisotropic charge screening and supercell size convergence of defect formation energies.* Physical Review B, 2013. **87**(9): p. 094111.

81. Baroni, S. and R. Resta, *Ab initio calculation of the macroscopic dielectric constant in silicon.* Physical Review B, 1986. **33**(10): p. 7017-7021.

82. Kunc, K. and R. Resta, *External Fields in the Self-Consistent Theory of Electronic States: A New Method for Direct Evaluation of Macroscopic and Microscopic Dielectric Response.* Physical Review Letters, 1983. **51**(8): p. 686-689.

83. Boeck, S., C. Freysoldt, A. Dick, L. Ismer, and J. Neugebauer, *The object-oriented DFT program library S/PHI/nX.* Computer Physics Communications, 2011. **182**(3): p. 543-554.

84. Komsa, H.-P., N. Berseneva, A.V. Krasheninnikov, and R.M. Nieminen, *Charged Point Defects in the Flatland: Accurate Formation Energy Calculations in Two-Dimensional Materials.* Physical Review X, 2014. **4**(3): p. 031044.

85. Komsa, H.-P., N. Berseneva, A.V. Krasheninnikov, and R.M. Nieminen, *Erratum: Charged Point Defects in the Flatland: Accurate Formation Energy Calculations in Two-Dimensional Materials [Phys. Rev. X 4, 031044 (2014)].* Physical Review X, 2018. **8**(3): p. 039902.

86. Deutschmann, O., *Modeling and Simulation of Heterogeneous Catalytic Reactions: From the Molecular Process to the Technical System*. 2011: Wiley-VCH Verlag GmbH & Co. KGaA

87. Oba, F. and Y. Kumagai, *Design and exploration of semiconductors from first principles: A review of recent advances.* Applied Physics Express, 2018. **11**(6): p. 060101.

88. Tachikawa, T., M. Fujitsuka, and T. Majima, *Mechanistic Insight into the TiO_2 Photocatalytic Reactions: Design of New Photocatalysts.* The Journal of Physical Chemistry C, 2007. **111**(14): p. 5259-5275.

89. Born, M. and R. Oppenheimer, *Zur Quantentheorie der Molekeln.* Annalen der Physik, 1927. **389**(20): p. 457-484.

90. Motta, M., D.M. Ceperley, G.K.-L. Chan, J.A. Gomez, E. Gull, S. Guo, C.A. Jiménez-Hoyos, T.N. Lan, J. Li, F. Ma, A.J. Millis, N.V. Prokof'ev, U. Ray, G.E. Scuseria, S. Sorella, E.M. Stoudenmire, Q. Sun, I.S. Tupitsyn, S.R. White, D. Zgid, and S. Zhang, *Towards the Solution of the Many-Electron Problem in Real Materials: Equation of State of the Hydrogen Chain with State-of-the-Art Many-Body Methods.* Physical Review X, 2017. **7**(3): p. 031059.

91. Nørskov, J.K., F. Abild-Pedersen, F. Studt, and T. Bligaard, *Density functional theory in surface chemistry and catalysis.* Proceedings of the National Academy of Sciences, 2011. **108**(3): p. 937-943.

92. Nørskov, J.K., T. Bligaard, J. Rossmeisl, and C.H. Christensen, *Towards the computational design of solid catalysts.* Nature Chemistry, 2009. **1**: p. 37.

93. Gross, E.K.U. and R.M. Dreizler, *Density functional theory*. NATO Advanced Science Institutes Series. 1995, Boston, MA: Springer US : Imprint : Springer.

94. Hohenberg, P. and W. Kohn, *Inhomogeneous Electron Gas.* Physical Review, 1964. **136**(3B): p. B864-B871.

95. Levy, M., *Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem.* Proceedings of the National Academy of Sciences, 1979. **76**(12): p. 6062-6065.

96. Perdew, J.P., R.G. Parr, M. Levy, and J.L. Balduz, *Density-Functional Theory for Fractional Particle Number: Derivative Discontinuities of the Energy.* Physical Review Letters, 1982. **49**(23): p. 1691-1694.

97. Chayes, J.T., L. Chayes, and M.B. Ruskai, *Density functional approach to quantum lattice systems.* Journal of Statistical Physics, 1985. **38**(3): p. 497-518.

98. Levy, M., *Electron densities in search of Hamiltonians.* Physical Review A, 1982. **26**(3): p. 1200-1208.

99. Kohn, W. and L.J. Sham, *Self-Consistent Equations Including Exchange and Correlation Effects.* Physical Review, 1965. **140**(4A): p. A1133-A1138.

100. Yang, C., J.C. Meza, and L.-W. Wang, *A constrained optimization algorithm for total energy minimization in electronic structure calculations.* Journal of Computational Physics, 2006. **217**(2): p. 709-721.

101. Weber, V., J. VandeVondele, J. Hutter, and A.M.N. Niklasson, *Direct energy functional minimization under orthogonality constraints.* The Journal of Chemical Physics, 2008. **128**(8): p. 084113.

102. Car, R. and M. Parrinello, *Unified Approach for Molecular Dynamics and Density-Functional Theory.* Physical Review Letters, 1985. **55**(22): p. 2471-2474.

103. Lee, M., K. Leiter, C. Eisner, J. Crone, and J. Knap, *Extended Hückel and Slater's rule initial guess for real space grid-based density functional theory.* Computational and Theoretical Chemistry, 2015. **1062**: p. 24-29.

104. Temperton, C., *Direct methods for the solution of the discrete Poisson equation: Some comparisons.* Journal of Computational Physics, 1979. **31**(1): p. 1-20.

105. Greengard, L. and V. Rokhlin, *A fast algorithm for particle simulations.* Journal of Computational Physics, 1987. **73**(2): p. 325-348.

106. Brandt, A., *Multi-level adaptive solutions to boundary-value problems.* Mathematics of Computation, 1977. **31**(138): p. 333-333.

107. García-Risueño, P., J. Alberdi-Rodriguez, M.J.T. Oliveira, X. Andrade, M. Pippig, J. Muguerza, A. Arruabarrena, and A. Rubio, *A survey of the parallel performance and accuracy of Poisson solvers for electronic structure calculations.* Journal of Computational Chemistry, 2014. **35**(6): p. 427-444.

108. Payne, M.C., M.P. Teter, D.C. Allan, T.A. Arias, and J.D. Joannopoulos, *Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients.* Reviews of Modern Physics, 1992. **64**(4): p. 1045-1097.

109. Wood, D.M. and A. Zunger, *A new method for diagonalising large matrices.* Journal of Physics A: Mathematical and General, 1985. **18**(9): p. 1343-1359.

110. Zhou, Y., Y. Saad, M.L. Tiago, and J.R. Chelikowsky, *Self-consistent-field calculations using Chebyshev-filtered subspace iteration.* Journal of Computational Physics, 2006. **219**(1): p. 172-184.

111. Davidson, E.R., *The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices.* Journal of Computational Physics, 1975. **17**(1): p. 87-94.

112. Knyazev, A.V., *Toward the Optimal Preconditioned Eigensolver: Locally Optimal Block Preconditioned Conjugate Gradient Method.* SIAM Journal on Scientific Computing, 2001. **23**(2): p. 517-541.

113. Auckenthaler, T., V. Blum, H.J. Bungartz, T. Huckle, R. Johanni, L. Krämer, B. Lang, H. Lederer, and P.R. Willems, *Parallel solution of partial symmetric eigenvalue problems from electronic structure calculations.* Parallel Computing, 2011. **37**(12): p. 783-794.

114. Kresse, G. and J. Furthmüller, *Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set.* Physical Review B, 1996. **54**(16): p. 11169-11186.

115. Broyden, C.G., *A Class of Methods for Solving Nonlinear Simultaneous Equations.* Mathematics of Computation, 1965. **19**(92): p. 577-593.

116. Anderson, D.G., *Iterative Procedures for Nonlinear Integral Equations.* J. ACM, 1965. **12**(4): p. 547–560.

117. Pulay, P., *Convergence acceleration of iterative sequences. the case of scf iteration.* Chemical Physics Letters, 1980. **73**(2): p. 393-398.

118. Wagner, L.O., E.M. Stoudenmire, K. Burke, and S.R. White, *Guaranteed Convergence of the Kohn-Sham Equations.* Physical Review Letters, 2013. **111**(9): p. 093003.

119. Gonze, X., J.M. Beuken, R. Caracas, F. Detraux, M. Fuchs, G.M. Rignanese, L. Sindic, M. Verstraete, G. Zerah, F. Jollet, M. Torrent, A. Roy, M. Mikami, P. Ghosez, J.Y. Raty, and D.C. Allan, *First-principles computation of material properties: the ABINIT software project.* Computational Materials Science, 2002. **25**(3): p. 478-492.

120. Clark Stewart, J., D. Segall Matthew, J. Pickard Chris, J. Hasnip Phil, I.J. Probert Matt, K. Refson, and C. Payne Mike, *First principles methods using CASTEP*, in *Zeitschrift für Kristallographie - Crystalline Materials*. 2005. p. 567.

121. VandeVondele, J., M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, *Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach.* Computer Physics Communications, 2005. **167**(2): p. 103-128.

122. Dovesi, R., A. Erba, R. Orlando, C.M. Zicovich-Wilson, B. Civalleri, L. Maschio, M. Rérat, S. Casassa, J. Baima, S. Salustro, and B. Kirtman, *Quantum-mechanical condensed matter simulations with CRYSTAL.* WIREs Computational Molecular Science, 2018. **8**(4): p. e1360.

123. Blum, V., R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, *Ab initio molecular simulations with numeric atom-centered orbitals.* Computer Physics Communications, 2009. **180**(11): p. 2175-2196.

124. Enkovaara, J., C. Rostgaard, J.J. Mortensen, J. Chen, M. Dułak, L. Ferrighi, J. Gavnholt, C. Glinsvad, V. Haikola, H.A. Hansen, H.H. Kristoffersen, M. Kuisma, A.H. Larsen, L. Lehtovaara, M. Ljungberg, O. Lopez-Acevedo, P.G. Moses, J. Ojanen, T. Olsen, V. Petzold, N.A. Romero, J. Stausholm-Møller, M. Strange, G.A. Tritsaris, M. Vanin, M. Walter, B. Hammer, H. Häkkinen, G.K.H. Madsen, R.M. Nieminen, J.K. Nørskov, M. Puska, T.T. Rantala, J. Schiøtz, K.S. Thygesen, and K.W. Jacobsen, *Electronic structure calculations with GPAW: a real-space implementation of the projector augmented-wave method.* Journal of Physics: Condensed Matter, 2010. **22**(25): p. 253202.

125. Giannozzi, P., S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G.L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S.

Scandolo, G. Sclauzero, A.P. Seitsonen, A. Smogunov, P. Umari, and R.M. Wentzcovitch, *QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials.* Journal of Physics: Condensed Matter, 2009. **21**(39): p. 395502.

126. Soler, J.M., E. Artacho, J.D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal, *The SIESTA method forab initioorder-Nmaterials simulation.* Journal of Physics: Condensed Matter, 2002. **14**(11): p. 2745-2779.
127. Kresse, G. and D. Joubert, *From ultrasoft pseudopotentials to the projector augmented-wave method.* Physical Review B, 1999. **59**(3): p. 1758-1775.
128. Schwarz, K. and P. Blaha, *Solid state calculations using WIEN2k.* Computational Materials Science, 2003. **28**(2): p. 259-273.
129. Baldereschi, A., *Mean-Value Point in the Brillouin Zone.* Physical Review B, 1973. **7**(12): p. 5212-5215.
130. Chadi, D.J. and M.L. Cohen, *Special Points in the Brillouin Zone.* Physical Review B, 1973. **8**(12): p. 5747-5753.
131. Monkhorst, H.J. and J.D. Pack, *Special points for Brillouin-zone integrations.* Physical Review B, 1976. **13**(12): p. 5188-5192.
132. Pack, J.D. and H.J. Monkhorst, *"Special points for Brillouin-zone integrations"---a reply.* Physical Review B, 1977. **16**(4): p. 1748-1749.
133. Wisesa, P., K.A. McGill, and T. Mueller, *Efficient generation of generalized Monkhorst-Pack grids through the use of informatics.* Physical Review B, 2016. **93**(15): p. 155109.
134. Hart, G.L.W., J.J. Jorgensen, W.S. Morgan, and R.W. Forcade, *A robust algorithm for k-point grid generation and symmetry reduction.* Journal of Physics Communications, 2019. **3**(6): p. 065009.
135. Blöchl, P.E., O. Jepsen, and O.K. Andersen, *Improved tetrahedron method for Brillouin-zone integrations.* Physical Review B, 1994. **49**(23): p. 16223-16233.
136. Lehmann, G. and M. Taut, *On the Numerical Calculation of the Density of States and Related Properties.* physica status solidi (b), 1972. **54**(2): p. 469-477.
137. Feynman, R.P., *Forces in Molecules.* Physical Review, 1939. **56**(4): p. 340-343.
138. Pulay, P., *Ab initio calculation of force constants and equilibrium geometries in polyatomic molecules.* Molecular Physics, 1969. **17**(2): p. 197-204.

139. Segall, M.D., C.J. Pickard, R. Shah, and M.C. Payne, *Population analysis in plane wave electronic structure calculations.* Molecular Physics, 1996. **89**(2): p. 571-577.

140. Jaffe, J.E. and A.C. Hess, *Gaussian basis density functional theory for systems periodic in two or three dimensions: Energy and forces.* The Journal of Chemical Physics, 1996. **105**(24): p. 10983-10998.

141. Slater, J.C., *Wave Functions in a Periodic Potential.* Physical Review, 1937. **51**(10): p. 846-851.

142. Andersen, O.K., *Linear methods in band theory.* Physical Review B, 1975. **12**(8): p. 3060-3083.

143. Hamann, D.R., *Generalized norm-conserving pseudopotentials.* Physical Review B, 1989. **40**(5): p. 2980-2987.

144. Vanderbilt, D., *Soft self-consistent pseudopotentials in a generalized eigenvalue formalism.* Physical Review B, 1990. **41**(11): p. 7892-7895.

145. Blöchl, P.E., *Projector augmented-wave method.* Physical Review B, 1994. **50**(24): p. 17953-17979.

146. Lieb, E.H., *Density functionals for coulomb systems.* International Journal of Quantum Chemistry, 1983. **24**(3): p. 243-277.

147. Mardirossian, N. and M. Head-Gordon, *Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals.* Molecular Physics, 2017. **115**(19): p. 2315-2372.

148. Peverati, R. and D.G. Truhlar, *Quest for a universal density functional: the accuracy of density functionals across a broad spectrum of databases in chemistry and physics.* Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2014. **372**(2011).

149. Barth, U.v. and L. Hedin, *A local exchange-correlation potential for the spin polarized case. i.* Journal of Physics C: Solid State Physics, 1972. **5**(13): p. 1629.

150. Ortiz, G. and P. Ballone, *Correlation energy, structure factor, radial distribution function, and momentum distribution of the spin-polarized uniform electron gas.* Physical Review B, 1994. **50**(3): p. 1391-1405.

151. Jones, R.O. and O. Gunnarsson, *The density functional formalism, its applications and prospects.* Reviews of Modern Physics, 1989. **61**(3): p. 689-746.

152. Herman, F., J.P. Van Dyke, and I.B. Ortenburger, *Improved Statistical Exchange Approximation for Inhomogeneous Many-Electron Systems.* Physical Review Letters, 1969. **22**(16): p. 807-811.

153. Langreth, D.C. and M.J. Mehl, *Beyond the local-density approximation in calculations of ground-state electronic properties.* Physical Review B, 1983. **28**(4): p. 1809-1834.

154. Ziesche, P., S. Kurth, and J.P. Perdew, *Density functionals from LDA to GGA.* Computational Materials Science, 1998. **11**(2): p. 122-127.

155. Lieb, E.H. and S. Oxford, *Improved lower bound on the indirect Coulomb energy.* International Journal of Quantum Chemistry, 1981. **19**(3): p. 427-439.

156. Odashima, M.M. and K. Capelle, *How tight is the Lieb-Oxford bound?* The Journal of Chemical Physics, 2007. **127**(5): p. 054106.

157. Zupan, A., J.P. Perdew, K. Burke, and M. Causà, *Density-gradient analysis for density functional theory: Application to atoms.* International Journal of Quantum Chemistry, 1997. **61**(5): p. 835-845.

158. Nagy, Á., *Hierarchy of equations in the generalized density functional theory.* International Journal of Quantum Chemistry, 2006. **106**(5): p. 1043-1051.

159. Perdew, J.P., A. Ruzsinszky, J. Tao, V.N. Staroverov, G.E. Scuseria, and G.I. Csonka, *Prescription for the design and selection of density functional approximations: More constraint satisfaction with fewer fits.* The Journal of Chemical Physics, 2005. **123**(6): p. 062201.

160. Cohen, A.J., P. Mori-Sánchez, and W. Yang, *Challenges for Density Functional Theory.* Chemical Reviews, 2012. **112**(1): p. 289-320.

161. Medvedev, M.G., I.S. Bushmarinov, J. Sun, J.P. Perdew, and K.A. Lyssenko, *Density functional theory is straying from the path toward the exact functional.* Science, 2017. **355**(6320): p. 49-52.

162. Ruzsinszky, A. and J.P. Perdew, *Twelve outstanding problems in ground-state density functional theory: A bouquet of puzzles.* Computational and Theoretical Chemistry, 2011. **963**(1): p. 2-6.

163. Perdew, J.P. and A. Zunger, *Self-interaction correction to density-functional approximations for many-electron systems.* Physical Review B, 1981. **23**(10): p. 5048-5079.

164. Mori-Sánchez, P., A.J. Cohen, and W. Yang, *Many-electron self-interaction error in approximate density functionals.* The Journal of Chemical Physics, 2006. **125**(20): p. 201102.

165. Li, C., X. Zheng, N.Q. Su, and W. Yang, *Localized orbital scaling correction for systematic elimination of delocalization error in density functional approximations.* National Science Review, 2018. **5**(2): p. 203-215.

166. Sai, N., P.F. Barbara, and K. Leung, *Hole Localization in Molecular Crystals from Hybrid Density Functional Theory.* Physical Review Letters, 2011. **106**(22): p. 226403.

167. Deák, P., A. Gali, B. Aradi, and T. Frauenheim, *Accurate gap levels and their role in the reliability of other calculated defect properties.* physica status solidi (b), 2011. **248**(4): p. 790-798.

168. Zhang, Y. and W. Yang, *A challenge for density functionals: Self-interaction error increases for systems with a noninteger number of electrons.* The Journal of Chemical Physics, 1998. **109**(7): p. 2604-2608.

169. Hamel, S., P. Duffy, M.E. Casida, and D.R. Salahub, *Kohn–Sham orbitals and orbital energies: fictitious constructs but good approximations all the same.* Journal of Electron Spectroscopy and Related Phenomena, 2002. **123**(2): p. 345-363.

170. Bickelhaupt, F.M. and E.J. Baerends, *Kohn-Sham Density Functional Theory: Predicting and Understanding Chemistry*, in *Reviews in Computational Chemistry*, K.B. Lipkowitz and D.B. Boyd, Editors. 2000, Wiley-VCH, Inc.

171. Janak, J.F., *Proof that* $\partial E\partial n_i=\varepsilon$ *in density-functional theory.* Physical Review B, 1978. **18**(12): p. 7165-7168.

172. Dabo, I., A. Ferretti, N. Poilvert, Y. Li, N. Marzari, and M. Cococcioni, *Koopmans' condition for density-functional theory.* Physical Review B, 2010. **82**(11): p. 115121.

173. Cohen, A.J., P. Mori-Sánchez, and W. Yang, *Fractional charge perspective on the band gap in density-functional theory.* Physical Review B, 2008. **77**(11): p. 115123.

174. Yang, W., A.J. Cohen, and P. Mori-Sánchez, *Derivative discontinuity, bandgap and lowest unoccupied molecular orbital in density functional theory.* The Journal of Chemical Physics, 2012. **136**(20): p. 204111.

175. Perdew, J.P. and M. Levy, *Physical Content of the Exact Kohn-Sham Orbital Energies: Band Gaps and Derivative Discontinuities.* Physical Review Letters, 1983. **51**(20): p. 1884-1887.

176. Li, C., X. Zheng, A.J. Cohen, P. Mori-Sánchez, and W. Yang, *Local Scaling Correction for Reducing Delocalization Error in Density Functional Approximations.* Physical Review Letters, 2015. **114**(5): p. 053001.

177. Koopmans, T., *Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den Einzelnen Elektronen Eines Atoms.* Physica, 1934. **1**(1–6): p. 104-113.

178. Almbladh, C.O. and U. von Barth, *Exact results for the charge and spin densities, exchange-correlation potentials, and density-functional eigenvalues.* Physical Review B, 1985. **31**(6): p. 3231-3244.

179. Levy, M., J.P. Perdew, and V. Sahni, *Exact differential equation for the density and ionization energy of a many-particle system.* Physical Review A, 1984. **30**(5): p. 2745-2748.

180. Perdew, J.P. and M. Levy, *Comment on "Significance of the highest occupied Kohn-Sham eigenvalue".* Physical Review B, 1997. **56**(24): p. 16021-16028.

181. Tsuneda, T., J.-W. Song, S. Suzuki, and K. Hirao, *On Koopmans' theorem in density functional theory.* The Journal of Chemical Physics, 2010. **133**(17): p. 174101.

182. Stein, T., J. Autschbach, N. Govind, L. Kronik, and R. Baer, *Curvature and Frontier Orbital Energies in Density Functional Theory.* The Journal of Physical Chemistry Letters, 2012. **3**(24): p. 3740-3744.

183. Savin, A., F. Colonna, and R. Pollet, *Adiabatic connection approach to density functional theory of electronic systems.* International Journal of Quantum Chemistry, 2003. **93**(3): p. 166-190.

184. Burke, K., M. Ernzerhof, and J.P. Perdew, *The adiabatic connection method: a non-empirical hybrid.* Chemical Physics Letters, 1997. **265**(1): p. 115-120.

185. Gritsenko, O.V., R. Van Leeuwen, and E.J. Baerends, *On the optimal mixing of the exchange energy and the electron-electron interaction part of the exchange-correlation energy.* International Journal of Quantum Chemistry, 1996. **60**(7): p. 1375-1384.

186. Paier, J., M. Marsman, K. Hummer, G. Kresse, I.C. Gerber, and J.G. Ángyán, *Screened hybrid density functionals applied to solids.* The Journal of Chemical Physics, 2006. **124**(15): p. 154709.

187. Janesko, B.G., T.M. Henderson, and G.E. Scuseria, *Screened hybrid density functionals for solid-state chemistry and physics.* Physical Chemistry Chemical Physics, 2009. **11**(3): p. 443-454.

188. Heyd, J. and G.E. Scuseria, *Efficient hybrid density functional calculations in solids: Assessment of the Heyd–Scuseria–Ernzerhof screened Coulomb hybrid functional.* The Journal of Chemical Physics, 2004. **121**(3): p. 1187-1192.

189. Deák, P., B. Aradi, T. Frauenheim, E. Janzén, and A. Gali, *Accurate defect levels obtained from the HSE06 range-separated hybrid functional.* Physical Review B, 2010. **81**(15): p. 153203.

190. Laura, E.R., A. Degomme, A.F.-L. José, G. Stefan, and G. Luigi, *Affordable and accurate large-scale hybrid-functional calculations on GPU-accelerated supercomputers.* Journal of Physics: Condensed Matter, 2018. **30**(9): p. 095901.

191. Anisimov, V.I., J. Zaanen, and O.K. Andersen, *Band theory and Mott insulators: Hubbard U instead of Stoner I.* Physical Review B, 1991. **44**(3): p. 943-954.

192. Hubbard, J., *Electron correlations in narrow energy bands.* Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 1963. **276**(1365): p. 238-257.

193. Vladimir, I.A., F. Aryasetiawan, and A.I. Lichtenstein, *First-principles calculations of the electronic structure and spectra of strongly correlated systems: the LDA+U method.* Journal of Physics: Condensed Matter, 1997. **9**(4): p. 767.

194. Janotti, A. and C.G. Van de Walle, *LDA + U and hybrid functional calculations for defects in ZnO, SnO_2, and TiO_2.* physica status solidi (b), 2011. **248**(4): p. 799-804.

195. Kulik, H.J., *Perspective: Treating electron over-delocalization with the DFT+U method.* The Journal of Chemical Physics, 2015. **142**(24): p. 240901.

196. Ylvisaker, E.R., W.E. Pickett, and K. Koepernik, *Anisotropy and magnetism in the LSDA+U method.* Physical Review B, 2009. **79**(3): p. 035103.

197. Liechtenstein, A.I., V.I. Anisimov, and J. Zaanen, *Density-functional theory and strong interactions: Orbital ordering in Mott-Hubbard insulators.* Physical Review B, 1995. **52**(8): p. R5467-R5470.

198. Himmetoglu, B., A. Floris, S. Gironcoli, and M. Cococcioni, *Hubbard-corrected DFT energy functionals: The LDA+U description of correlated systems.* International Journal of Quantum Chemistry, 2014. **114**(1): p. 14-49.

199. Stausholm-Møller, J., H.H. Kristoffersen, B. Hinnemann, G.K.H. Madsen, and B. Hammer, *DFT+U study of defects in bulk rutile TiO_2.* The Journal of Chemical Physics, 2010. **133**(14): p. 144708.

200. Arroyo-de Dompablo, M.E., A. Morales-García, and M. Taravillo, *DFT+U calculations of crystal lattice, electronic structure, and phase stability under pressure of TiO_2 polymorphs.* The Journal of Chemical Physics, 2011. **135**(5): p. 054503.

201. Morgan, B.J. and G.W. Watson, *A Density Functional Theory+U Study of Oxygen Vacancy Formation at the (110), (100), (101), and (001) Surfaces of Rutile TiO_2.* The Journal of Physical Chemistry C, 2009. **113**(17): p. 7322-7328.

202. Taizo, S., Y. Kenji, M. Susanne, and S. Biplab, *A systematic study of polarons due to oxygen vacancy formation at the rutile TiO_2 (110) surface by GGA+U and HSE06 methods.* Journal of Physics: Condensed Matter, 2012. **24**(43): p. 435504.

203. Kulik, H.J., M. Cococcioni, D.A. Scherlis, and N. Marzari, *Density Functional Theory in Transition-Metal Chemistry: A Self-Consistent Hubbard U Approach.* Physical Review Letters, 2006. **97**(10): p. 103001.

204. Solovyev, I.V. and P.H. Dederichs, *Ab initio calculations of Coulomb U parameters for transition-metal impurities.* Physical Review B, 1994. **49**(10): p. 6736-6740.

205. Stevanović, V., S. Lany, X. Zhang, and A. Zunger, *Correcting density functional theory for accurate predictions of compound enthalpies of formation: Fitted elemental-phase reference energies.* Physical Review B, 2012. **85**(11): p. 115104.

206. Hu, Z. and H. Metiu, *Choice of U for DFT+U Calculations for Titanium Oxides.* The Journal of Physical Chemistry C, 2011. **115**(13): p. 5841-5845.

207. Verma, P. and D.G. Truhlar, *Does DFT+U mimic hybrid density functionals?* Theoretical Chemistry Accounts, 2016. **135**(8): p. 182.

208. Wang, Z., C. Brock, A. Matt, and K.H. Bevan, *Implications of the DFT+U method on polaron properties in energy materials.* Physical Review B, 2017. **96**(12): p. 125150.
209. Christensen, N., *Electronic structure of GaAs under strain.* Physical Review B, 1984. **30**(10): p. 5753.
210. Lany, S., H. Raebiger, and A. Zunger, *Magnetic interactions of Cr-Cr and Co-Co impurity pairs in ZnO within a band-gap corrected density functional approach.* Physical Review B, 2008. **77**(24): p. 241201.
211. Lany, S., *Predicting polaronic defect states by means of generalized Koopmans density functional calculations.* physica status solidi (b), 2011. **248**(5): p. 1052-1060.
212. Lany, S., *Semiconducting transition metal oxides.* Journal of Physics: Condensed Matter, 2015. **27**(28): p. 283203.
213. Schlegel, H.B., *Geometry optimization.* Wiley Interdisciplinary Reviews: Computational Molecular Science, 2011. **1**(5): p. 790-809.
214. Maseras, F. and K. Morokuma, *IMOMM: A new integrated ab initio + molecular mechanics geometry optimization scheme of equilibrium structures and transition states.* Journal of Computational Chemistry, 1995. **16**(9): p. 1170-1179.
215. Vreven, T., K. Morokuma, Ö. Farkas, H.B. Schlegel, and M.J. Frisch, *Geometry optimization with QM/MM, ONIOM, and other combined methods. I. Microiterations and constraints.* Journal of Computational Chemistry, 2003. **24**(6): p. 760-769.
216. Sun, J.Q. and K. Ruedenberg, *Quadratic steepest descent on potential energy surfaces. I. Basic formalism and quantitative assessment.* The Journal of Chemical Physics, 1993. **99**(7): p. 5257-5268.
217. Fletcher, R. and C.M. Reeves, *Function minimization by conjugate gradients.* The Computer Journal, 1964. **7**(2): p. 149-154.
218. Filippone, F., S. Meloni, and M. Parrinello, *A novel implicit Newton–Raphson geometry optimization method for density functional theory calculations.* The Journal of Chemical Physics, 2001. **115**(2): p. 636-642.
219. Császár, P. and P. Pulay, *Geometry optimization by direct inversion in the iterative subspace.* Journal of Molecular Structure, 1984. **114**: p. 31-34.
220. Murnaghan, F.D., *The Compressibility of Media under Extreme Pressures.* Proceedings of the National Academy of Sciences, 1944. **30**(9): p. 244-247.

221. Heidrich, D., *An Introduction to the Nomenclature and Usage of the Reaction Path Concept*, in *The Reaction Path in Chemistry: Current Approaches and Perspectives*, D. Heidrich, Editor. 1995, Springer Netherlands: Dordrecht. p. 1-10.

222. McIver, J.W. and A. Komornicki, *Structure of transition states in organic reactions. General theory and an application to the cyclobutene-butadiene isomerization using a semiempirical molecular orbital method.* Journal of the American Chemical Society, 1972. **94**(8): p. 2625-2633.

223. Schlegel, H.B., *Optimization of equilibrium geometries and transition structures.* Journal of Computational Chemistry, 1982. **3**(2): p. 214-218.

224. E, W., W. Ren, and E. Vanden-Eijnden, *String method for the study of rare events.* Physical Review B, 2002. **66**(5): p. 052301.

225. Peters, B., A. Heyden, A.T. Bell, and A. Chakraborty, *A growing string method for determining transition states: Comparison to the nudged elastic band and string methods.* The Journal of Chemical Physics, 2004. **120**(17): p. 7877-7886.

226. Jónsson, H., G. Mills, and K.W. Jacobsen, *Nudged elastic band method for finding minimum energy paths of transitions*, in *Classical and Quantum Dynamics in Condensed Phase Simulations*, B.J. Berne, G. Ciccotti, and D.F. Coker, Editors. 1998, World Scientific. p. 385-404.

227. Elber, R. and M. Karplus, *A method for determining reaction paths in large molecules: Application to myoglobin.* Chemical Physics Letters, 1987. **139**(5): p. 375-380.

228. Gillilan, R.E. and K.R. Wilson, *Shadowing, rare events, and rubber bands. A variational Verlet algorithm for molecular dynamics.* The Journal of Chemical Physics, 1992. **97**(3): p. 1757-1772.

229. Mills, G. and H. Jónsson, *Quantum and thermal effects in* H_2 *dissociative adsorption: Evaluation of free energy barriers in multidimensional quantum systems.* Physical Review Letters, 1994. **72**(7): p. 1124-1127.

230. Henkelman, G. and H. Jónsson, *Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points.* The Journal of Chemical Physics, 2000. **113**(22): p. 9978-9985.

231. Sheppard, D. and G. Henkelman, *Paths to which the nudged elastic band converges.* Journal of Computational Chemistry, 2011. **32**(8): p. 1769-1771.

232. Sheppard, D., R. Terrell, and G. Henkelman, *Optimization methods for finding minimum energy paths.* The Journal of Chemical Physics, 2008. **128**(13): p. 134106.

233. Herbol, H.C., J. Stevenson, and P. Clancy, *Computational Implementation of Nudged Elastic Band, Rigid Rotation, and Corresponding Force Optimization.* Journal of Chemical Theory and Computation, 2017. **13**(7): p. 3250-3259.

234. Henkelman, G., B.P. Uberuaga, and H. Jónsson, *A climbing image nudged elastic band method for finding saddle points and minimum energy paths.* The Journal of Chemical Physics, 2000. **113**(22): p. 9901-9904.

235. Monserrat, B., *Electron–phonon coupling from finite differences.* Journal of Physics: Condensed Matter, 2018. **30**(8): p. 083001.

236. Hafner, J., *Materials simulations using VASP—a quantum perspective to materials science.* Computer Physics Communications, 2007. **177**(1): p. 6-13.

237. Hafner, J., *Ab-initio simulations of materials using VASP: Density-functional theory and beyond.* Journal of Computational Chemistry, 2008. **29**(13): p. 2044-2078.

238. Sun, G., J. Kürti, P. Rajczy, M. Kertesz, J. Hafner, and G. Kresse, *Performance of the Vienna ab initio simulation package (VASP) in chemical applications.* Journal of Molecular Structure: THEOCHEM, 2003. **624**(1): p. 37-45.

239. Wende, F., M. Marsman, J. Kim, F. Vasilev, Z. Zhao, and T. Steinke, *OpenMP in VASP: Threading and SIMD.* International Journal of Quantum Chemistry, 2019. **119**(12): p. e25851.

240. Maniopoulou, A., E.R.M. Davidson, R. Grau-Crespo, A. Walsh, I.J. Bush, C.R.A. Catlow, and S.M. Woodley, *Introducing k-point parallelism into VASP.* Computer Physics Communications, 2012. **183**(8): p. 1696-1701.

241. Hacene, M., A. Anciaux-Sedrakian, X. Rozanska, D. Klahr, T. Guignon, and P. Fleurat-Lessard, *Accelerating VASP electronic structure calculations using graphic processing units.* Journal of Computational Chemistry, 2012. **33**(32): p. 2581-2589.

242. Maintz, S., B. Eck, and R. Dronskowski, *Speeding up plane-wave electronic-structure calculations using graphics-processing units.* Computer Physics Communications, 2011. **182**(7): p. 1421-1427.

243. Stegailov, V., G. Smirnov, and V. Vecher, *VASP hits the memory wall: Processors efficiency comparison.* Concurrency and Computation: Practice and Experience, 2019. **31**(19): p. e5136.

244. Shiraishi, K., *A New Slab Model Approach for Electronic Structure Calculation of Polar Semiconductor Surface.* journal of the physical society of japan, 1990. **59**(10): p. 3455-3458.

245. Perdew, J.P., K. Burke, and M. Ernzerhof, *Generalized Gradient Approximation Made Simple.* Physical Review Letters, 1996. **77**(18): p. 3865-3868.

246. Lany, S. and A. Zunger, *Polaronic hole localization and multiple hole binding of acceptors in oxide wide-gap semiconductors.* Physical Review B, 2009. **80**(8): p. 085202.

247. Bengone, O., M. Alouani, P. Blöchl, and J. Hugel, *Implementation of the projector augmented-wave LDA+U method: Application to the electronic structure of NiO.* Physical Review B, 2000. **62**(24): p. 16392-16401.

248. Nocedal, J., *Updating quasi-Newton matrices with limited storage.* Mathematics of Computation, 1980. **35**(151): p. 773-773.

249. Bitzek, E., P. Koskinen, F. Gähler, M. Moseler, and P. Gumbsch, *Structural Relaxation Made Simple.* Physical Review Letters, 2006. **97**(17): p. 170201.

250. Farzalipour Tabriz, M., B. Aradi, T. Frauenheim, and P. Deák, *SLABCC: Total energy correction code for charged periodic slab models.* Computer Physics Communications, 2019. **240**: p. 101-105.

251. Larson, J., M. Menickelly, and S.M. Wild, *Derivative-free optimization methods.* Acta Numerica, 2019. **28**: p. 287-404.

252. Deák, P., M. Han, M. Lorke, M. Farzalipour Tabriz, and T. Frauenheim, *Intrinsic defects of GaSe.* Journal of Physics: Condensed Matter, 2020. **32**(28): p. 285503.

253. Deák, P., E. Khorasani, M. Lorke, M. Farzalipour Tabriz, B. Aradi, and T. Frauenheim, *Defect calculations with hybrid functionals in layered compounds and in slab models.* Physical Review B, 2019. **100**(23): p. 235304.

254. Cao, T. and A. Bongiorno, *Atomistic Corrective Scheme for Supercell Density Functional Theory Calculations of Charged Defects.* Scientific Reports, 2017. **7**(1): p. 2834.

255. Noh, J.-Y., H. Kim, and Y.-S. Kim, *Stability and electronic structures of native defects in single-layer* MoS_2. Physical Review B, 2014. **89**(20): p. 205417.

256. Gajdoš, M., K. Hummer, G. Kresse, J. Furthmüller, and F. Bechstedt, *Linear optical properties in the projector-augmented wave methodology.* Physical Review B, 2006. **73**(4): p. 045112.

257. Krukau, A.V., O.A. Vydrov, A.F. Izmaylov, and G.E. Scuseria, *Influence of the exchange screening parameter on the performance of screened hybrid functionals.* The Journal of Chemical Physics, 2006. **125**(22): p. 224106.

258. Topsakal, M. and S. Ciraci, *Effects of static charging and exfoliation of layered crystals.* Physical Review B, 2012. **85**(4): p. 045121.

259. Chagas da Silva, M., M. Lorke, B. Aradi, M. Farzalipour Tabriz, T. Frauenheim, A. Rubio, D. Rocca, and P. Deák, *Self-Consistent Potential Correction for Charged Periodic Systems.* Physical Review Letters, 2021. **126**(7): p. 076401.

260. Womack, J.C., L. Anton, J. Dziedzic, P.J. Hasnip, M.I.J. Probert, and C.-K. Skylaris, *DL_MG: A Parallel Multigrid Poisson and Poisson–Boltzmann Solver for Electronic Structure Calculations in Vacuum and Solution.* Journal of Chemical Theory and Computation, 2018. **14**(3): p. 1412-1432.

261. Fisicaro, G., L. Genovese, O. Andreussi, N. Marzari, and S. Goedecker, *A generalized Poisson and Poisson-Boltzmann solver for electrostatic environments.* The Journal of Chemical Physics, 2016. **144**(1).

262. Budiardja, R.D. and C.Y. Cardall, *Parallel FFT-based Poisson solver for isolated three-dimensional systems.* Computer Physics Communications, 2011. **182**(10): p. 2265-2275.

263. Tang, H., F. Lévy, H. Berger, and P.E. Schmid, *Urbach tail of anatase TiO_2.* Physical Review B, 1995. **52**(11): p. 7771-7774.

264. Pascual, J., J. Camassel, and H. Mathieu, *Fine structure in the intrinsic absorption edge of TiO_2.* Physical Review B, 1978. **18**(10): p. 5606-5614.

265. Labat, F., P. Baranek, C. Domain, C. Minot, and C. Adamo, *Density functional theory analysis of the structural and electronic properties of TiO_2 rutile and anatase polytypes: Performances of different exchange-correlation functionals.* The Journal of Chemical Physics, 2007. **126**(15): p. 154703.

266. Mei, Z.-G., Y. Wang, S.-L. Shang, and Z.-K. Liu, *First-Principles Study of Lattice Dynamics and Thermodynamics of TiO_2 Polymorphs.* Inorganic Chemistry, 2011. **50**(15): p. 6996-7003.

267. Estefania, G., F. Ricardo, and W.M. Alvaro, *A DFT+U study on structural, electronic, vibrational and thermodynamic properties of TiO_2 polymorphs and hydrogen titanate: tuning the Hubbard ' U -term'.* Journal of Physics Communications, 2017. **1**(5): p. 055006.

268. Alkauskas, A., P. Broqvist, and A. Pasquarello, *Defect levels through hybrid density functionals: Insights and applications.* physica status solidi (b), 2011. **248**(4): p. 775-789.

269. Deák, P., B. Aradi, and T. Frauenheim, *Quantitative theory of the oxygen vacancy and carrier self-trapping in bulk TiO_2.* Physical Review B, 2012. **86**(19): p. 195206.

270. Shibuya, T., K. Yasuoka, S. Mirbt, and B. Sanyal, *A systematic study of polarons due to oxygen vacancy formation at the rutile TiO_2(110) surface by GGA+U and HSE06 methods.* Journal of Physics: Condensed Matter, 2012. **24**(43): p. 435504.

271. Deák, P., M. Lorke, B. Aradi, and T. Frauenheim, *Optimized hybrid functionals for defect calculations in semiconductors.* Journal of Applied Physics, 2019. **126**(13): p. 130901.

272. Kang, Y., H. Peelaers, and C.G. Van de Walle, *First-principles study of electron-phonon interactions and transport in anatase TiO_2.* Physical Review B, 2019. **100**(12): p. 121113.

273. Selçuk, S. and A. Selloni, *Excess electrons at anatase TiO_2 surfaces and interfaces: insights from first principles simulations.* Journal of Physics D: Applied Physics, 2017. **50**(27): p. 273002.

274. Tersoff, J. and D.R. Hamann, *Theory and Application for the Scanning Tunneling Microscope.* Physical Review Letters, 1983. **50**(25): p. 1998-2001.

275. Tersoff, J. and D.R. Hamann, *Theory of the scanning tunneling microscope.* Physical Review B, 1985. **31**(2): p. 805-813.

276. Selloni, A., P. Carnevali, E. Tosatti, and C.D. Chen, *Voltage-dependent scanning-tunneling microscopy of a crystal surface: Graphite.* Physical Review B, 1985. **31**(4): p. 2602-2605.

277. Farzalipour Tabriz, M., B. Aradi, T. Frauenheim, and P. Deák, *Application of the Lany–Zunger polaron correction for calculating surface charge trapping.* Journal of Physics: Condensed Matter, 2017. **29**(39): p. 394001.

278. Setvin, M., M. Buchholz, W. Hou, C. Zhang, B. Stöger, J. Hulva, T. Simschitz, X. Shi, J. Pavelec, G.S. Parkinson, M. Xu, Y. Wang, M. Schmid, C. Wöll, A. Selloni, and U. Diebold, *A Multitechnique Study of CO Adsorption on the TiO_2 Anatase (101) Surface.* The Journal of Physical Chemistry C, 2015. **119**(36): p. 21044-21052.

279. Fujishima, A., X. Zhang, and D.A. Tryk, *TiO_2 photocatalysis and related surface phenomena.* Surface Science Reports, 2008. **63**(12): p. 515-582.

280. Aschauer, U., J. Chen, and A. Selloni, *Peroxide and superoxide states of adsorbed O_2 on anatase TiO_2 (101) with subsurface defects.* Physical Chemistry Chemical Physics, 2010. **12**(40): p. 12956-12960.

281. Li, Y.-F. and A. Selloni, *Theoretical Study of Interfacial Electron Transfer from Reduced Anatase TiO_2(101) to Adsorbed O_2.* Journal of the American Chemical Society, 2013. **135**(24): p. 9195-9199.

282. Dudarev, S.L., G.A. Botton, S.Y. Savrasov, C.J. Humphreys, and A.P. Sutton, *Electron-energy-loss spectra and the structural stability of nickel oxide: An LSDA+U study.* Physical Review B, 1998. **57**(3): p. 1505-1509.

283. Mou, T., V.K. Gupta, M.F. Tabriz, T. Frauenheim, and P. Deák, *Size of electron polarons in anatase TiO_2 and their role in photocatalysis.* Physical Review B, 2023. **107**(15): p. 155127.

284. Sanderson, C. and R. Curtin, *Armadillo: a template-based C++ library for linear algebra.* Journal of Open Source Software, 2016. **1**: p. 26.

285. Powell, M.J.D., *Direct search algorithms for optimization calculations.* Acta Numerica, 2008. **7**: p. 287-336.

286. Powell, M.J.D., *The BOBYQA algorithm for bound constrained optimization without derivatives*, in *Department of Applied Mathematics and Theoretical Physics, Cambridge England, technical report*. 2009.

287. Rowan, T.H., *Functional Stability Analysis of Numerical Algorithms*, in *Department of Computer Sciences*. 1990, University of Texas at Austin.

288. Johnson, S.G., *The NLopt nonlinear-optimization package*. 2008-2014.

289. Nelder, J.A. and R. Mead, *A Simplex Method for Function Minimization.* The Computer Journal, 1965. **7**(4): p. 308-313.

290. Powell, M.J.D., *On trust region methods for unconstrained minimization without derivatives.* Mathematical Programming, 2003. **97**(3): p. 605-623.

291. Powell, M.J.D., *A Direct Search Optimization Method That Models the Objective and Constraint Functions by Linear Interpolation*, in *Advances in Optimization and Numerical*

Analysis, S. Gomez and J.-P. Hennart, Editors. 1994, Springer Netherlands: Dordrecht. p. 51-67.

292. Sanderson, C. and R. Curtin. *An Adaptive Solver for Systems of Linear Equations*. in *2020 14th International Conference on Signal Processing and Communication Systems (ICSPCS)*. 2020.

List of publications

1. Farzalipour Tabriz, M., K. Rehling, B. Aradi, T. Frauenheim, and P. Deák, *Photocatalytic oxidation of carbon monoxide on the anatase $TiO_2(101)$ surface*, In progress.
2. Mou, T., V. Kristin Gupta, M. Farzalipour Tabriz, T. Frauenheim, and P. Deák, *Size of electron polarons in anatase TiO_2 and their role in photocatalysis*, Physical Review B, 2023. **107**(15): p. 155127.
3. Chagas da Silva, M., M. Lorke, B. Aradi, M. Farzalipour Tabriz, T. Frauenheim, A. Rubio, D. Rocca, and P. Deák, *Self-Consistent Potential Correction for Charged Periodic Systems.* Physical Review Letters, 2021. **126**(7): p. 076401.
4. Deák, P., M. Han, M. Lorke, M. Farzalipour Tabriz, and T. Frauenheim, *Intrinsic defects of GaSe.* Journal of Physics: Condensed Matter, 2020. **32**(28): p. 285503.
5. Deák, P., E. Khorasani, M. Lorke, M. Farzalipour Tabriz, B. Aradi, and T. Frauenheim, *Defect calculations with hybrid functionals in layered compounds and in slab models.* Physical Review B, 2019. **100**(23): p. 235304.
6. Farzalipour Tabriz, M., B. Aradi, T. Frauenheim, and P. Deák, *SLABCC: Total energy correction code for charged periodic slab models.* Computer Physics Communications, 2019. **240**: p. 101-105.
7. Farzalipour Tabriz, M., B. Aradi, T. Frauenheim, and P. Deák, *Application of the Lany–Zunger polaron correction for calculating surface charge trapping.* Journal of Physics: Condensed Matter, 2017. **29**(39): p. 394001.