POLITECNICO DI MILANO
DEPARTMENT OF ENERGY
DOCTORAL PROGRAMME IN
ENERGY AND NUCLEAR SCIENCE AND TECHNOLOGY

Ab Initio Studies of Bulk and Defective Oxides Using Nonempirical Hybrid Density Functionals

s.s.d. FIS/03 – Physics of Matter

Doctoral Dissertation of:
Matteo Gerosa

Supervisors:
Prof. Carlo E. Bottani
Prof. Giovanni Onida
Prof. Gianfranco Pacchioni

Tutor:
Prof. Matteo Passoni

The Chair of the Doctoral Program:
Prof. Carlo E. Bottani

Year 2016 – XXVIII Cycle

To my parents

“Indeed, the only truly serious questions are ones that even a child can formulate. Only the most naive of questions are truly serious.”

– M. Kundera, *The Unbearable Lightness of Being*

Acknowledgements

First, I would like to express my sincere gratitude to my supervisors Prof. Carlo Bottani, Prof. Giovanni Onida, and Prof. Gianfranco Pacchioni for their enduring support and mentoring throughout this PhD, and for having introduced me to a new, exciting research field. I also would like to thank them for their time spent discussing with me career opportunities after graduation, which has helped me a lot sorting out future plans. I am also thankful to my tutor Prof. Matteo Passoni for his acute advice and guidance. I am greatly indebted to Prof. Cristiana Di Valentin for her always insightful supervision and advice on my work, and for accepting to serve as member of the committee for my final PhD defense. I am deeply grateful to Dr. Lucia Caramella for the time spent introducing me to practical aspects of first-principles calculations at the very beginning of this project, as well as for the continuous support especially during hard times in my PhD.

I would like to thank Prof. Giulia Galli for giving me the opportunity to spend the last three months of this PhD working in her research group at the University of Chicago. Thanks to all the group's people for what I could learn during this time, and for always being kind and welcoming to me.

Many thanks to all the Nanolab people for always contributing to make CeSNEF a friendly and enjoyable place to work in. In particular, I would like to thank my longtime officemates, Irene and Valeria, for the nice time spent together in the past three years, and for their friendship and support.

Ringrazio gli amici di Canzo e dintorni per i momenti condivisi in questi anni, per avermi aiutato a distrarmi dalle fatiche del dottorato, e per averne condiviso con me le soddisfazioni. Grazie in particolare a Angelo, Romina, Chiara e Mattia per l'amicizia che continua da più di dieci anni, nonostante le diverse strade intraprese e, a volte, la lontananza. Grazie a Ally per l'amicizia cresciuta nelle passeggiate in montagna.

Un ringraziamento particolare a chi, nella mia famiglia, mi ha sempre incoraggiato e sostenuto nel mio percorso nonostante la mia incorreggibile incapacità di spiegare in modo comprensibile cosa stessi facendo, grazie quindi ancor più per la fiducia e il supporto instancabili, che sono stati per me forte motivo per perseverare nei momenti difficili, e che mi hanno reso orgoglioso dei risultati raggiunti. Infine, grazie ai miei genitori per tanti e troppi motivi, ma in particolare per avermi insegnato tutto quanto c'è di veramente importante e per la fiducia che sempre mi danno. Grazie alle mie sorelle Francesca e Giulia per aver condiviso questo, e per farmi sentire, da "fratellone", qualche volta un po' più "saggio".

Canzo, 22 dicembre 2015
M. G.

Abstract

Metal oxides have been proposed as candidate materials for a number of renewable energy technologies, ranging from catalysis, to photocatalysis and solar cells. In applications involving the conversion of solar energy into chemical or electrical energy (e.g., in the case of photoelectrochemical devices or solar cells), a thorough understanding of the electronic structure and electronic excitation mechanisms is necessary. In catalytic applications instead, the chemical reactivity of the material, as well as the energetics relative to catalytic processes in which the material is actively involved, should be adequately characterized. Great challenges are posed to theory when these materials properties are to be computed using first-principles methods, with particular regard to density-functional theory (DFT) and many-body perturbation theory (MBPT).

In this thesis, DFT with hybrid exchange-correlation (xc) functionals are employed for calculation of the electronic structure and ground-state properties of wide-gap oxide materials. The employed hybrid functional can be considered as nonempirical, as the fraction of exact exchange to be admixed in the xc potential is evaluated *ab initio* for each material using a procedure which can be rigorously justified within MBPT. Results of various benchmarks of the method are reported for both bulk pristine and defective oxides. For the latter case, oxygen vacancies are investigated as a prototypical intrinsic point defect in transition metal oxides: the calculated properties are critically compared with various experiments characterizing charge localization and excitation mechanisms in the defective material. Furthermore, aluminum-doped

silicon dioxide is studied as a well-known system in which popular hybrid functional approximations fail in reproducing well-established experimental findings, due to inaccurate description of the ground state, which can be ultimately traced to the arbitrariness of the amount of exact exchange built in these functionals.

Sommario

I materiali ossidi sono promettenti candidati per un gran numero di applicazioni tecnologiche di rilevanza nel campo delle energie rinnovabili, dalla catalisi, alla fotocatalisi, alle celle solari. Per applicazioni in cui si vede necessario convertire l'energia solare in energia chimica o elettrica (ad esempio nei dispositivi fotoelettrochimici e nelle celle solari), è richiesta una conoscenza accurata della struttura elettronica del materiale, così come dei meccanismi di eccitazione elettronica. Per le applicazioni in catalisi, invece, è necessario ottenere una buona caratterizzazione della reattività chimica del materiale e dell'energetica associata a processi catalitici che coinvolgono direttamente il materiale stesso. La descrizione quantitativa da principi primi di queste proprietà, in particolare nell'ambito della teoria del funzionale di densità (*density-functional theory*, DFT) e della teoria delle perturbazioni a molti corpi (*many-body perturbation theory*, MBPT), è ancor oggi oggetto di intensi sforzi di ricerca.

In questa tesi, la teoria del funzionale di densità in combinazione con funzionali ibridi di scambio e correlazione sono utilizzati per il calcolo della struttura elettronica e delle proprietà di stato fondamentale di materiali ossidi ad ampio gap. Il funzionale ibrido utilizzato può essere considerato non empirico, essendo la frazione di scambio esatto valutata da principi primi sulla base di una procedura che può essere rigorosamente giustificata nel contesto della teoria delle perturbazioni a molti corpi. Viene riportata una validazione del metodo per il calcolo di diverse proprietà di ossidi *bulk* ideali e difettivi. Nel secondo caso, le vacanze di ossigeno sono studiate come difetto di

punto prototipo negli ossidi di metalli di transizione: le quantità calcolate vengono criticamente messe a confronto con i risultati degli esperimenti caratterizzanti la localizzazione di carica e i meccanismi di eccitazione nel materiale difettivo. Inoltre, il diossido di silicio drogato con alluminio viene analizzato come materiale modello in cui i più comuni funzionali ibridi non sono in grado di dare una descrizione corretta dello stato fondamentale del sistema, una caratteristica che deriva dall'arbitarietà della frazione di scambio esatto presente in tali funzionali.

Contents

Acknowledgements **V**

Abstract (English/Italian) **VII**

1 Introduction **1**
1.1 The Electronic Structure Problem 1
1.2 Electronic Structure Through DFT and MBPT: State of the Art 4
1.3 Oxide Materials: Theoretical Challenges and Applications 7
1.3.1 Electronic Structure and Photocatalysis 8
1.3.2 Thermochemical Properties and Catalysis 8
1.3.3 Defective Oxide Materials: Oxygen Vacancies and Al-Doped SiO_2 as Case Studies 10
1.4 Structure of the Thesis 12

I Theoretical Methods **14**

2 Density-Functional Theory Methods **15**
2.1 Fundamentals of Density-Functional Theory 15
2.1.1 The Hohenberg-Kohn Theorem 15
2.1.2 The Kohn-Sham Ansatz 19
2.2 Approximate Treatments of Exchange and Correlation . 23
2.2.1 Local Density Approximation 24
2.2.2 Generalized Gradient Approximations 25
2.2.3 Hybrid Functionals 26

2.3 Interpretation of the Kohn-Sham Eigenvalues: Excited States and Fundamental Gap in DFT 32
2.4 Practical DFT Calculations 35
2.4.1 Plane Waves and Pseudopotential Method 36
2.4.2 Linear Combination of Atomic Orbitals Method . 41

3 Many-Body Perturbation Theory (MBPT), and MBPT-Derived DFT Methods 45
3.1 Green's Function Theory 45
3.1.1 Single-Particle Green's Function and the Self-Energy 46
3.1.2 Linear Response and Hedin's Equations 49
3.1.3 The GW Approximation 51
3.1.4 The COHSEX Approximation 52
3.2 Practical GW Calculations 53
3.2.1 The G_0W_0 Approximation 53
3.2.2 Plane waves implementation of the G_0W_0 scheme 54
3.3 Combining MBPT and DFT: Self-Consistent Dielectric-Dependent Hybrid Functionals 57
3.3.1 Literature Review of Approaches to Determine the Exchange Fraction 59
3.3.2 Method Used for Calculation of the Dielectric Constant 59
3.3.3 Self-Consistent Hybrid Functional (sc-PBE0α_{ϵ_∞}) 61

4 Modeling of Point Defects in Insulating Materials 62
4.1 Introduction: The Supercell Approximation 62
4.2 Defect Formation Energy 63
4.2.1 Correction Term for Electrostatic Interactions: the Makov-Payne Scheme 65
4.3 Optical and Electrical Properties 67
4.3.1 Charge Transition Levels 69

II Investigation of Bulk and Defective Oxides 74

5 Bulk Oxides: Band Gaps and Phase Stabilities 75
5.1 Band Gap: Local/Semilocal, Hybrid Functionals and GW 76
5.1.1 Experimental Data 77
5.1.2 DFT LDA, GGA and Standard Hybrids: Comparison Between PW and LCAO 78
5.1.3 Many-Body Perturbation Theory: G_0W_0 79

5.1.4 DFT Dielectric-Dependent Hybrid Functionals . . 82
5.2 Oxide Polymorphs: Band Gaps, Crystal Structures and Phase Stability 83
5.2.1 Polymorphs Structures and Relative Stability . . . 85
5.2.2 Band Gap Dependence on Polymorphic Structure 88
5.2.3 Equilibrium Geometries 90
5.3 Stoichiometry Transformations: Chemical Reduction of Bulk TiO_2 91
5.4 Summary and Conclusions 94

6 Oxygen Vacancies in Wide-Gap Metal Oxides 99
6.1 Reduced Titanium Dioxide: TiO_{2-x} 102
6.1.1 Rutile 102
6.1.2 Anatase 105
6.2 Reduced Tungsten Trioxide: WO_{3-x} 108
6.3 Reduced Zirconium Dioxide: ZrO_{2-x} 111
6.4 Summary and Conclusions 114

7 Hole Localization in Al-Doped SiO_2 117
7.1 sc-PBE0α_{ϵ_∞} Results for Bulk SiO_2 119
7.2 Structural Deformation and Hole Localization 120
7.3 Analysis of EPR Parameters 120
7.4 Optical properties of the Al impurity 122
7.5 Summary and Conclusions 125

8 Conclusions 127

List of Figures 133

List of Tables 135

Bibliography 137

CHAPTER 1

Introduction

1.1 The Electronic Structure Problem

Electronic structure theory constitutes a fundamental tool for investigating electronic, optical and electrical properties of crystals, in which qualitatively new, complex physical phenomena emerge when a huge number of atoms are brought together to form the solid in the crystalline phase. The overlapping of atomic orbitals gives rise to extended, energy-dispersive electronic bands.

The theory of electronic bands was pioneered in 1928 by Felix Bloch [23], who first realized that the simplest and very fundamental description of electrons in crystals is that of a system of independent particles moving in a periodic potential. In a perfect crystal, the single-electron wavefunction obeys Bloch's theorem and is an eigenstate of the crystal momentum $\mathbf{k}$, the conserved quantity associated with the discrete translational symmetry of the underlying atomic lattice.

Early quantitative calculations of electronic bands for simple metals were performed in the 1930s and 1940s [166,284,324] using the cellular method, a precursor of the atomic sphere methods which constitute the basis of some of today's state-of-the-art approaches for *ab initio* band structure calculations. In the early 1950s, the first calculations of the

band structure of semiconductors also appeared [129].

Most of the current research efforts in the theory and computation of the electronic structure and optical properties of materials are pursued within density-functional theory [147] and many-body perturbation theory [226]. Through these approaches, first-principles calculation[1] of the structural, thermodynamic, and optoelectronic properties of complex, realistic condensed-matter systems can be pursued. The objective of the Part I of this thesis concerning "Theoretical Methods" is to overview the fundamental concepts underlying these two different theoretical schemes, as well as their application to investigation of both perfect and defective[2] insulating materials.

The aim of any theoretical approach to the study of condensed matter, and in particular of its electronic properties, is ultimately that of solving the many-body Schrödinger equation for a system of electrons (coordinates $\mathbf{r}_i$, momenta $\mathbf{p}_i$, charge $-e$, mass m) and nuclei (coordinates $\mathbf{R}_I$, momenta $\mathbf{P}_I$, charge $+z_I e$, mass M_I) interacting via the Coulomb force, whose Hamiltonian reads

$$\begin{aligned}\hat{H} = \sum_i \frac{|\hat{\mathbf{p}}_i|^2}{2m} + \sum_I \frac{|\hat{\mathbf{P}}_I|^2}{2M_I} + \sum_i V_{\text{nucl}}(\hat{\mathbf{r}}_i) \\ + \frac{1}{2}\sum_{i\neq j} \frac{e^2}{|\hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j|} + \frac{1}{2}\sum_{I\neq J} \frac{z_I z_J e^2}{|\hat{\mathbf{R}}_I - \hat{\mathbf{R}}_J|}\end{aligned} \tag{1.1}$$

with

$$V_{\text{nucl}}(\hat{\mathbf{r}}) = -\sum_I \frac{z_I e^2}{|\hat{\mathbf{r}} - \hat{\mathbf{R}}_I|}.$$

The above Hamiltonian is exact in the nonrelativistic limit, in the absence of external electromagnetic fields, and if energy terms related to the spin and the magnetic moment of electrons and nuclei are neglected. Thus, the terms appearing in its expression represent, respectively, the kinetic energy of the electrons, the kinetic energy of the nuclei, the electron-nucleus attractive potential energy, the electron-electron and nucleus-nucleus repulsive potential energy.

Within the adiabatic Born-Oppenheimer approximation [26] the nuclei are assumed frozen in a given configuration,[3] and the nuclear

[1] In the present context, a method is referred to as first-principles, or *ab initio*, if information from experiments is not used (i.e., the defining model Hamiltonian does not contain empirically fitted parameters); in principle, only the chemical nature of the atomic species constituting the system is to be known in advance.

[2] In particular, point defects such as atomic impurities or vacancies are addressed in this thesis.

[3] This configuration may correspond, for example, to the equilibrium one, in which all the atomic forces vanish.

coordinates thus play the role of parameters in the Hamiltonian (1.1). Correspondingly, the electron-nucleus potential energy becomes a constant external field acting on the electronic subsystem, and the nuclear repulsion term is a constant for any fixed configuration, which may be canceled by appropriate choice of the energy reference. The original problem is thus reduced to that of solving the many-body Schrödinger equation of N interacting electrons in the external Coulomb field created by nuclei,

$$\hat{H}_e \Psi(\mathbf{r}_1\,\sigma_1, \ldots, \mathbf{r}_N\,\sigma_N) = E\Psi(\mathbf{r}_1\,\sigma_1, \ldots, \mathbf{r}_N\,\sigma_N), \tag{1.2}$$

where σ_i is the spin coordinate of the i-th electron. The many-electron Hamiltonian reads

$$\hat{H}_e = \sum_i \frac{|\hat{\mathbf{p}}_i|^2}{2m} + \sum_i V_{\text{nucl}}(\hat{\mathbf{r}}_i) + \frac{1}{2}\sum_{i\neq j} \frac{e^2}{|\hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j|}, \tag{1.3}$$

where, in the coordinate representation, $\hat{\mathbf{r}} = \mathbf{r}$, and $\hat{\mathbf{p}} = -i\hbar\nabla$.

Even within the adiabatic approximation, the problem is not fundamentally reduced in its complexity, as the actual difficulties stem from the electron-electron interaction term, which makes solving the Schrödinger equation (1.2) an exceptionally hard task to tackle, even for the simplest molecules and solids. In fact, through this interaction term, the motion of each electron is intricately correlated to that of all the others, and the many-electron wavefunction cannot be expressed as a (antisymmetric) product of one-electron wavefunctions (i.e., a Slater determinant).[4] However, such correlation effects can be equally represented by recasting the many-electron problem into an equivalent one for independent particles moving in an effective generalized potential, which incorporates all the nontrivial many-body interactions. *Density-functional theory* (DFT) and *many-body perturbation theory* (MBPT) provide different, but equally theoretically justified, approaches to perform this mapping: the resulting single-particle Schrödinger-like equation to be solved is known as the *Kohn-Sham equation* and the *quasiparticle equation*, and the potential containing correlation effects are the *exchange-correlation potential* and the *electron self-energy*, respectively. Either within DFT or MBPT, an approximation for the relevant effective potential must be devised: some of the most popular approximations will be presented in Chapters 2 and 3. The

[4] The ansatz of a product wavefunction leads to the Hartree (simple product) [122], or the Hartree-Fock (antisymmetrized product) [91] approximations, in which dynamical electronic correlation is completely missing.

reformulated many-body problem can thus be solved by resorting to standard numerical algorithms, and the availability of high-performance computing resources today makes it possible to investigate the properties of realistic condensed-matter systems (containing hundreds of atoms).

However, the scope of the two methods is actually different: in its most common implementation, based on the ideas of Kohn and Sham [157], DFT is in principle only able to provide the exact ground state of the quantum system, thus allowing one to compute the corresponding total energy and other quantities directly related to it. Instead, the MBPT scheme is naturally suited to address excited states, as some quantities defined within it (most notably, quasiparticle energies) are apt to direct and rigorous comparison with spectroscopic observables. Using both approaches, or an appropriate combination of the two, first-principles calculation of the electronic structure of solid-state systems can be pursued, often resulting in satisfactory agreement with experiments.

1.2 Electronic Structure Through DFT and MBPT: State of the Art

The success of DFT [136] as the today's standard computational approach to investigation of a broad variety of solid-state and molecular systems is justified by its relatively low computational cost, together with the often good accuracy achieved in the calculation of ground-state properties [147]. In contrast, excited states are usually poorly described, being beyond the scope of DFT in the usual Kohn-Sham (KS) implementation [157]. The *KS equations* (a system of N coupled eigenvalue equations)[5]

$$\left(-\frac{\hbar^2\nabla^2}{2m} + v_{\text{ext}}(\mathbf{r}) + v_{\text{H}}[n](\mathbf{r}) + v_{\text{xc}}[n](\mathbf{r})\right)\phi_i(\mathbf{r}\sigma) = \varepsilon_i\phi_i(\mathbf{r}\sigma) \quad (1.4)$$

represent the (exact) Schrödinger problem for a system of effectively independent electrons (described by single-particle orbitals ϕ_i), into which the original many-body problem can be rigorously mapped. KS electrons are subject to an effective potential $v_{\text{s}}[n]$ (the *KS potential*) expressible as a functional of the electronic density.[6] $v_{\text{s}}[n]$ can be separated into different contributions: the external potential v_{ext}, typically describing the electrostatic field created by the nuclei,

[5] The index i is shorthand notation for both orbital and spin quantum numbers, and labels single-particle states in order of increasing energy.

[6] More precisely, the total energy is a density functional, and the potential is then obtained by differentiation.

the classical (Hartree) contribution coming from electron-electron interactions, $v_{\rm H}[n]$, and the unknown exchange-correlation (xc) potential $v_{\rm xc}[n]$, which accounts for all the remaining quantum many-body effects (electronic exchange and correlation).[7]

Plenty of approximations have been proposed for $v_{\rm xc}[n]$ in the literature; the simplest one relies on the assumption that at each point in the system, the xc energy is given by the xc energy of an homogeneous electron gas with density equal to the local density [157]. An extension of this *local density approximation* (LDA) to include gradient terms in the local expansion of the density leads to formulation of a new class of functionals, known as *generalized gradient approximations* (GGA) [177]. While working surprisingly well for many ground-state properties of a wide range of materials, both LDA and GGA fail in providing electronic band structures in quantitative agreement with experiments. Remarkably, the band gap, obtained from the eigenvalue difference of the first unoccupied and last occupied KS orbital, $[\varepsilon_{N+1}(\mathbf{k}) - \varepsilon_N(\mathbf{k})]$, turns out to be severely underestimated with respect to experimental values in most semiconductors and insulators. Such behavior is intimately connected with the fundamental properties of the standard KS scheme, as discussed Section 2.3).

The need to improve calculated thermochemical properties of molecular systems led to formulation of a new class of xc functionals, known as *hybrid functionals* [19, 20, 244, 292]. Hybrid xc energy functionals are typically constructed by admixing a GGA-type xc functional with a portion of the exact Hartree-Fock (HF) exchange energy. In fact, quantitative deviation of calculated molecular ground-state properties from experiments can be attributed to the *self-interaction error*, i.e. the incomplete cancellation of the Hartree self-interaction energy by the corresponding term in the exchange term, which affects both LDA and GGA [248], while being absent by construction in HF. Hybrid functionals thus constitute an attempt to partially compensate for such error. They are now routinely applied to extended, solid-state systems, for which they proved superior to LDA and GGA in the calculation of electronic band structures, often providing band gaps in quantitative agreement with experiments [54, 200, 216, 235]. A rigorous rationalization of this improvement can be given in the context of a generalized KS scheme [271], allowing for inclusion of nonlocal potentials in the class of admissible xc potentials; incorporating a portion of nonlocal HF exchange, hybrid functionals do

[7] For a more detailed definition of these quantities, see Section 2.1.2.

belong to this class (see Section 2.2.3).

Clearly, the performance of the hybrid functional is closely related to the amount of exact exchange (EXX) used for its construction. Popular hybrid functionals have this exchange fraction set to a constant value of 20% (B3LYP [292]) or 25% (PBE0 [244] and HSE06 [132, 133]), but at the time such choice was ultimately guided (for PBE0 and B3LYP) by comparison with experimental data (for atomization energies of standard set of molecules), so that the exchange fraction may be in fact considered as a semiempirical parameter in these functionals. On the other hand, application of the B3LYP or PBE0 to solid-state systems is not completely justified, since these functionals were not originally conceived for materials, and hence the value of the exchange fraction therein is not necessarily relevant to calculation of extended systems. In the end, one needs to seek for the actual physical significance of the exchange fraction in solids, and to devise an approach to evaluate it from first principles, i.e. not via direct or indirect relation with experiments. Such a program can be carried out by analyzing the formal similarities between the KS equation (1.4) and the quasiparticle equation derivable within MBPT [141],

$$\begin{aligned}\left(-\frac{\hbar^2\nabla^2}{2m}+v_{\text{ext}}(\mathbf{r})+v_{\text{H}}(\mathbf{r})\right)f_i(\mathbf{r}\sigma)\\ +\sum_{\sigma'}\int d^3r'\,\Sigma(\mathbf{r}\sigma,\mathbf{r}'\sigma',\varepsilon_i^{\text{QP}})f_i(\mathbf{r}'\sigma')=\varepsilon_i^{\text{QP}}f_i(\mathbf{r}\sigma),\end{aligned} \tag{1.5}$$

where the electron self-energy Σ is a generalized effective potential containing all the effects of exchange and correlation among electrons.[8] f_i and $\varepsilon_i^{\text{QP}}$ are the quasiparticle wavefunction and eigenvalues, which characterize effective single-particle states in the interacting many-body system (see Section 3.1). Notice that in the case of a crystal, the difference $[\varepsilon_{N+1}^{\text{QP}}(\mathbf{k})-\varepsilon_N^{\text{QP}}(\mathbf{k})]$ would exactly correspond to the measurable quasiparticle gap as obtained from photoemission experiments. Systematic approximations to the self-energy Σ can be derived within MBPT, and the well-known GW approximation [126] is the most commonly used one for state-of-the-art electronic structure calculations of materials [141] (see Section 3.1.3).

Given the manifest analogy between the KS equation (1.4) and the quasiparticle equation (1.5) (within the hybrid functional scheme, $v_{\text{xc}}[n](\mathbf{r},\mathbf{r}')$ can be viewed as an energy-independent, Hermitian, but

[8] The electron self-energy in general is a nonlocal, energy-dependent, and non-Hermitian operator.

possibly nonlocal approximation to Σ), one can seek for a formal relationship between the expression of the hybrid v_{xc}, and Σ as obtained in the static limit of GW, i.e. within the Coulomb-hole and screened-exchange (COHSEX) approximation proposed by Hedin [126]. Remarkably, the exchange fraction can thus be identified as the inverse macroscopic dielectric constant of the material [6, 198] (see Section 3.3). Hence, the exchange fraction is defined for each material, and can be evaluated within DFT starting from a chosen xc approximation. Furthermore, it has been demonstrated that the exchange fraction and the dielectric constant can be obtained self-consistently, leading to definition of a *self-consistent dielectric-dependent hybrid functional* [283]. This method proved superior to standard hybrid functional formulations in providing band gaps in quantitative agreement with experiments [198, 283], besides ensuring at least the same accuracy in the calculation of structural properties and total energies for a wide range of bulk semiconductors and insulators (see Chapter 5). Demonstrating the capabilities of this approach in investigating the properties of pristine and defective bulk oxide materials is the principal objective of this thesis.

1.3 Oxide Materials: Theoretical Challenges and Applications

Metal oxides, and in particular transition metal oxides (TMOs), are certainly one of today's most investigated chemical compounds in condensed matter physics. This interest is motivated both by the exceptionally rich physical phenomena encountered in these materials, going from magnetism to high-temperature superconductivity and strongly correlated phenomena [55, 152], and by the many potential technological applications (mainly utilizing semiconductors), ranging from microelectronics to the development of devices for solar energy harvesting and conversion.

In TMOs, the presence of localized, narrow d electronic bands plays a fundamental role in determining many interesting electronic and magnetic properties, making at the same time theoretical modeling of these materials a challenging task. In the following, a broad discussion on the present challenges concerning calculation of various properties of bulk and defective oxide materials is presented; selected examples of applications are also discussed in connection with some of these properties.

1.3.1 Electronic Structure and Photocatalysis

Characterizing the electronic structure of materials is certainly paramount for photocatalytic applications [253], in which certain semiconductors can catalyze conversion of sunlight into chemical energy (producing fuels) [114]. For example, production of hydrogen through photoelectrochemical splitting of water is a promising field of application for TMOs [308]. In a typical configuration for water splitting devices, the semiconductor is placed in contact with an aqueous electrolyte, and serves as a component for separation of the negative and positive charge carriers (electrons and holes) upon light absorption and application of a bias voltage. In order to optimize device performance, the material should efficiently absorb light in the visible range, and this limits its band gap to a maximum value of ~ 3.1 eV. Instead, the minimum band gap value is determined by the energy required to split water, corresponding to its redox potential of 1.23 eV, which in devices is increased up to ~ 1.9 eV in order to compensate for thermodynamic losses and the overpotential that needs to be applied in order to boost reaction kinetics.

Another aspect to be taken into account is the alignment of the semiconductor band edges with respect to water oxidation and reduction potentials; in fact, the valence band maximum should be below the oxidation potential, and the conduction band minimum above the reduction potential, in order to allow for oxidation and reduction of water, respectively.

Wide band gap oxide semiconductors have been identified as promising materials for photocatalytic applications, due to their suitable electronic structure features [168]. Theoretical investigations have demonstrated that state-of-the-art DFT (e.g., hybrid functionals) and many-body (GW) methods generally achieve a good description of bulk (see Chapter 5 and references therein) and surface and interface [209, 210, 252, 312, 317] band structures of oxide materials. Moreover, band engineering by incorporation of dopant species in the host material offers an attractive way to improve light absorption efficiency and band edge positioning, and constitutes a very active field of theoretical research [72].

1.3.2 Thermochemical Properties and Catalysis

In many other applications, such as in catalysis, one needs to characterize the chemical reactivity of the material. Oxide materials can

serve as active catalysts for promoting oxidation and dehydrogenation reactions. For example, oxidation of carbon monoxide (CO) can be efficiently achieved by exposure to a rutenium dioxide (RuO_2) surface [228]. Here the Mars-van Kravelen mechanism [199] takes place: the CO molecules get adsorbed at the $RuO_2(110)$ surface, which is highly reactive thanks to the presence of unsaturated dangling bonds belonging to the surface Ru atoms; when the CO-covered surface is heated, not only CO, but also CO_2 molecules detach from the surface, which are formed by surface O atoms reacting with the adsorbed CO. Thus, an oxygen vacancy (O vacancy) is created at the surface, which can be filled by exposure to gas phase oxygen, leading to regeneration of the catalyst. Clearly, the efficiency of the overall catalytic process is influenced by the energy cost of removing an O atom from the material (i.e. of creating an O vacancy). The presence of excess electrons consequent to removal of the O atom leads to reduction of the original oxide material,[9] eventually changing its stoichiometry. One then is interested in evaluating reaction energies associated to prototypical reduction pathways, such as, in the case of bulk titanium dioxide (TiO_2),

$$2\,TiO_2 + H_2 \rightarrow Ti_2O_3 + H_2O \tag{1.6}$$

$$2\,TiO_2 \rightarrow Ti_2O_3 + {}^1\!/_2\,O_2. \tag{1.7}$$

The oxidation state of the Ti cation passes from +4 in TiO_2 to +3 in the reduced oxide, Ti_2O_3, determining a transformation in the oxide stoichiometry.

A number of theoretical investigations have attempted to address the calculation of thermochemical quantities in oxides (for a review, see Ref. [232]), of which the energetics of model reduction processes like (5.1) and (5.2) is a prototypical example. However, it has been found that density-functional methods able to describe the thermochemistry are not necessarily equally capable of providing the correct description of the electronic band structure, and viceversa. This is dramatically evidenced when one tries to study the reduction of TiO_2 by means of DFT + U methodologies [189], where the Hubbard parameter U represents the effective Coulomb interaction for the d-electron subsystem. One then notices that U values that allow for a good description of the band structure of TiO_2 give reduction reaction energies in stark disagreement with experiment [139]. This questions on the actual predictivity of these methods, as long as they contain adjustable parameters whose value is

[9] This chemical reduction manifests itself in the reduction of the formal charge attributed to the metal ionic species in the oxide compound.

to be optimized according to the analyzed property.

In this thesis, the dielectric-dependent hybrid functional method briefly introduced in Section 1.2 is proposed as a fully *ab initio* approach to address the above issues. Benchmarks of its performance are reported for several polymorphs of five different oxide materials in their bulk structure.

1.3.3 Defective Oxide Materials: Oxygen Vacancies and Al-Doped SiO_2 as Case Studies

Clearly, the chemical reductions (5.1) and (5.2) constitute just representative mechanisms of modification of the original oxide stoichiometry. Actually, the formation of an O vacancy in oxides (occurring, e.g., through the Mars-van Kravelen mechanism described above) involves a much richer phenomenology, which cannot be captured by such a simple model. Thus, one has to directly tackle the modeling of a nonstoichiometric oxide, which calls for studying the effects of removal of one single O atom from an otherwise perfect crystal (see Section 4.1). In fact, this implies not only addressing the energetics (calculation of the vacancy formation energy), but also adequately describing the building up of defect levels in the band gap and the distribution of the related excess charge in the neighborhood of the defect. The presence of O vacancies substantially alters the optical, electrical, and magnetic properties of oxide semiconductors [98, 231], and a correct interpretation of experiments requires these modifications to be taken into account. A careful control of the defectivity plays a fundamental role in many applications other than catalysis [232], such as information technology [265], gas sensors [160], and smart windows [219]. In catalysis, oxide doping with heteroatoms is a promising way to reduce the energy cost for oxygen removal (i.e., O vacancy formation energy) [205], leading to possibly increased catalytic efficiencies.

Modeling of O vacancies in oxide materials is a challenging task, and the accuracy of the description achieved within DFT is strongly dependent on the level of approximation used for exchange and correlation [233]. Local or semilocal functionals such as LDA or GGA tend to produce delocalization of the excess charge in the host material (i.e., to poorly describe the defect wavefunction), as a consequence of the self-interaction error inherent in these approximations [73]. Furthermore, defect levels experimentally found in the bulk band gap are actually predicted by these methods to hybridize with the band edges, leading to an incorrect picture. Hybrid functionals generally

help correcting this problem [64], thanks to incorporation of a portion of EXX in the density functional. This leads to opening of the band gap, correcting the position of the band edges [5], and as a result localized defect levels in the gap show up. Actually, if one wants to characterize electronic transitions involving such defect levels, a quantitatively accurate description of the bulk electronic structure is a necessary prerequisite. Hence, the dielectric-dependent hybrid functional, being able to provide band gaps in quantitative agreement with experiments for wide-gap oxides [100, 198, 283], is a promising candidate for modeling defects; O vacancies represent a prototypical intrinsic point defect in these materials, where they are inevitably formed throughout the synthesis process, as a result of the exposition of the samples to high temperatures and strongly reducing atmospheres. Hence, they stand out as an ideal test for the performance of the method in describing defective oxides.

Finally, it is worthwhile to quantify to what extent the self-interaction error is avoided in the dielectric-dependent hybrid approach. In fact, this is an important source of inaccuracy in the description of defective semiconductors and insulators, as explained above. A paradigmatic system in which incomplete cancellation of self-interaction leads to a clearly incorrect ground state is aluminum-doped silicon dioxide (Al-doped SiO_2) [172,234]. In this system, the unpaired electron created by substitution of a tetravalent Si cation with a trivalent Al atom is predicted by most density-functional methods to be delocalized over two or more O atoms surrounding the Al impurity, at odds with the experimental evidence which suggests it to be trapped at a single O atom. Popular hybrid functionals such as B3LYP also fail in reproducing the correct experimental picture; it has thus been argued that the amount of EXX admixed is critical to obtain a truly self-interaction-free functional. The self-consistent dielectric-dependent functional is here proposed as a nonempirical approach to evaluate the optimal exchange fraction. Being this functional an approximate DFT implementation of the many-body COHSEX potential, in which by construction the exchange contribution is treated exactly, it is an almost obvious question to wonder if this property translates into the absence of self-interaction error in the directly derived DFT method.

1.4 Structure of the Thesis

This thesis is organized in two parts: the first one presents a general overview of *ab initio* theoretical methods which can be applied to investigation of bulk and defective materials, and in particular:

- In Chapter 2, a review of density-functional theory methods is reported, providing a rigorous introduction to the hybrid functional formalism; different numerical implementations of DFT employed in this work, corresponding to usage of different basis sets for the expansion of KS orbitals, are also discussed.
- Chapter 3 provides an overview of many-body perturbation theory methods based on Green's function theory, with particular reference to the GW approximation, which represents the state-of-the-art approach for electronic structure calculations. Connection of MBPT with DFT is also discussed, leading to the notion, central to this thesis, of (self-consistent) dielectric-dependent hybrid functional.
- Chapter 4 deals with application of DFT methods within the supercell approximation to characterization of point defects in insulating materials. The charge-transition level formalism is also discussed as a useful and rigorous approach to calculation of excitation and emission energies associated to electronic transitions involving defect levels and the bulk conduction and valence bands; connection to experimentally measurable spectroscopic quantities is discussed.

In the second part of the thesis results of calculations for several bulk and defective oxide semiconductors and insulators are collected.

- In Chapter 5, band gaps and phase stabilities of five different oxide materials (MgO, ZnO, TiO_2, ZrO_2, and WO_3) in several crystallographic structures are investigated by comparing the performance of different DFT functionals against computed GW band structures and experiments.
- In Chapter 6, the behavior of O vacancies in rutile and anatase TiO_2, room-temperature monoclinic WO_3, and tetragonal ZrO_2 is studied using the self-consistent dielectric-dependent hybrid functional. Calculated charge-transition levels relative to the O-vacancy defect are compared with the available experimental

results, with the goal of shedding light on the underlying excitation mechanism, and related materials properties.

- In Chapter 7, the long-standing problem of the hole localization at the Al impurity in quartz SiO_2 is addressed. Agreement of theoretical results with electron paramagnetic resonance and optical experiments is discussed.

Most results discussed in this thesis are adapted from the following papers by the author, which have been published or will be shortly submitted for publication in peer-reviewed journals:

[C1] M. Gerosa, C. E. Bottani, L. Caramella, G. Onida, C. Di Valentin, and G. Pacchioni. Electronic structure and phase stability of oxide semiconductors: Performance of dielectric-dependent hybrid functional DFT, benchmarked against GW bandstructure calculations and experiments. *Phys. Rev. B* 91:155201, 2015.

[C2] M. Gerosa, C. E. Bottani, L. Caramella, G. Onida, C. Di Valentin, and G. Pacchioni. Defect calculations in semiconductors through a dielectric-dependent hybrid DFT functional: the case of oxygen vacancies in metal oxides. *J. Chem. Phys.*, 143:134702, 2015.

[C3] M. Gerosa, C. Di Valentin, C. E. Bottani, G. Onida, and G. Pacchioni. Hole localization in Al-doped quartz SiO_2 within ab initio hybrid-functional DFT. *J. Chem. Phys.*, 143:111103, 2015.

[C4] M. Gerosa, C. E. Bottani, G. Onida, C. Di Valentin, and G. Pacchioni. Electrochromic properties and conductivity of reduced γ-monoclinic WO_3: The anisotropic behavior of the oxygen vacancy. *In preparation*.

Chapter 4 contains material from [C1], Chapters 3 and Chapter 5 from [C2], and Chapter 6 from [C3]. [C4] is partially included in Chapter 5. In the following, atomic units $\hbar = m = e = 1$ are understood when not stated otherwise.

Part I

Theoretical Methods

CHAPTER 2

Density-Functional Theory Methods

2.1 Fundamentals of Density-Functional Theory

Density-functional theory (DFT), in the approach developed by Hohenberg and Kohn [136] which addresses the static properties of an interacting quantum system,[1] allows one to look at the many-body problem from a new, quite revolutionary perspective. In fact, instead of directly tackling the solution of the many-electron problem (1.2) (consisting of computing the energy spectrum E_k and the corresponding wavefunctions $|\Psi_k\rangle$), it focuses on the general structure of the many-body Schrödinger equation and on the fundamental relationships among its ingredients.

2.1.1 The Hohenberg-Kohn Theorem

A rigorous definition for such relationships is provided by the Hohenberg-Kohn (HK) theorem [136], whose fundamental statements are reported in the following. Consider a system of N interacting spin-$\frac{1}{2}$ particles (typically electrons) with Hamiltonian $\hat{H} = \hat{T} + \hat{V}_{\text{int}} + \hat{V}_{\text{ext}}$, where $\hat{T} = \sum_i |\hat{\mathbf{p}}_i|^2/2$ is the kinetic energy operator,

[1] It thus tackles the solution of the time-independent many-body Schrödinger equation.

$\hat{V}_{\text{int}} = (1/2)\sum_{i\neq j} v_{\text{int}}(\hat{\mathbf{r}}_i, \hat{\mathbf{r}}_j)$ denotes the total particle-particle potential energy (it would correspond to the electron-electron repulsive potential of Eq. (1.3) for a system of electrons), and $\hat{V}_{\text{ext}} = \int d^3r\, v_{\text{ext}}(\mathbf{r})\hat{n}(\mathbf{r})$ is a (spin-independent) external potential acting on the system (identified with the nuclear electrostatic field of Eq. (1.3) for a system of electrons in a crystal); $\hat{n}(\mathbf{r})$ is the particle density operator. The following statements can be proven.

(a) There is a one-to-one correspondence between $v_{\text{ext}}(\mathbf{r})$, the (nondegenerate) ground state $|\Psi_0\rangle$ resulting from solution of the Schrödinger equation with Hamiltonian $\hat{H}$, and the ground-state density $n_0(\mathbf{r}) = \langle\Psi_0|\,\hat{n}(\mathbf{r})\,|\Psi_0\rangle$,

$$v_{\text{ext}}(\mathbf{r}) \iff |\Psi_0\rangle \iff n_0(\mathbf{r}), \tag{2.1}$$

up to a constant shift in $v_{\text{ext}}(\mathbf{r})$. Mathematically, one can state that the ground state is a unique functional of the ground-state density,

$$|\Psi\rangle = |\Psi[n]\rangle \quad \Longrightarrow \quad |\Psi_0\rangle = |\Psi[n_0]\rangle\,. \tag{2.2}$$

(b) Any ground state observable is a density functional,

$$O[n] = \langle\Psi[n]|\,\hat{O}\,|\Psi[n]\rangle \tag{2.3}$$

and, in particular, the ground state energy reads

$$E[n] = \langle\Psi[n]|\,\hat{H}\,|\Psi[n]\rangle = F[n] + \int d^3r\, v_{\text{ext}}(\mathbf{r})n(\mathbf{r}) \tag{2.4}$$

where $F[n] = \langle\Psi[n]|\,\hat{T} + \hat{V}_{\text{int}}\,|\Psi[n]\rangle$ is a universal functional, in the sense that it only depends on the nature of the particle-particle interaction potential.

(c) The energy functional $E[n]$ satisfies a variational principle: if n_0 is the ground-state density corresponding to v_{ext}, then

$$E[n_0] < E[n] \quad \text{for any} \quad n(\mathbf{r}) \neq n_0(\mathbf{r}). \tag{2.5}$$

The corresponding variational equation,

$$\frac{\delta}{\delta n(\mathbf{r})}\left\{E[n] - \mu\left(\int d^3r\, n(\mathbf{r}) - N\right)\right\}\Bigg|_{n(\mathbf{r})=n_0(\mathbf{r})} = 0 \tag{2.6}$$

in principle allows one to determine the ground-state density of the many-body system constituted by a fixed number N of particles (μ is the Langrange multiplier enforcing this condition, and can thus be identified with the chemical potential of the system).

The above statements are easily proved via *reductio ad absurdum* arguments [79]. Statement (a) can be extended to also encompass degenerate ground states. Notice that $E[n]$ is defined only for ground state densities resulting from solution of the Schrödinger equation $\hat{H}\,|\Psi_k\rangle = E_k\,|\Psi_k\rangle$, once v_{ext} has been specified. Such densities are termed interacting v-representable; the conditions that they should obey are not known in general, and this makes the problem of the v-representability a nontrivial one: one can conceive situations in which, although the density is perfectly admissible, the generating external potential does not exist or is nonanalytic [79, 80]. On the contrary, it can be proven that any arbitrary nonnegative density of a N-particle system can be derived from a suitable antisymmetric, normalized N-particle wavefunction $\left|\Psi^N\right\rangle$ [104]. Accordingly, it is possible to define a generalized energy functional (the Levy-Lieb functional), whose domain is constituted by such interacting N-representable densities [183, 184, 188]; finally, a variational principle analogous to (2.5) can be formulated for the generalized energy functional. This alternative formulation of DFT does not suffer from the v-representability issue, as the Levy-Lieb functional is well-defined for any density n in the vicinity of some ground state density n_0; moreover, it reduces to the usual HK functional for v-representable densities (for more details about this generalization of HK-DFT formulation, see Ref. [79]).

Spin-Polarized Systems

The HK theorem can also be extended to describe spin-polarized systems, e.g. electronic systems subject to an external magnetic field $\mathbf{B}_{\text{ext}}(\mathbf{r})$. The magnetic field couples to the magnetization density $\mathbf{m}(\mathbf{r})$, resulting in an additional energy term in the Hamiltonian, $\int d^3r\,\mathbf{B}_{\text{ext}}(\mathbf{r})\,\cdot\,\hat{\mathbf{m}}(\mathbf{r})$. Consequently, the ground state $|\Psi_0\rangle$ is in one-to-one correspondence with the ground-state particle and magnetization densities, n_0 and $\mathbf{m}_0$,

$$|\Psi\rangle = |\Psi[n,\mathbf{m}]\rangle \quad \Longrightarrow \quad |\Psi_0\rangle = |\Psi[n_0,\mathbf{m}_0]\rangle\,. \tag{2.7}$$

The corresponding energy functional $E[n,\mathbf{m}]$ satisfies the variational principle $E[n_0,\mathbf{m}_0] < E[n,\mathbf{m}]$ for any $(n,\mathbf{m}) \neq (n_0,\mathbf{m}_0)$.

If one considers magnetic fields having a single nonvanishing Cartesian component, $\mathbf{B}_{\text{ext}}(\mathbf{r}) = (0,0,B_{\text{ext}}(\mathbf{r}))$, only the z-component of the magnetization density, m_z, enters the above relations (spin-collinear case). Since $m_z(\mathbf{r}) = \mu_{\text{B}}\,(n_\uparrow(\mathbf{r}) - n_\downarrow(\mathbf{r}))$,[2] with n_σ ($\sigma = \uparrow,\downarrow$) being the

[2] $\mu_{\text{B}} = 1/2$ is the Bohr magneton.

spin density of the electron subsystem with spin parallel (antiparallel) to $\mathbf{B}_{\text{ext}}$, the HK energy functional can be expressed in terms of the set of variables $(n_\uparrow, n_\downarrow)$, which is completely equivalent to the set (n, m_z),

$$E[n_\uparrow, n_\downarrow] = F[n_\uparrow, n_\downarrow] + \int d^3r \, \{v_{\text{ext}}(\mathbf{r}) \, (n_\uparrow(\mathbf{r}) + n_\downarrow(\mathbf{r})) + \mu_{\text{B}} B_{\text{ext}}(\mathbf{r}) \, (n_\uparrow(\mathbf{r}) - n_\downarrow(\mathbf{r}))\} \tag{2.8}$$

where $F[n_\uparrow, n_\downarrow] = \langle \Psi[n_\uparrow, n_\downarrow] | \, \hat{T} + \hat{V}_{\text{int}} \, | \Psi[n_\uparrow, n_\downarrow] \rangle$. Since in the spin-collinear case the Hamiltonian commutes with the total particle number for a given spin, $\hat{N}_\sigma = \int d^3r \, \hat{n}_\sigma(\mathbf{r})$, the number of particles N_σ with spin σ can be fixed, as long as the total number of particles is constrained to satisfy the condition $N = N_\uparrow + N_\downarrow$. Hence, the DFT minimum principle can be applied separately to each of the σ sector of Fock space; the ground state is obtained for the pair $(N_\uparrow, N_\downarrow)$ which yields the minimum energy.

Even in the absence of any magnetic field, $B_{\text{ext}} = 0$, the discussed formalism can be usefully applied to study of spin-polarized systems, i.e. systems with nonvanishing magnetic moment. Such situation occurs, for example, in open-shell atomic and molecular systems, as well as in magnetic solids and defective materials. In these cases, the lowest energy solution is such that $n_\uparrow(\mathbf{r}) \neq n_\downarrow(\mathbf{r})$, and the DFT minimization is conveniently performed using spin densities as fundamental variables.

Either in the standard or the spin-polarized case, the ground state $|\Psi[\cdot]\rangle$ has the same functional form for all the many-body systems with the same particle-particle interaction: this translates into the already mentioned universality of $F[\cdot]$. Applied to typical condensed-matter systems, this means that electrons in atoms, molecules and solids (under the Born-Oppenheimer approximation) will be described by the same $F[\cdot]$. Any information on the specific system under consideration is contained in the form of the external potential v_{ext}. As a consequence, the properties of any atomic, molecular and solid-state system are completely determined by the number of electrons (as obtained by integration of the density over the volume of the system) and the types and positions of atoms within them. The latter fact is of fundamental importance for practical implications of the theory, namely the possibility of a first-principle description of condensed-matter systems.

In spite of its great conceptual power, DFT does not provide a direct way to retrieve the properties of a many-particle system from

knowledge of its density profile alone. For example, in the case of N noninteracting electrons in an external potential, the $F[n]$ term in (2.4) is simply given by the kinetic energy. However, there is no known way to exactly express the kinetic energy as an explicit functional of the density. An early attempt to construction of an approximate kinetic energy functional was made by Thomas [299] and Fermi [84]; in the Thomas-Fermi theory, the kinetic energy is locally approximated to that of a noninteracting homogeneous electron gas. However, within this approach, even basic properties of quantum systems, namely the shell structure of atoms, are not properly reproduced. For a noninteracting system, the kinetic energy may be written in terms of single-particle wavefunctions; yet, the same does not hold true for interacting systems. The crucial step forward, made by Kohn and Sham in 1965 [157], is to formulate the kinetic energy in terms of suitable single-particle orbitals (the Kohn-Sham orbitals), which provide a basis for a convenient (and nontrivial) representation of the interacting density.

2.1.2 The Kohn-Sham Ansatz

The Kohn-Sham (KS) approach [157] to DFT consists in establishing a connection between the properties of the original many-body system and those of an auxiliary noninteracting system, by means of which the former can be conveniently expressed and computed. The KS ansatz assumes that such noninteracting system can be chosen so that its ground-state density equals that of the original interacting system. Hence, by virtue of the HK theorem,[3] it suffices to solve the resulting auxiliary single-particle equations to obtain an in principle equivalent description of the interacting system. In practice, all the nontrivial part of the many-body problem is mapped into an exchange-correlation (xc) functional of the density, which eventually needs to be approximated.

The auxiliary system is chosen to be realized by N noninteracting electrons[4] subject to a multiplicative external potential v_{s},

$$\hat{H}_{\mathrm{s}} = \hat{T} + \hat{V}_{\mathrm{s}}, \qquad \hat{V}_{\mathrm{s}} = \int d^3r\, \hat{n}(\mathbf{r}) v_{\mathrm{s}}(\mathbf{r}). \tag{2.9}$$

The corresponding N-particle (nondegenerate) ground state is a Slater

[3] "Any ground state observable is a density functional, which takes its true value when evaluated for the true ground state density".

[4] From now on, the original many-body system is assumed to be constituted by electrons purely interacting via the Coulomb force, $v_{\mathrm{int}}(\mathbf{r}, \mathbf{r}') \equiv v(\mathbf{r}, \mathbf{r}') = 1/|\mathbf{r} - \mathbf{r}'|$. The nature of the chosen KS system follows from this assumption.

determinant,

$$\Phi_0(\mathbf{r}_1\sigma_1,\ldots,\mathbf{r}_N\sigma_N) = \frac{1}{\sqrt{N!}}\begin{pmatrix} \phi_1(\mathbf{r}_1\sigma_1) & \ldots & \phi_N(\mathbf{r}_1\sigma_1) \\ \vdots & & \vdots \\ \phi_1(\mathbf{r}_N\sigma_N) & \ldots & \phi_N(\mathbf{r}_N\sigma_N) \end{pmatrix} \quad (2.10)$$

constructed from the energetically lowest energy eigenstates ϕ_i of the single-particle Schrödinger equation,

$$\left(-\frac{\nabla^2}{2} + v_{\mathrm{s}}(\mathbf{r})\right)\phi_i(\mathbf{r}) = \varepsilon_i\phi_i(\mathbf{r}), \quad (2.11)$$

where the eigenvalues ε_i are ordered according to $\varepsilon_1 \leq \varepsilon_2 \leq \ldots \leq \varepsilon_N = \varepsilon_F < \varepsilon_{N+1} \leq \varepsilon_{N+1} \leq \ldots$, and ε_F is the Fermi energy, which is taken equal to the eigenvalue ε_N of the highest-occupied single-particle state.[5] The quantum number $i = (\alpha, s)$ labels both spatial and spin quantum numbers; since $\hat{H}_{\mathrm{s}}$ is spin-independent, the total spin operator commutes with it and ϕ_i can be factorized into orbital and spin parts, $\phi_i(\mathbf{r}\sigma) = \phi_\alpha(\mathbf{r})\chi_s(\sigma)$, where $\chi_{\mathrm{s}}(\sigma) = \delta_{s\sigma}$ is the Pauli spinor. The eigenvalues ε_i are at least twofold degenerate due to spin; the ground-density can thus be written

$$n_{0s}(\mathbf{r}) = \sum_{\sigma=\uparrow,\downarrow}\sum_i \Theta_i \left|\phi_i(\mathbf{r}\sigma)\right|^2 = 2\sum_{\alpha=1}^{N/2}\left|\phi_\alpha(\mathbf{r})\right|^2 \quad (2.12)$$

where the occupation function at zero temperature is $\Theta_i = \theta(\varepsilon_F - \varepsilon_i)$, and $\theta(x)$ is the step function; at finite temperature it is identified with the Fermi function.

The crucial observation is that, by virtue of the HK theorem, n_{0s} determines the ground state uniquely: $|\Phi\rangle = |\Phi[n]\rangle \implies |\Phi_0\rangle = |\Phi[n_{0s}]\rangle$. The HK ground-state energy functional for the noninteracting system can be written as

$$E_s[n] = \langle\Phi[n]|\,\hat{T} + \hat{V}_{\mathrm{s}}\,|\Phi[n]\rangle = \langle\Phi[n]|\,\hat{T}\,|\Phi[n]\rangle + \int d^3r\, v_{\mathrm{s}}(\mathbf{r})n(\mathbf{r}), \quad (2.13)$$

and its minimum is obtained for at $n(\mathbf{r}) = n_{0s}(\mathbf{r})$: the true ground-state energy of the noninteracting system is thus $E_{\mathrm{s},0} = E_s[n_{0\mathrm{s}}]$.

For any density n resulting from diagonalization of the Hamiltonian (2.9), i.e. for any noninteracting v-representable density, the kinetic

[5] The restriction $\varepsilon_N < \varepsilon_{N+1}$ is necessary to ensure that Φ_0 is nondegenerate. The degenerate case is more subtle to be treated, see Ref. [79].

energy functional is defined as

$$T_{\mathrm{s}}[n] = \langle \Phi[n] | \, \hat{T} \, | \Phi[n] \rangle \,, \tag{2.14}$$

which is universal, since so $|\Phi[n]\rangle$ is. For the true ground state, the noninteracting kinetic energy can be explicitly expressed in terms of the single-particle orbitals ϕ_i,

$$T_{\mathrm{s},0} = T_{\mathrm{s}}[n_{0\mathrm{s}}] = \sum_{\sigma=\uparrow,\downarrow} \sum_i \Theta_i \int d^3r \, \phi_i^*(\mathbf{r}\sigma) \left(-\frac{\nabla^2}{2} \right) \phi_i(\mathbf{r}\sigma). \tag{2.15}$$

The KS ansatz can be stated as follows: for any admissible potential v_{ext}, the interacting ground-state density $n_0(\mathbf{r})$ is simultaneously the ground-density of the auxiliary noninteracting system subject to an appropriate external potential v_{s}. The existence of such KS system is related to the issue of the noninteracting v-representability of n_0: given an arbitrary n_0, is it always possible to construct v_{s} such that n_{0s} equals n_0? The answer to this question involves a rather advanced mathematical treatment [79]. In the following, it is assumed that this ansatz holds true for all the physically relevant interacting densities; the KS system thus yields a highly nontrivial representation of the interacting density in terms of single-particle orbitals,

$$n_0(\mathbf{r}) \equiv n_{0\mathrm{s}}(\mathbf{r}) = \sum_{\sigma=\uparrow,\downarrow} \sum_i \Theta_i \left| \phi_i(\mathbf{r},\sigma) \right|^2 . \tag{2.16}$$

The ground-state energy functional of the interacting system can be partitioned as follows

$$E[n] = T_{\mathrm{s}}[n] + E_{\mathrm{H}}[n] + E_{\mathrm{ext}}[n] + E_{\mathrm{xc}}[n], \tag{2.17}$$

which isolates contributions readily expressible within the KS system from all the complicated many-body effects, which are incorporated into the xc energy functional $E_{\mathrm{xc}}[n]$ defined through Eq. (2.17). The noninteracting kinetic energy is given by Eq. (2.15). The classical Hartree interaction energy between N electrons with density n is

$$E_{\mathrm{H}}[n] = \frac{1}{2} \int d^3r \int d^3r' n(\mathbf{r}) v(\mathbf{r},\mathbf{r}') n(\mathbf{r}') \tag{2.18}$$

and includes the self-interaction of the charge density with itself. The energy related to coupling of the external potential to particle density is

$$E_{\mathrm{ext}}[n] = \int d^3r \, v_{\mathrm{ext}}(\mathbf{r}) n(\mathbf{r}). \tag{2.19}$$

$E_{\rm xc}[n]$ is a universal functional of the density, since the same is true for $F[n]$, $T_{\rm s}[n]$ (by virtue of the HK theorem applied to the interacting and the auxiliary noninteracting systems, respectively), $E_{\rm H}[n]$ and $E_{\rm ext}[n]$; hence, the same xc functional applies to all systems in which particle-particle interaction is mediated by Coulomb force. Notice that $E_{\rm xc}[n]$ does not simply incorporate the quantum many-body effects beyond the classical Hartree contribution, but also the difference between the interacting and the noninteracting kinetic energy. This contribution is generally of the same order of magnitude as the whole correlation energy.

In order to specify the form of $v_{\rm s}$, one finally applies the DFT variational principle: $E[n]$ has a minimum at the true ground-state density $n_0(\mathbf{r}) \equiv n_{0s}(\mathbf{r})$ reproduced by the ϕ_i orbitals. If the density is infinitesimally perturbed from its ground-state value,

$$n(\mathbf{r}) = n_0(\mathbf{r}) + \delta n(\mathbf{r}) \tag{2.20}$$

then $E[n]$ departs from the minimum. By performing a Taylor expansion of the energy functional $E[n_0 + \delta n]$ around $n_0(\mathbf{r})$, and imposing the difference $E[n_0 + \delta n] - E[n_0]$ to be zero to second order in δn, one finally arrives at an explicit form for the noninteracting potential,

$$v_{\rm s}[n](\mathbf{r}) = v_{\rm ext}(\mathbf{r}) + v_{\rm H}[n](\mathbf{r}) + v_{\rm xc}[n](\mathbf{r}) \tag{2.21}$$

which must be determined self-consistently with the orbitals ϕ_i by solving the KS equations

$$\left(-\frac{\nabla^2}{2} + v_{\rm ext}(\mathbf{r}) + v_{\rm H}[n](\mathbf{r}) + v_{\rm xc}[n](\mathbf{r})\right)\phi_i(\mathbf{r},\sigma) = \varepsilon_i\phi_i(\mathbf{r},\sigma). \tag{2.22}$$

$v_{\rm H}$ is the Hartree potential

$$v_{\rm H}[n](\mathbf{r}) = \int d^3r'\, v(\mathbf{r},\mathbf{r}')n(\mathbf{r}'), \tag{2.23}$$

and the xc potential $v_{\rm xc}$ is defined as

$$v_{\rm xc}[n](\mathbf{r}) = \frac{\delta E_{\rm xc}[n]}{\delta n(\mathbf{r})}. \tag{2.24}$$

Spin-Polarized Systems

The KS formalism can be readily extended to spin-polarized systems; in the simplest spin-collinear case, in which the external magnetic field is

restricted to a pure z-component, the noninteracting potential explicitly depends on the spin state σ,

$$\hat{V}_{\mathrm{s}} = \sum_{\sigma=\uparrow,\downarrow} \int d^3r \, v_{\mathrm{s}}^{\sigma}(\mathbf{r}) \hat{n}_{\sigma}(\mathbf{r}). \tag{2.25}$$

As a consequence, the spatial part of the single-particle orbitals ϕ_i also depends on the spin quantum number s, $\phi_i(\mathbf{r}\sigma) = \phi_{\alpha s}(\mathbf{r})\chi_s(\sigma)$. The resulting KS equations of spin-density-functional theory embody a spin-dependent potential,

$$\left(-\frac{\nabla^2}{2} + v_{\mathrm{s}}^{\sigma}[n_\uparrow, n_\downarrow](\mathbf{r}) \right) \phi_{\alpha\sigma}(\mathbf{r}) = \varepsilon_{\alpha\sigma}\phi_{\alpha\sigma}(\mathbf{r}), \tag{2.26}$$

where v_{s} generalizes to

$$v_{\mathrm{s}}^{\sigma}[n_\uparrow, n_\downarrow](\mathbf{r}) = v_{\mathrm{ext}}(\mathbf{r}) + \mathrm{sign}(\sigma)\mu_B B_{\mathrm{ext}}(\mathbf{r}) + v_{\mathrm{H}}[n](\mathbf{r}) + v_{\mathrm{xc}}^{\sigma}[n_\uparrow, n_\downarrow](\mathbf{r}), \tag{2.27}$$

$$v_{\mathrm{xc}}^{\sigma}[n_\uparrow, n_\downarrow](\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}[n_\uparrow, n_\downarrow]}{\delta n_{\sigma}(\mathbf{r})}. \tag{2.28}$$

Thus, apart from the term proportional to the magnetic field, the xc potential also contributes to the different potential experienced by each spin component. This formalism turns out to be particularly useful in the limiting case of $B_{\mathrm{ext}} = 0$, i.e. when dealing with systems with a ground state exhibiting a finite magnetic moment, e.g. paramagnetic point defects in solids.

2.2 Approximate Treatments of Exchange and Correlation

The fundamental advantage of re-expressing the ground-state energy as in Eq. (2.17) is that the noninteracting kinetic energy and the Hartree term (which can be evaluated exactly) are isolated from the remaining many-body contributions. The former two terms, together with the energy associated to the physical external potential, are dominating over the remaining xc energy. Thus, one can hope to obtain reasonably accurate results by devising appropriate approximations to the xc functional. In the following, the most popular xc approximations are reviewed, also in light of their relevance for the rest of the thesis. A more comprehensive discussion of other xc functionals commonly used in DFT applications can be found in Refs. [79, 201].

2.2.1 Local Density Approximation

The simplest approach is to assume that the xc energy density at point $\mathbf{r}$, where density is $n(\mathbf{r})$, is locally equal to the xc energy density of an homogeneous electron gas (HEG), $e_{\text{xc}}^{\text{HEG}}$, with density $n_0 \equiv n(\mathbf{r})$. The resulting local density approximation (LDA) [157] to the xc energy reads

$$E_{\text{xc}}^{\text{LDA}}[n] = \int d^3r\, e_{\text{xc}}^{\text{HEG}}(n(\mathbf{r})). \tag{2.29}$$

The xc energy density of the HEG can be partitioned into exchange and correlation contributions, $e_{\text{xc}}^{\text{HEG}}(n_0) = e_{\text{x}}^{\text{HEG}}(n_0) + e_{\text{c}}^{\text{HEG}}(n_0)$. The exchange energy density e_{x} is easily expressed in terms of the density [109],

$$e_{\text{x}}^{\text{HEG}}(n_0) = -\frac{3(3\pi^2)^{1/3}}{4\pi} n_0^{4/3}. \tag{2.30}$$

The correlation energy density e_{c} is not known analytically for arbitrary densities of the HEG. One can obtain explicit expressions in terms of n_0 in the limit of low and high densities [109]; however, intermediate density regimes are typically the relevant ones to, e.g., materials and molecules. e_{c} can be extracted from the HEG total energy computed by means of quantum Monte Carlo simulations [40], and interpolation formulas have been proposed to parametrize the correlation energy density at arbitrary densities, the most popular of which are due to the work of Vosko, Wilk and Nusair [313], and of Perdew and Zunger [248]. The LDA xc correlation energy can then be explicitly evaluated, and the corresponding xc potential is a simple function of the local density,

$$v_{\text{xc}}^{\text{LDA}}(\mathbf{r}) = \frac{\delta E_{\text{xc}}^{\text{LDA}}[n]}{\delta n(\mathbf{r})} = \left.\frac{de_{\text{xc}}^{\text{HEG}}(n_0)}{dn_0}\right|_{n_0=n(\mathbf{r})}. \tag{2.31}$$

One issue which is readily realized with this approximation is that the LDA potential does not reproduce the correct asymptotic behavior for $|\mathbf{r}| \to \infty$ in finite systems. In atoms, for example, at large distances from the nucleus an electron should see the Coulomb potential of the positively charged ion, $-1/|\mathbf{r}|$; in contrast, the LDA predicts an exponential decay of the exchange potential, as a consequence of the same behavior followed by the density in finite systems. An analogous exponential behavior is obtained for $v_{\text{c}}^{\text{LDA}}(\mathbf{r})$ at large $\mathbf{r}$, in contrast with the correct power-law behavior. One important consequence of this failure is the incorrect description of Rydberg states in atoms, and the consequent prediction for negative ions to be unbound [282].

On the other hand, the LDA has proved able to yield exceptionally accurate results for a wide range of molecular and condensed-matter systems, regardless the degree of inhomogeneity. This success can be rationalized on the basis of the generally short-range behavior of the xc contribution to total energy, the long-range Hartree term being treated separately and exactly within the KS scheme. Furthermore, one can show that the LDA satisfies an important sum rule on the global strength of the exact xc hole [109].

One can similarly construct a local approximation to the (collinear) spin-density xc functional $E_{\rm xc}[n_\uparrow, n_\downarrow]$, by generalizing the LDA to the local spin-density approximation (LSDA). For the exchange energy functional written in terms of the KS orbitals,[6]

$$E_{\rm x}[n_\uparrow, n_\downarrow] = -\frac{1}{2} \sum_{\sigma=\uparrow,\downarrow} \sum_{\alpha\beta} \Theta_{\alpha\sigma}\Theta_{\beta\sigma} \times \int d^3r \int d^3r' \frac{\phi^*_{\alpha\sigma}(\mathbf{r})\phi_{\beta\sigma}(\mathbf{r})\phi^*_{\beta\sigma}(\mathbf{r}')\phi_{\alpha\sigma}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{2.32}$$

the following factorization holds

$$E_{\rm x}[n_\uparrow, n_\downarrow] = \frac{1}{2}\left(E_{\rm x}[2n_\uparrow] + E_{\rm x}[2n_\downarrow]\right), \tag{2.33}$$

and thus the the LSDA exchange energy can be evaluated in an analogous way, by using the LDA exchange (2.29),

$$E_{\rm x}^{\rm LSDA}[n_\uparrow, n_\downarrow] = \frac{1}{2}\left(E_{\rm x}^{\rm LDA}[2n_\uparrow] + E_{\rm x}^{\rm LDA}[2n_\downarrow]\right). \tag{2.34}$$

On the contrary, the spin-dependent correlation energy is more involved to be evaluated. Starting from the available quantum Monte Carlo data for the spin-unpolarized or the fully spin-polarized HEG, various parametrizations of $e_{\rm c}^{\rm HEG}(n_\uparrow, n_\downarrow)$ have been suggested for intermediate spin-polarizations, and the corresponding $E_{\rm c}^{\rm LSDA}$ could then be evaluated along the same lines as in LDA [247, 248].

2.2.2 Generalized Gradient Approximations

The most obvious extension to L(S)DA would be apparently obtained by constructing an energy functional explicitly dependent on the gradient of

[6] The spin-resolved occupation function $\Theta_{\alpha\sigma}$ is defined by generalizing the occupation function Θ_i of Eq. (2.12).

the density. However, it has been realized that a naive gradient expansion actually turns out to yield less accurate results than the LDA. The crucial point is that the xc hole sum rule obeyed by LDA becomes violated if the xc energy is simply expanded in the density fluctuation around $n(\mathbf{r})$. Generalized gradient approximations (GGAs) have been proposed to overcome such limitation [177]. The GGA xc energy functional is expressible in the form [241]

$$E_{\rm x}^{\rm GGA}[n_\uparrow, n_\downarrow] = \int d^3r\, n(\mathbf{r}) e_{\rm x}^{\rm HEG}(n(\mathbf{r})) F_{\rm xc}\left(n_\uparrow, n_\downarrow, \nabla n_\uparrow, \nabla n_\downarrow\right), \quad (2.35)$$

where $F_{\rm xc}$ depends, to lowest order, on the dimensionless variable $s \sim |\nabla n|/n$. Various forms have been proposed for $F_{\rm xc}$, the most popular of which are contained in the works of Perdew and Wang [247], and Perdew, Burke and Ernzerhof [242]: the corresponding functionals are known as PW91 and PBE, respectively. Different forms of $F_{\rm x}$ (which is dominating over the correlation part) give different asymptotic behaviors for large s; however, in the region of $0 < s \lesssim 3$, which is the most relevant for most systems of interest, different approximations give practically the same values [241]. Remarkably, in this region, $F_{\rm x} \geq 1$, so the GGAs tend to yield an exchange energy lower than the LDA. Since this reduction is more substantial in atoms (larger inhomogeneity) than in molecules and solids, this eventually results in a reduction of the binding energy evaluated within GGAs, which corrects the LDA overbinding and improves agreement with experiment [79].

2.2.3 Hybrid Functionals

A completely different approach to treatment of the xc energy makes use of the concept of orbital-dependent functionals [169]: $E_{\rm xc}$ is represented in terms of the KS orbitals and eigenvalues, and is thus only an implicit functional of the total density. Such a development is motivated by the intrinsic limitations of the standard DFT-KS scheme when used in conjunction with the LDA or the GGA.

In particular, one realizes that the wrong asymptotic behavior of the LDA and GGA exchange potential is related to incomplete cancellation of the density self-interaction included in the Hartree potential (2.18) by the other terms in the approximated KS potential. Such cancellation occurs exactly in the HF theory, in which the exchange potential, expressed on a per-orbital basis, is nonlocal in space. This indicates that some nonlocality should be introduced in the exchange functional, in order to correctly retrieve the Coulomb $1/r$-behavior (the Coulomb

matrix elements leading to self-interaction in the nonlocal Hartree term should thus be present, with opposite sign, in the exchange term). The prototypical functional satisfying this condition is the exact-exchange (EXX) energy of KS electrons, Eq. (2.32), which can be evaluated explicitly once a set of KS orbitals is specified.

Another deficiency of the LDA and GGA approximations is the lack of derivative discontinuity in the xc functional. On the contrary, the exact functional (suitably generalized to systems with a noninteger number of particles) exhibits a change in its slope when varying the particle number across the integer number N [246]. The absence of the derivative discontinuity results, for example, in a qualitatively wrong description of dissociation mechanisms in molecules [41]. By construction, orbital-dependent functionals allow for infinitesimal occupation of a previously unoccupied orbital; they are thus naturally suitable to best reproducing the derivative discontinuity of the exact xc functional.

A rigorous implementation of the concept of orbital-dependent xc functionals is beyond the scope of the standard KS approach described in Section 2.1.2. To date, the most popular and computationally efficient way to achieve such description is provided by the generalized Kohn-Sham (GKS) scheme developed by Seidl *et al.* [271].[7] In this framework, the physical system is mapped onto an auxiliary interacting system, which incorporates part of the xc contribution to total energy, while being still representable through a Slater determinant. The total energy functional is partitioned so that this (in general, nonlocal and orbital-dependent) part can be exactly evaluated. For example, such representation can be chosen so that both the kinetic and the HF contributions are represented exactly through the auxiliary system, leaving only correlation to be approximated (Hartree-Fock-Kohn-Sham equations are thus obtained). Hence, the GKS approach constitutes a rigorous basis for employing nonlocal, orbital-dependent, xc potentials, whose exchange component is inspired to the exact HF-like exchange (2.32).

One of the most popular realizations of orbital-dependent functionals is represented by hybrid functionals, which were originally introduced as a method to improve prediction of thermochemical properties of molecules [19, 20]. The hybrid xc energy is written as a

[7] Another suitable, but computationally more demanding, approach is represented by the optimized potential method [275], in which the KS equations are solved simultaneously with an integral equation for v_{xc}.

one-parameter-dependent admixture of the KS EXX to GGA-type functionals,

$$E_{\mathrm{xc}} = \alpha E_{\mathrm{x}} + (1-\alpha) E_{\mathrm{x}}^{\mathrm{GGA}} + E_{\mathrm{c}}^{\mathrm{GGA}}, \tag{2.36}$$

where $0 < \alpha < 1$ is the exchange fraction, and E_{x} is the EXX energy of the KS system, Eq. (2.32). A formal justification of an expression of the form (2.36) for E_{xc} is provided by the adiabatic connection theorem [120]: the exact xc energy of the electronic system can be expressed in terms of the xc energy of an auxiliary system with variable interaction strength $\lambda\, v$,

$$E_{\mathrm{xc}}[n] = \int_0^1 d\lambda\, E_{\mathrm{xc},\lambda}[n], \tag{2.37}$$

provided that, for any $0 \leq \lambda \leq 1$, the true ground state density n_0 is reproduced in the auxiliary system by appropriate choice of the auxiliary external potential $v_{\mathrm{ext},\lambda}$, so that $n_{0,\lambda} = \left\langle \Psi^{\lambda} \middle| \Psi^{\lambda} \right\rangle \equiv n_0$. In the limit $\lambda = 0$ one obtains the KS system, while for $\lambda = 1$ the physical, fully-interacting system is approached. The xc energy of the auxiliary system is thus

$$E_{\mathrm{xc},\lambda}[n] = \left\langle \Psi^{\lambda} \middle| \hat{V}_{\mathrm{int}} \middle| \Psi^{\lambda} \right\rangle - E_{\mathrm{H}}[n], \tag{2.38}$$

which leads to a representation of the exact xc energy that allows to circumvent the evaluation of the kinetic energy term, at the cost of having to evaluate the potential energy contribution for all the adiabatic states connecting the KS system to the physical one.

"Half-and-Half" Functional

Different ansatzs have been suggested for approximating the integral (2.37). If one assumes that $E_{\mathrm{xc},\lambda}$ varies linearly with λ, the so-called half-and-half approximation [19] is obtained,

$$E_{\mathrm{xc}} = \frac{1}{2} E_{\mathrm{xc},\lambda=0} + \frac{1}{2} E_{\mathrm{xc},\lambda=1}. \tag{2.39}$$

In the noninteracting limit, $\lambda = 0$, the KS system is recovered and the xc energy, given by

$$E_{\mathrm{xc},\lambda=0}[n] = \left\langle \Psi^{0} \middle| \hat{V}_{\mathrm{int}} \middle| \Psi^{0} \right\rangle - E_{\mathrm{H}}[n] \tag{2.40}$$

where Ψ^0 is a Slater determinant, and coincides with the KS EXX energy (2.32), $E_{\mathrm{x}}[n]$, which is independent of the interaction strength, as it arises only from the antisymmetry of the wavefunction. In the fully

interacting limit, $\lambda = 1$, the physical system is obtained, for which the GGA provides a reasonably accurate representation of the xc energy, since at full coupling strength, the xc hole is expected to localize more strongly around its electron [243].[8] Eq. (2.39) thus becomes

$$E_{\mathrm{xc}} = \frac{1}{2}E_{\mathrm{x}} + \frac{1}{2}E_{\mathrm{xc}}^{\mathrm{GGA}}. \tag{2.41}$$

PBE0 Functional

A more refined, yet still simple, model for the coupling strength dependence of the xc energy has been proposed in Ref. [244],

$$E_{\mathrm{xc},\lambda} = E_{\mathrm{xc},\lambda}^{\mathrm{GGA}} + \left(E_{\mathrm{x}} - E_{\mathrm{x}}^{\mathrm{GGA}}\right)(1-\lambda)^{n-1} \tag{2.42}$$

which is exact for $\lambda = 0$ ($E_{\mathrm{xc},\lambda=0}^{\mathrm{GGA}} = E_{\mathrm{x}}^{\mathrm{GGA}}$), and recovers the GGA limit for $\lambda = 1$, similarly to the half-and-half approach. However, the form (2.42) is more general, since it allows for an arbitrary power-law behavior to be reproduced for intermediate strengths λ. For molecules, the fourth-order Møller-Plesset proves able to yield accurate atomization energies; if the assumption is made that the xc energy is equally well evaluated to within the same perturbation order, one can estimate $n \approx 4$, and, after integration of (2.42) over the connection states ensemble, the following approximations are obtained

$$E_{\mathrm{x}} = \frac{1}{4}E_{\mathrm{x}} + \frac{3}{4}E_{\mathrm{x}}^{\mathrm{GGA}} \tag{2.43}$$

$$E_{\mathrm{c}} = E_{\mathrm{c}}^{\mathrm{GGA}}. \tag{2.44}$$

If the GGA contribution is chosen so that the HEG limit is correctly reproduced, i.e. through the PBE functional, one finally obtains the so-called PBE0 functional [244],

$$E_{\mathrm{xc}}^{\mathrm{PBE0}} = \frac{1}{4}E_{\mathrm{x}} + \frac{3}{4}E_{\mathrm{x}}^{\mathrm{PBE}} + E_{\mathrm{c}}^{\mathrm{PBE}} \tag{2.45}$$

which is obtained from the general form (2.36) after identifying the exchange fraction to be $\alpha = 1/4$, and choosing the GGA approximation to be PBE. The PBE0 functional has proven to be superior to the LDA or GGA in reproducing several thermochemical and spectroscopic properties of molecules [2, 82, 244].

[8] This argument has been proven valid by explicit calculations in real molecular systems: in Ref. [81], the coupling strength dependence of the xc dissociation energy of the N_2 molecule is reported to be well approximated by the LDA or GGA in the $\lambda = 1$ limit, while being substantially underestimated for $\lambda = 0$. It is in this limit that local and semilocal approximations cannot provide an accurate description (the exchange energy is highly nonlocal in the noninteracting case).

Three-Parameter Hybrid Functionals: B3LYP

Another very popular form of hybrid functionals is based on the three-parameter mixing [20]

$$E_{\mathrm{xc}} = E_{\mathrm{xc}}^{\mathrm{LDA}} + a_0 \left(E_{\mathrm{x}} - E_{\mathrm{x}}^{\mathrm{LDA}}\right) + a_{\mathrm{x}} \Delta E_{\mathrm{x}}^{\mathrm{GGA}} + a_{\mathrm{c}} \Delta E_{\mathrm{c}}^{\mathrm{GGA}}, \qquad (2.46)$$

where $\Delta E_{\mathrm{x}}^{\mathrm{GGA}}$ and $\Delta E_{\mathrm{c}}^{\mathrm{GGA}}$ are the exchange and correlation GGA corrections to $E_{\mathrm{x}}^{\mathrm{LDA}}$ and $E_{\mathrm{c}}^{\mathrm{LDA}}$, respectively. a_0, a_{x}, and a_{c} are empirical parameters whose values were obtained by fitting to molecular data sets; in the so-called B3PW91 functional [20], in which the B88 [18] and the PW91 [247] GGA functionals are chosen for exchange and correlation, respectively, Becke suggested the optimal parametrization to be $a_0 = 0.2$, $a_{\mathrm{x}} = 0.72$, and $a_{\mathrm{c}} = 0.81$. If the same parametrization is employed, but the PW91 correlation is replaced with the Lee-Yang-Parr (LYP) GGA correlation functional [180], one obtains the B3LYP functional [292],

$$\begin{aligned} E_{\mathrm{xc}}^{\mathrm{B3LYP}} = {} & (1 - a_0) E_{\mathrm{x}}^{\mathrm{LDA}} + a_0 E_{\mathrm{x}} \\ & + a_{\mathrm{x}} \Delta E_{\mathrm{x}}^{\mathrm{B88}} + a_{\mathrm{c}} E_{\mathrm{x}}^{\mathrm{LYP}} + (1 - a_c) E_{\mathrm{c}}^{\mathrm{LDA}}, \end{aligned} \qquad (2.47)$$

where for $E_{\mathrm{c}}^{\mathrm{LDA}}$, the Vosko-Wilk-Nusair (VWN) parametrization [313] is used. The B3LYP functional soon became the default choice for hybrid functional applications in the chemistry community, also due to its outstanding accuracy in predicting properties of several molecular [20] and solid-state systems [29, 54, 216]. However, from the conceptual point of view, it presents some disadvantages. Firstly, the three-parameter form (2.46) cannot be justified by adiabatic connection theorem arguments. Secondly, the inclusion of the semiempirical GGA-LYP correlation results in a functional that does not reduce to the LDA in the HEG limit. Thus, three-parameter hybrid functionals should in fact be regarded as semi-empirical xc approximations.

Screened-Exchange Hybrid Functionals

While full-range hybrid functionals, such as PBE0 and B3LYP, have been successfully applied for calculations of both molecules and solids, serious limitations are encountered when metallic systems are addressed. In this case, in fact, one has to deal with the logarithmic divergence of the derivative of the HF eigenvalues with respect to the crystal momentum at the Fermi surface [109]: just as in the HF scheme, the GKS approach suffers from this divergence. Furthermore, it has been recognized that the decay of the range of the exchange interaction

is only algebraically for metallic systems, compared to an exponential decay for gapped systems, and this makes GKS calculations particularly cumbersome in systems with small or vanishing gaps [132].

In order to overcome this difficulty, Heyd *et al.* proposed the concept of screened-exchange functionals, in which the exchange component is separated into a long-range (LR) and a short-range (SR) part,

$$\begin{aligned} E_{\text{xc}} = \alpha E_{\text{x}}^{\text{SR}}(\omega) + (1-\alpha) E_{\text{x}}^{\text{SR,GGA}}(\omega) \\ + \beta E_{\text{x}}^{\text{LR}}(\omega) + (1-\beta) E_{\text{x}}^{\text{LR,GGA}}(\omega) + E_{\text{c}}^{\text{GGA}}, \end{aligned} \tag{2.48}$$

where ω is an empirical parameter determining the range of the Coulomb interaction, which is split into a SR and a LR part, respectively defined, in a spatial decomposition of the Coulomb kernel, as

$$\frac{1}{|\mathbf{r}|} = \frac{\text{erfc}(\omega|\mathbf{r}|)}{|\mathbf{r}|} + \frac{\text{erf}(\omega|\mathbf{r}|)}{|\mathbf{r}|}, \tag{2.49}$$

where $\text{erf}(x) = (2/\sqrt{\pi}) \int_0^x dt\, e^{-t^2}$ and $\text{erfc}(x) = 1 - \text{erf}(x)$ are the error function and its complement to one. Various values can be chosen for the parameters α and β controlling the portion of EXX in the LR and SR exchange contributions [321, 327]. Given the problematic description of the long-range tail of the exchange interaction in periodic systems, one possibility is to constrain such contribution to be finite-ranged, i.e. to set $\beta = 0$ in Eq. (2.48). If moreover α is set to 0.25, as in the PBE0 functional (2.45), one obtains the short-range Heyd-Scuseria-Ernzerhof (HSE) functional [132]; in this functional the range is set to $\omega = 0.11\,\text{bohr}^{-1}$ [133] so as to perform at least as good as the PBE0 in the description of various molecular and solid-state properties [131, 200, 235], while being computationally more convenient.

The formulation of screened-exchange functionals definitely opened the way toward application of the hybrid functional approach to solid-state problems [29, 200, 216, 235], where, for example, they partially solve the well-known problem of band gap underestimation in semiconductors and insulators, faced within both LDA and GGA (see discussion in Section 5.1). However, there is apparently no formal justification to the success of hybrid functionals constructed for molecular calculations when applied to extended systems: the exchange fraction α in Eq. (2.36), whose value can be rationalized in molecular systems based on analogy with post-HF methods [244],

should assume a different meaning in solids. This conceptual difficulty can be satisfactorily solved by combining hybrid DFT with MBPT methods, and will be more specifically addressed in Section 3.3, after introducing the fundamentals of MBPT.

2.3 Interpretation of the Kohn-Sham Eigenvalues: Excited States and Fundamental Gap in DFT

It is often argued that static DFT is a ground-state theory: in fact, this statement applies only for DFT within the KS approach. By virtue of the HK theorem, the external potential v_{ext} entering some physical Hamiltonian $\hat{H}$ is uniquely determined by the ground-state density n_0; once v_{ext} is known, all excited states are unambiguously determined by solution of the Schrödinger equation. Hence, the ground-state density ultimately determines the full spectrum of $\hat{H}$, so that all of its eigenstates and eigenvalues can be regarded as density functionals. However, this property can be hardly exploited in practice, except in very specific cases [79] (mostly not relevant to realistic condensed-matter systems). This is due to the fact that (i) it is more difficult to derive an explicit energy functional for excited states, and (ii) there exists no variational principle for these states.

When coming to the DFT-KS approach, the fundamental issue is to devise a correct interpretation to the KS single-particle states. In general, neither the KS eigenvalues, nor the KS orbitals retain a well-defined physical significance. However, there are exceptions to this statement. In finite systems, such as atoms or molecules, it can be proven that the ionization potential I coincides with the negative eigenvalue of the highest occupied KS orbital, ε_N [9].[9] A generalization of this result to extended systems with a vanishing gap at the Fermi level (i.e., metals) is possible: in this case, it can be proven that ε_N gives the true Fermi energy of the metal, although the shape of the KS Fermi surface is in general different from the physical one [206].

In insulating systems, the question naturally arises of whether the energy gap can be characterized by means of analysis of the associated KS spectrum. The fundamental band gap E_{g} is defined as (identifying

[9] The proof makes use of the fact that in a finite system the exact density asymptotically decays as $e^{-\sqrt{2I}r}$ for $r \to \infty$, while, in the same limit, the behavior of the KS density is dominated by the eigenvalue of the highest occupied KS orbital, ε_N, $e^{-\sqrt{-2\varepsilon_N}r}$. By virtue of the KS ansatz, the two densities are identical, and as a consequence $I = -\varepsilon_N$.

the neutral solid with the N-particle system)

$$E_{\mathrm{g}} = -\left[(E_0(N) - E_0(N-1)) - (E_0(N+1) - E_0(N))\right], \quad (2.50)$$

where $E_0(N)$ is the ground-state energy of the N-particle system, so that the energy differences in parenthesis represent the binding energy of the most weakly bound electron in the N and the $(N+1)$-particle systems, corresponding to the negative ionization potential (I) and electron affinity (A), respectively.

In order to derive an analogous expression for E_{g} within HK-DFT, one needs to generalize the definition of the HK energy functional to systems with a noninteger number of particles [246]. For example, if the particle number is let to vary in the interval $[N, N+1)$, then the total energy can be expressed in the form

$$E(N+\eta) = (1-\eta)E(N) + \eta E(N+1) \quad (2.51)$$

with $0 \leq \eta < 1$. One then recognizes that the total energy depends on the fractional particle number in a peculiar way: (i) it is linear in the interval $(N-1, N]$ with slope $\mu(N^-)$, and in the interval $[N, N+1)$ with slope $\mu(N^+)$, (ii) it is continuous at N, (iii) its derivative with respect to particle number is discontinuous, $\mu(N^-) \neq \mu(N^+)$. $\mu(N^-)$ is identified with the chemical potential of a system with particle number approaching N from the left, in $(N-1, N]$, which defines the ionization potential I. Similarly, $\mu(N^+)$ is the chemical potential of a system with particle number approaching N from the right, in $[N, N+1)$: it is identified with the electron affinity A; in summary,

$$\lim_{\eta \to 0^+} \mu(N-\eta) = \mu(N^-) = E(N) - E(N-1) = -I \quad (2.52)$$

$$\lim_{\eta \to 0^+} \mu(N+\eta) = \mu(N^+) = E(N+1) - E(N) = -A. \quad (2.53)$$

One can interpret $\mu(N \pm \eta)$ as the Lagrange multiplier enforcing the proper normalization of the density for the system with particle number $N \pm \eta$, in a similar way as in the variational equation (2.6). The variational principle thus yields

$$\mu(N \pm \eta) = \left.\frac{\delta E[n]}{\delta n(\mathbf{r})}\right|_{N \pm \eta} \quad (2.54)$$

where $E[n]$ is the generalization of the HK energy functional to fractional particle numbers. Thus, one can re-express the fundamental

band gap as

$$E_{\mathrm{g}} = -(I - A) = \lim_{\eta \to 0^+} \left[\mu(N + \eta) - \mu(N - \eta)\right]. \tag{2.55}$$

By using Eq. (2.54), together with the standard decomposition of the total energy functional within the KS approach, Eq. (2.17), one arrives at

$$\begin{aligned} E_{\mathrm{g}} = \lim_{\eta \to 0^+} \Bigg\{ & \left[\left. \frac{\delta T_{\mathrm{s}}[n]}{\delta n(\mathbf{r})} \right|_{N+\eta} - \left. \frac{\delta T_{\mathrm{s}}[n]}{\delta n(\mathbf{r})} \right|_{N-\eta} \right] \\ & + \left[\left. \frac{\delta E_{\mathrm{xc}}[n]}{\delta n(\mathbf{r})} \right|_{N+\eta} - \left. \frac{\delta E_{\mathrm{xc}}[n]}{\delta n(\mathbf{r})} \right|_{N-\eta} \right] \Bigg\} \end{aligned} \tag{2.56}$$

in which the continuous dependence of E_{ext} and E_{H} on particle number leads to cancellation of the corresponding derivative terms.

For the noninteracting KS system, the fundamental gap is simply the difference in energy between the first unoccupied and the last occupied eigenvalues,

$$E_{\mathrm{g}}^{\mathrm{KS}} = \varepsilon_{N+1} - \varepsilon_N, \tag{2.57}$$

since in this case ionization energies are identical to KS eigenvalues, $E_0(N) - E_0(N-1) = \varepsilon_N$ in the noninteracting system. The KS gap is equivalently expressed as the derivative discontinuity of the noninteracting kinetic energy T_{s}, corresponding to the first term in parenthesis in Eq. (2.56). Hence, the fundamental gap is the sum of two contributions [245, 274],

$$E_{\mathrm{g}} = E_{\mathrm{g}}^{\mathrm{KS}} + \Delta_{\mathrm{xc}} \tag{2.58}$$

where

$$\Delta_{\mathrm{xc}} = \lim_{\eta \to 0^+} \left[\left. \frac{\delta E_{\mathrm{xc}}[n]}{\delta n(\mathbf{r})} \right|_{N+\eta} - \left. \frac{\delta E_{\mathrm{xc}}[n]}{\delta n(\mathbf{r})} \right|_{N-\eta} \right] \tag{2.59}$$

is the derivative discontinuity of the xc energy functional. Thus, the KS gap differs from E_{g} by Δ_{xc}, which is expected to account for a substantial part of the discrepancy between the LDA/GGA computed band gap[10] and the experimental one (the exchange part of the discontinuity, Δ_{x} is as large as 5.62 eV in silicon [289]). Notice that, even if the exact energy functional was known, the gap obtained with a standard KS calculation needs not be equal to the fundamental gap.

[10] Or any other xc functional well defined within the pure KS scheme.

This is because the KS approach applied to systems with integer particle numbers is inherently unable to provide information about Δ_{xc}.

While the LDA or GGA are known to yield band gaps significantly underestimated with respect to experiments, orbital-dependent functionals, such as hybrid functionals, have been found to provide values in more quantitative agreement with experiment [29, 200, 216, 235]. How can one rationalize this improvement? In fact, since the KS gap is not uniquely related to the fundamental gap, this may be at first interpreted as due to a naive compensation of the GGA gap underestimation with the corresponding overestimation provided by the HF-like contribution. However, a more precise and rigorous answer can be given by resorting to the GKS scheme [271]. Since hybrid functionals are well defined within the GKS scheme, the above derivative discontinuity argument does not directly apply. Instead, it can be proved that the GKS gap (as defined in a system with a fixed particle number N) already contains a portion of Δ_{xc}, which is instead completely missing in the standard KS gap: the discrepancy between the GKS gap and the fundamental band gap is thus expected to be smaller. For example, within the Hartree-Fock-Kohn-Sham scheme,[11] the whole exchange contribution to the derivative discontinuity would be incorporated in the GKS gap [271],

$$E_{\mathrm{g}}^{\mathrm{GKS}} = E_{\mathrm{g}}^{\mathrm{KS}} + \Delta_{\mathrm{x}}. \tag{2.60}$$

Thus, the generally better performance of hybrid functionals in electronic structure calculations is formally justified within the GKS scheme, and should not be regarded as due to a fortuitous cancellation of errors [169].

2.4 Practical DFT Calculations

In practical molecular and solid-state applications, the solution of the KS equations (2.22) needs to be pursued via numerical methods. The iterative scheme adopted for obtaining such solution generally comprises the following steps [201]:

- Starting from an initial guess for the density $n_\sigma^{(0)}(\mathbf{r}) = n_\sigma^{\mathrm{in}}(\mathbf{r})$, the KS potential $v_{\mathrm{s}}^{\mathrm{in}}$, Eq. (2.27), is evaluated (given a chosen xc approximation).

[11] A particular case of this approach is obtained by setting $\alpha = 1$ in the hybrid functional form (2.36).

- The KS equations (2.22) are solved, yielding the KS orbitals, $\phi_{\alpha\sigma}(\mathbf{r})$ and eigenvalues, $\varepsilon_{\alpha\sigma}$.
- The electronic density is computed, $n_\sigma^{\text{out}}(\mathbf{r}) = \sum_\alpha \Theta_{\alpha\sigma} \left|\phi_{\alpha\sigma}(\mathbf{r})\right|^2$.
- A new input density is constructed, usually through a suitable mixing of n^{out} with the density found at the preceding step (or steps); thus, the new KS potential is evaluated, and the self-consistency loop is reiterated until convergence at some determined level of accuracy is achieved.

Convergence to the minimum energy solution is checked by looking at the variation of total energy functional as the density or the KS potential are varied from the input to the output values.

Once the self-consistent solution is obtained, energies, eigenvalues, forces, and stresses can be calculated, and their explicit expressions depend on the details of the employed numerical scheme [201]. Plenty of numerical schemes have been developed for practical DFT calculations. Here the fundamentals of the approaches adopted in the present thesis are outlined, namely the first-principles plane waves and pseudopotential (PW-PP) and linear combination of atomic orbitals (LCAO) methods.

2.4.1 Plane Waves and Pseudopotential Method

Within the PW-PP approach, the KS equations (2.22) are represented in a plane waves basis. It is convenient to look for single-particle states that are normalized over a large volume $\mathcal{N}\Omega$ (with Ω being the volume of the primitive cell, and $\mathcal{N}$ the number of primitive cells, which is allowed to go to infinity), subject to Born-von Karman periodic boundary conditions [116]. Thus, the single-particle wavefunctions can be expanded in a basis of orthonormal plane waves characterized by a wavevector $\mathbf{q}$,

$$\phi_{\alpha\sigma}(\mathbf{r}) = \sum_{\mathbf{q}} c_{\alpha\sigma}(\mathbf{q}) \left\{ \frac{1}{\sqrt{\Omega}} \mathrm{e}^{i\mathbf{q}\cdot\mathbf{r}} \right\} \tag{2.61}$$

where $c_{\alpha\sigma}(\mathbf{q})$ are expansion coefficients. In a crystal, due to translational invariance, the effective KS potential is periodic (with the periodicity fixed by the primitive cell) and can thus be expressed as

$$v_{\text{s}}(\mathbf{r}) = \sum_{\mathbf{G}_n} v_{\text{s}}(\mathbf{G}_n) \mathrm{e}^{i\mathbf{q}\cdot\mathbf{r}} \tag{2.62}$$

where $\mathbf{G}_n$ are the reciprocal lattice vectors and $v_{\rm s}(\mathbf{G}_n)$ is the Fourier transform of the KS potential,

$$v_{\rm s}(\mathbf{G}) = \frac{1}{\Omega}\int_{\Omega} d^3r\, v_{\rm s}(\mathbf{r}) \mathrm{e}^{-i\mathbf{G}\cdot\mathbf{r}}. \tag{2.63}$$

If one restricts the crystalline momentum $\mathbf{k}$ to lie in the first Brillouin zone (BZ), $\mathbf{q} = \mathbf{k}+\mathbf{G}_n$, then a matrix representation of the KS equations is obtained in reciprocal space,

$$\sum_{n'} H_{nn'}(\mathbf{k}) c_{\alpha\sigma,n'}(\mathbf{k}) = \varepsilon_{\alpha\sigma}(\mathbf{k}) c_{\alpha\sigma,n}(\mathbf{k}) \tag{2.64}$$

where

$$H_{nn'}(\mathbf{k}) = \langle \mathbf{k}+\mathbf{G}_n | \, \hat{H}_{\rm s} \, | \mathbf{k}+\mathbf{G}_{n'} \rangle = \frac{1}{2} |\mathbf{k}+\mathbf{G}_n|^2 \, \delta_{nn'} + v_{\rm s}(\mathbf{G}_n - \mathbf{G}_{n'}). \tag{2.65}$$

Thus, a separate Schrödinger-like equation has to be solved for each $\mathbf{k}$ point in the BZ, from which one obtains the single-particle KS eigenvalues $\varepsilon_{\alpha\sigma}(\mathbf{k})$ and orbitals,

$$\phi_{\alpha\sigma,\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}_n} c_{\alpha\sigma,n}(\mathbf{k}) \left\{ \frac{1}{\sqrt{\Omega}} \mathrm{e}^{i(\mathbf{k}+\mathbf{G}_n)\cdot\mathbf{r}} \right\}. \tag{2.66}$$

In the limit of large N, the $\mathbf{k}$-point ensemble becomes a dense continuum and the eigenvalues $\varepsilon_{\alpha\sigma}(\mathbf{k})$ give rise to continuous electronic bands. Actually, once the point-group symmetries of the crystal are taken into account, one needs to compute solutions of (2.64) only for $\mathbf{k}$ points belonging to the irreducible BZ (IBZ).

In actual calculations, the expansion (2.66) is truncated to a finite number of plane wave components, by limiting the kinetic energy through a energy cutoff $E_{\rm cut}$,

$$\frac{1}{2} |\mathbf{k}+\mathbf{G}_n|^2 < E_{\rm cut}. \tag{2.67}$$

The choice of $E_{\rm cut}$ clearly affects the accuracy of the computed KS orbitals, as well as of any other quantity, such as the density and total energy. In particular, calculation of core electron wavefunctions requires a huge number of plane waves to be included in the sum (2.66). The energy cutoff determines the computational cost of DFT calculations within the PW-PP approach: for example, the number of operations needed to evaluate the density scales roughly as $N_G \log N_G$

using fast Fourier transform techniques, with N_G being the number of plane wave components needed to describe the density (which is roughly an order of magnitude larger than the number of plane wave components needed in (2.66) at the same level of accuracy). For this reason, application of plane waves methods inevitably calls for usage of pseudopotentials to treat the electron-ion Coulomb interaction. Pseudopotentials allow one to focus only on description of the smooth valence electron wavefunctions, and thus to limit the number of plane wave components included in the expansion (2.66).

In hybrid functional calculations, the evaluation of EXX introduces further complications. In fact, one has to compute the nonlocal EXX energy,

$$E_{\mathrm{x}} = -\frac{1}{\mathcal{N}} \sum_{\substack{\alpha\sigma,\alpha'\sigma' \\ \mathbf{k},\mathbf{k}'}} \Theta_{\alpha\sigma}\Theta_{\alpha'\sigma'} \times \int d^3r d^3r' \frac{\phi^*_{\alpha\sigma,\mathbf{k}}(\mathbf{r})\phi_{\alpha'\sigma',\mathbf{k}'}(\mathbf{r})\phi^*_{\alpha'\sigma',\mathbf{k}'}(\mathbf{r}')\phi_{\alpha\sigma,\mathbf{k}}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \tag{2.68}$$

where the summations run over all the occupied bands and the $\mathbf{k}$ points belonging to the IBZ. In a PW-PP implementation, an auxiliary grid of reciprocal space points $\mathbf{q}$ is defined in the BZ, so that for each $\mathbf{k}$ point, the summation is restricted to the set $\mathbf{k}' = \mathbf{k} + \mathbf{q}$ (this approach is implemented in the QUANTUM ESPRESSO code [103]). The divergence of the resulting matrix elements appearing for $\mathbf{q} = 0$ would require very dense reciprocal space grids in order to accurately evaluate the EXX contribution: the issue can be circumvented by applying the Gygi-Baldareschi integration technique [118]. In general, however, the evaluation of the EXX term remains computationally demanding within the PW-PP scheme, making hybrid functional calculations with plane waves basis sets at least one order of magnitude more costly than nonhybrid ones.

Pseudopotentials

Pseudopotentials (PPs) represent a fundamental ingredient of any electronic structure methods based on plane waves. The basic idea is to replace the strong Coulomb potential of the nucleus with a softer effective ionic potential acting on the valence electrons. One then has to define a core region, in which the single-electron wavefunction is not

treated exactly, and a scattering or valence region, for which one aims at reproducing some desired scattering properties of the physical nuclear potential, as obtained either from experiments or from first-principles calculations.

The concept of PP originates from the orthogonalized plane wave (OPW) method [130], which, under very general assumptions, allows one to partition the valence atomic wavefunction $\psi^v_{lm}(\mathbf{r})$ (labeled by the angular momentum quantum numbers l and m) into a spatially smooth part $\widetilde{\psi}^v_{lm}(\mathbf{r})$ (with the same angular momentum), and functions $u_{lmj}(\mathbf{r})$ localized around the nucleus and orthogonal to $\psi^v_{lm}(\mathbf{r})$,

$$\psi^v_{lm}(\mathbf{r}) = \widetilde{\psi}^v_{lm}(\mathbf{r}) + \sum_j B_{lmj} u_{lmj}(\mathbf{r}). \tag{2.69}$$

If the localized functions are chosen to be the core atomic orbitals, $u_{lmj} = \psi^c_{lm}$, then by virtue of the orthogonalization property, one obtains a smooth function $\widetilde{\psi}^v_{lm}$ which has no radial nodes, thus being smoother than the true ψ^v_{lm}. The core orbitals ψ^c_{lm} can largely be considered as unchanged when passing from the isolated atom to a molecular or solid-state environment. If the ansatz (2.69) is inserted into the atomic Schrödinger equation for the valence wavefunctions

$$\hat{H}\psi^v_{lm}(\mathbf{r}) = \left[-\frac{\nabla^2}{2} + v_{\mathrm{a}}(\mathbf{r})\right]\psi^v_{lm}(\mathbf{r}) = \varepsilon^v_{lm}\psi^v_{lm}(\mathbf{r}) \tag{2.70}$$

where v_{a} is the total atomic potential,[12] one obtains an effective Schrödinger equation for the smooth wavefunctions through a transformation proposed by Phillips and Kleinman [250] and Antoncik [10] (PKA transformation)

$$\hat{H}_{\mathrm{PKA}}\widetilde{\psi}^v_{lm}(\mathbf{r}) = \left[-\frac{\nabla^2}{2} + \hat{v}_{\mathrm{PKA}}\right]\widetilde{\psi}^v_{lm}(\mathbf{r}) = \varepsilon^v_{lm}\widetilde{\psi}^v_{lm}(\mathbf{r}) \tag{2.71}$$

where $\hat{v}_{\mathrm{PKA}} = v_{\mathrm{a}} + \hat{v}_{\mathrm{R}}$ is a nonlocal potential operator which turns out to be much weaker than the bare atomic potential v_{a}. In fact, the nonlocal operator $\hat{v}_{\mathrm{R}}$ acts on the smooth valence wavefunctions $\widetilde{\psi}^v_{lm}(\mathbf{r})$ by introducing a repulsive potential which largely cancels the strong bare Coulomb attraction of the nucleus. As a result, the PP often turns out to be very weak inside the core region. Exploiting both the formal properties of the PKA potential and the possibility of choosing different

[12] The potential $v_{\mathrm{a}}(\mathbf{r})$, which embodies the nuclear attractive Coulomb potential and the electron-electron interaction, is the KS potential obtained from an atomic DFT calculation.

PPs reproducing the same scattering properties outside the core region, one can construct PPs which are smoother and weaker than the original atomic potential $v_{\rm a}$.

Model ionic PPs have been developed which replace the potential of the nucleus and of core electrons, while reproducing the scattering properties of the PKA PP operator. By virtue of the atomic spherical symmetry, each angular momentum sector (l, m) can be treated separately, so that one can define a semilocal (SL) PP of the form

$$\hat{v}_{\rm SL} = \sum_{lm} |Y_{lm}\rangle \, v_l(\mathbf{r}) \, \langle Y_{lm}| \,, \tag{2.72}$$

where $|Y_{lm}\rangle$ denotes the spherical harmonic function (l, m), and $v_l(\mathbf{r})$ is a l-dependent potential which is local in the radial variable. For each l one then has to construct a different potential, a task that can be pursued either by enforcing fitting to atomic data (leading to definition of empirical pseudopotentials) [121] or by reproducing the valence properties calculated *ab initio* for the atom. The latter represents the so-called *ab initio* pseudopotential method: among the many methods to construct *ab initio* PPs, the norm-conservation approach proves successful in providing PPs that are excellently transferable to molecular and solid-state environments.

In the norm-conserving PP method, the valence pseudo-wavefunctions are chosen to be orthonormal and normalized (at odds with the valence functions of the PKA approach), being solutions of a model potential constructed to reproduce the valence atomic properties as obtained from an all-electron calculation. Hamann, Schlüter, and Chiang proposed in Ref. [119] a list of mathematical conditions that the valence pseudo-wavefunctions should satisfy in order to fit the behavior of the all-electron wavefunctions, and to enhance transferability of the resulting PP.

After performing an all-electron atomic DFT calculation with a chosen approximation for exchange and correlation, one has to identify the valence states and to generate the PP components $v_l(\mathbf{r})$ and the corresponding pseudo-wavefunctions $\psi_l^{\rm PS}(\mathbf{r})$ for each l. The $\psi_l^{\rm PS}(\mathbf{r})$ functions are numerically constructed to fit the desired properties of the all-electron valence wavefunctions outside a core radius $R_{\rm c}$. The choice of $R_{\rm c}$ ultimately determines the level of accuracy and transferability of the PP (the smaller is $R_{\rm c}$, the more accurate is the description of the wavefunction close to the nucleus), as well as its smoothness (a large $R_{\rm c}$ allows one to describe the valence pseudo-wavefunction using a limited

number of plane wave components). The potential $v_l(\mathbf{r})$ is then obtained by numerical inversion of the Schrödinger equation for each ψ_l^{PS}. In the final step, the Hartree and xc potentials for the valence electrons have to be subtracted from the potential $v_l(\mathbf{r})$, in order to obtain a bare ionic PP that is transferable to different environments. Inclusion of scalar relativistic and spin-orbit coupling effects can be performed by solving the Dirac-Kohn-Sham equation instead of the nonrelativistic one in the atomic calculation. A number of recipes have been proposed in the literature for constructing *ab initio* norm-conserving PPs [13, 151, 304, 310], which differ by the choice of the core radius, the conditions required to be satisfied by the pseudo-wavefunctions, and the numerical description of the PP components.

In an approach developed by Kleinman and Bylander [154], the semilocal part of the potential $v_l(\mathbf{r})$, Eq. (2.72), is approximated through a fully nonlocal PP operator $\hat{v}_{\mathrm{NL}}(\mathbf{r}, \mathbf{r}')$, which is separable as a sum of products of the form $\sum_i f_i(\mathbf{r}) g_i(\mathbf{r}')$. Construction of this type of PP for several atoms has been performed by Goedecker and coworkers [111, 123]. PPs generated within this scheme are mainly adopted in the PW-PP based calculations presented in this thesis.

Alternative procedures for constructing PPs that do not belong to the class of norm-conserving PPs have been devised, and the corresponding PPs are also commonly used in plane waves based electronic structure calculations. Ultrasoft pseudopotentials [24, 311] are constructed to maximize the smoothness of the potential, i.e. to minimize the number of plane waves components needed to describe the valence properties to a given accuracy. Another method which has become very popular in electronic structure calculations of solids is the projector augmented wave (PAW) method [25, 164], which also allows one to reconstruct the full all-electron wavefunction once calculation of the smooth part has been carried out. In the calculations presented in the following chapters, only norm-conserving PPs have been used.

2.4.2 Linear Combination of Atomic Orbitals Method

A different approach to representation of the electronic wavefunction in density functional calculations of crystalline systems is based on usage of localized atom-centered orbitals as basis sets. This method is implemented, for example, in the CRYSTAL09 code [77, 78] used for most of the hybrid functional calculations presented in this thesis. In fact, usage of small-sized localized orbitals basis sets allows one to

considerably reduce the cost of evaluating EXX with respect, e.g., to plane waves based methods. The reduction in CPU time is of roughly two orders of magnitude when passing from PW-PP to LCAO methods if hybrid functional calculations are performed.

On the other hand, the LCAO method also has its own limitations: the choice of the basis set should be tailored to the specific system in order to be accurate and efficient. Convergence issues can arise if this choice is not made properly. This system-based approach conceptually implies loss of generality with respect to PW-PP methods, in which a single form of the basis set is in principle suitable to describing any system. However, from the practical point of view, great experience has been acquired in constructing accurate localized basis sets, so that the *ab initio* LCAO method generally provides accurate results (see Section 5.1.2).

In the tight-binding LCAO approach [116], the solutions of the single-particle KS equations (2.22) are sought for in the form of a linear combination of a set of orbitals $\chi_\mu(\mathbf{r}-\mathbf{R}_I)$ localized at atomic position $\mathbf{R}_I$; in a periodic system $\mathbf{R}_I$ can be equivalently written as

$$\mathbf{R}_I = \tau_{\kappa,j} + \mathbf{T}, \tag{2.73}$$

where $\tau_{\kappa,j}$ ($j = 1, \ldots, n^\kappa$) are the atomic positions of the atoms of type κ in the primitive cell, and $\mathbf{T}$ is a direct lattice vector (expressed with respect to some chosen origin of the coordinate system). Together with the translation vector $\mathbf{T}$, the composite index $\{\kappa, j, \mu\} \equiv n$ allows the entire basis set to be be specified; thus, one can define a basis Bloch function with wavevector $\mathbf{k}$ through a lattice sum,

$$\chi_{n,\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{T}} \mathrm{e}^{i\mathbf{k}\cdot\mathbf{T}} \chi_n[\mathbf{r} - (\tau_n + \mathbf{T})], \tag{2.74}$$

where the summation is extended to all the direct lattice vectors $\mathbf{T}$. The KS wavefunction with wavevector $\mathbf{k}$ is written as a linear combination of the basis functions $\chi_{n,\mathbf{k}}(\mathbf{r})$,

$$\phi_{\alpha\sigma,\mathbf{k}}(\mathbf{r}) = \sum_n c_{\alpha\sigma,n}(\mathbf{k}) \chi_{n,\mathbf{k}}(\mathbf{r}), \tag{2.75}$$

where the coefficients $c_{\alpha\sigma,n}(\mathbf{k})$ are solutions of the Schrödinger-like equation,

$$\sum_{n'} \left[H_{nn'}(\mathbf{k}) - \varepsilon_{\alpha\sigma}(\mathbf{k}) S_{nn'}(\mathbf{k}) \right] c_{\alpha\sigma,n}(\mathbf{k}) = 0 \tag{2.76}$$

and the matrix elements of the KS Hamiltonian read

$$H_{nn'}(\mathbf{k}) = \langle \chi_{n,\mathbf{k}} | \, \hat{H}_{\mathrm{s}} \, | \chi_{n',\mathbf{k}} \rangle \, , \tag{2.77}$$

and

$$S_{nn'}(\mathbf{k}) = \langle \chi_{n\mathbf{k}} | \chi_{n'\mathbf{k}} \rangle \, . \tag{2.78}$$

One possible choice for the basis set is done by expressing the radial part of the localized basis functions χ_n as a linear combination of a certain number n_{G} of normalized Gaussian-type functions [27, 294],

$$\chi_n(\mathbf{r}) = \sum_i^{n_{\mathrm{G}}} d_i \, G(\alpha_i; \mathbf{r}), \tag{2.79}$$

where α_i are the exponents of the Gaussian, and d_i are the coefficients of the expansion, which are defined for any atom type (or, more generally, atom) in the system. In the CRYSTAL09 code, the localized basis functions χ_n are grouped into shells, each of which is identified by the principal and the angular momentum quantum numbers. The angular part is described by spherical harmonics.

Due to the mathematical simplicity of Gaussian functions, all the multi-center integrals involved in evaluation of the matrix elements of the Hamiltonian (and, in particular, of the Coulomb operator) can be expressed analytically [264], and numerically approximated to within a determined accuracy [77]. The possibility of analytical computation of the exchange Coulomb integral makes Gaussian basis sets particularly appealing for solid state calculations within HF and hybrid DFT.

Today, a number of Gaussian basis sets optimized for solid state calculations are available in the literature, which are obtained from molecular or atomic basis sets after removing the most diffuse functions. Since the atomic environment in a molecule or a crystal is not spherically symmetric, it is in general necessary to augment the basis set with higher angular momentum functions (polarization functions) with respect to the minimal basis set needed to describe the atom. Another possibility is to describe polarization effects by modifying the spatial dependence of the radial function at fixed angular momentum. The latter approach turns out to be best able to capture the lowest order effects of the nonspherical environment, and is often adopted in solid state calculations.

The limited computational cost of solving KS equations within the LCAO approach makes it possible to perform all-electron calculations on extended systems containing as many as hundreds electrons. On the other hand, the use of shape-consistent PPs [48, 163], analogous

to norm-conserving PPs, or energy-consistent [76] PPs (effective core potentials), typically generated at the HF level, turns out to be necessary when systems with heavy atoms are to be treated. In fact, on the one hand the computational cost of an all-electron calculation is too high for these systems even using the LCAO method, on the other hand relativistic effects, which are important for core electrons of heavy atoms, are not taken into account in standard all-electron calculations, while they can be consistently included in the generation of the PP.

CHAPTER 3

Many-Body Perturbation Theory (MBPT), and MBPT-Derived DFT Methods

3.1 Green's Function Theory

Thanks to its versatility and affordable computational cost, DFT has become a very popular method for calculating ground state properties and electronic structures of realistic condensed-matter systems. However in principle, DFT does not lend itself to a rigorous description of electronic excitations, i.e. it does not provide a formal theoretical apparatus through which one can derive consistent approximations to electron and hole excitation and ionization energies involved, e.g., in optical processes [226] (exceptions to this statement has been discussed in Section 2.3). In order to account for such phenomena, one takes advantage of the concept of a quasiparticle (QP) as a fictitious particle subject to an effective interaction [174, 175] which embodies all the quantum many-body effects. One then derives an effective Schrödinger-like single-particle equation (the QP equation), whose solution would give effective single-particle energy levels; excitation and decay processes involving such quasielectron and quasihole states may then be rigorously described.

Although this picture may at first seem similar to the one obtained within the DFT-KS theoretical scheme, the underlying rationale is indeed different: the KS theory is to be regarded as a purely mathematical construct to conveniently represent the ground-state density of a quantum system, and, with some exceptions, cannot give a fundamental description of its excited states.

A rigorous mathematical treatment of QP properties in condensed matter is the subject of many-body perturbation theory [88, 127, 293] (also known as Green's function theory), whereby one can in principle construct systematically improved approximations to the QP equation and thus calculate excitation and ionization energies, optical absorption spectra, and related quantities in molecular and condensed-matter systems [226]. One of such approximations is known as the GW approximation [11, 126]. This method constitutes the state-of-the-art approach for electronic structure calculations of materials, and the basis for all the MBPT-based calculations presented in this work. QP energies computed within the GW scheme are suitable to comparison with the measured ones obtained from photoemission spectroscopies.

3.1.1 Single-Particle Green's Function and the Self-Energy

In the simplest case, in which effective single-particle states in the many-electron system are addressed, one is interested in calculating the single-electron Green's function G, which at zero temperature is defined as an expectation value over the N-electron ground state $\left|\Psi_0^N\right\rangle$,

$$G(\mathbf{r}\sigma t, \mathbf{r}'\sigma' t') = -i \left\langle \Psi_0^N \right| T\hat{\psi}(\mathbf{r}\sigma t)\hat{\psi}^\dagger(\mathbf{r}'\sigma' t') \left| \Psi_0^N \right\rangle \tag{3.1}$$

where $\hat{\psi}(\mathbf{r}\sigma t)$ ($\hat{\psi}^\dagger(\mathbf{r}\sigma t)$) is the field operator in the Heisenberg representation annihilating (creating) one electron of spin σ, at position $\mathbf{r}$ and time t,[1]

$$\hat{\psi}(\mathbf{r}\sigma t) = \mathrm{e}^{i\hat{H}t}\hat{\psi}(\mathbf{r}\sigma)\mathrm{e}^{-i\hat{H}t} \tag{3.2}$$

and T denotes time-ordering of the two field operators. The following interpretation is apt to Eq. (3.1). For $t > t'$, an electron with spin σ' is added at time t' and position $\mathbf{r}'$ to the N-electron ground state $\left|\Psi_0^N\right\rangle$; the resulting $(N+1)$-electron state evolves in time according to the Hamiltonian $\hat{H}$ until time t, at which an electron with spin σ is removed from the system at position $\mathbf{r}$. For $t < t'$, an electron with spin σ is

[1] In a molecular or solid-state system, the N-electron Hamiltonian $\hat{H}$ should be identified with the electronic part $\hat{H}_\mathrm{e}$ of the total Hamiltonian, Eq. (1.3), obtained under the Born-Oppenheimer approximation. The associated eigenvalue problem for $\hat{H}$ is denoted $\hat{H}\left|\Psi_l^N\right\rangle = E_l^N\left|\Psi_l^N\right\rangle$.

annihilated at time t and position $\mathbf{r}$; after evolution of the resulting $(N-1)$-electron state until time t', the missing electron is reintroduced into the system at position $\mathbf{r}'$ with spin σ'. The Green's function G thus describes the probability amplitude for transitions, under the action of $\hat{H}$, of an electron or a hole from the initial to the final state throughout the fully interacting many-electron system.

The Green's function contains information about ground-state fundamental properties such as the charge density,

$$n_0(\mathbf{r}t) = -i \sum_{\sigma} \lim_{\eta \to 0^+} G(\mathbf{r}\sigma t, \mathbf{r}\sigma t + \eta), \tag{3.3}$$

and the ground-state total energy (through the Galitskii-Migdal formula) [88]. In view of analysis of excited states, one also introduces the Lehmann representation of the Green's function, which reads (in frequency space)

$$G(\mathbf{r}\sigma, \mathbf{r}'\sigma, \omega) = \sum_{l} \frac{f_l(\mathbf{r}\sigma) f_l^*(\mathbf{r}'\sigma')}{\omega - [\varepsilon_l^{\text{QP}} + i\eta \,\text{sgn}(\mu - \varepsilon_l^{\text{QP}})]}, \tag{3.4}$$

where the total energy differences

$$\varepsilon_l^{\text{QP}} = \begin{cases} E_l^{N+1} - E_0^N, & \varepsilon_l^{\text{QP}} > \mu \\ E_0^N - E_l^{N-1}, & \varepsilon_l^{\text{QP}} < \mu \end{cases} \tag{3.5}$$

refer the energy of the excited state l of the $(N+1)$- and $(N-1)$-electron systems to the ground state of the N-electron system[2] ($\mu \equiv \mu(N) = \mu(N+1) + \mathcal{O}(1/N)$ is the chemical potential), and the QP amplitudes f_l are defined as

$$f_l(\mathbf{r}\sigma) = \begin{cases} \left\langle \Psi_0^N \right| \hat{\psi}(\mathbf{r}\sigma) \left| \Psi_l^{N+1} \right\rangle, & \varepsilon_l^{\text{QP}} > \mu \\ \left\langle \Psi_l^{N-1} \right| \hat{\psi}(\mathbf{r}\sigma) \left| \Psi_0^N \right\rangle, & \varepsilon_l^{\text{QP}} < \mu. \end{cases} \tag{3.6}$$

Thus, the poles of G correspond to electron addition and removal energies which can be measured in photoemission experiments in which an electron is added to (inverse photoemission) and subtracted from (direct photoemission) the original system. The associated QP amplitudes satisfy the set of coupled equations

$$\begin{aligned} &\left[-\frac{\nabla^2}{2} + v_{\text{ext}}(\mathbf{r}) - \varepsilon_l^{\text{QP}} \right] f_l(\mathbf{r}\sigma) \\ &+ \sum_{l'} \int d^3r' \, v(\mathbf{r}, \mathbf{r}') \left\langle \Psi_l^{N-1} \right| \hat{n}(\mathbf{r}') \left| \Psi_{l'}^{N-1} \right\rangle f_{l'}(\mathbf{r}\sigma) = 0. \end{aligned} \tag{3.7}$$

[2]The label l understands both spatial and spin quantum numbers.

In the limit of vanishing Coulomb interaction between electrons, $v \to 0$, the N-, $(N+1)$-, and $(N-1)$-electron wavefunctions are Slater determinants constructed from the single-particle orbitals ϕ_l of the noninteracting system with external potential v_{ext} (described by the Hamiltonian $\hat{H}_0 = \sum_{i=1}^{N} \hat{h}_0(\mathbf{r}_i)$, with $h_0(\mathbf{r}_i) = -(1/2)\nabla_i^2 + v_{\text{ext}}(\mathbf{r}_i)$), and the energy differences $\varepsilon_l^{\text{QP}}$ are identified with the corresponding single-particle eigenvalues ε_l, obtained from solution of $\hat{h}_0 \left|\phi_l\right\rangle = \varepsilon_l \left|\phi_l\right\rangle$. Hence the Green's function reduces to its noninteracting form,

$$G_0(\mathbf{r}\sigma, \mathbf{r}'\sigma, \omega) = \sum_l \frac{\phi_l(\mathbf{r}\sigma)\phi_l^*(\mathbf{r}'\sigma')}{\omega - [\varepsilon_l + i\eta\,\text{sgn}(\mu - \varepsilon_l)]}. \tag{3.8}$$

From the mathematical point of view, G is the resolvent of $\hat{H}$, $G^{-1} = \omega - \hat{H}$; similarly, G_0 is the resolvent of the noninteracting Hamiltonian $\hat{H}_0$, $G_0^{-1} = \omega - \hat{H}_0$. The two Green's functions are connected through the Dyson's equation,

$$\sum_{\sigma''} \int d^3r'' \left\{\delta_{\sigma\sigma''}\delta(\mathbf{r} - \mathbf{r}'')\left[\varepsilon_l - H_0(\mathbf{r}\sigma)\right] - \Sigma(\mathbf{r}\sigma, \mathbf{r}'\sigma', \varepsilon_l)\right\} f_l(\mathbf{r}''\sigma'') = 0, \tag{3.9}$$

where the electron self-energy Σ is defined to include all the many-body Coulomb interaction effects. To lowest order in the Coulomb interaction v, Σ consists of the direct Coulomb (Hartree) term and the Fock exchange contribution. More conveniently for practical computational purposes, the self-energy can be defined to include only interaction contributions beyond some mean-field description of the many-electron system. Its expression of course depends on the choice made for the reference noninteracting system. Most often, one chooses $\hat{H}_0$ to be the Hartree Hamiltonian: hence, $v_{\text{ext}}(\mathbf{r}) \to v_{\text{ext}}(\mathbf{r}) + v_{\text{H}}(\mathbf{r})$, ϕ_l and ε_l entering Eq. (3.8) are the Hartree eigenfunctions and eigenvalues, and the self-energy incorporates the Fock exchange and correlation contributions.

Despite the form of Eq. (3.9) is apparently analogous to that of a single-particle Schrödinger-like equation, the presence of the self-energy operator complicates it substantially. In fact, Σ is a nonlocal, energy-dependent operator, which has to be approximated by resorting to perturbative methods. If an analytic continuation of G, Σ, and related quantities is performed from the real frequency axis to the whole complex plane, one observes that the QP energies (3.5) become complex. From a physical point of view, the real part of the QP energy

gives the position of a peak in the photoemission spectrum, i.e. in the spectral function

$$A(\mathbf{r}\sigma, \mathbf{r}'\sigma', \omega) = \frac{1}{\pi} \left| \mathrm{Im}\, G(\mathbf{r}\sigma, \mathbf{r}'\sigma, \omega) \right| = \sum_l f_l(\mathbf{r}\sigma) f_l^*(\mathbf{r}'\sigma') \delta(\omega - \varepsilon_l^{\mathrm{QP}}), \tag{3.10}$$

while its imaginary part describes the peak width, i.e. measures the QP lifetime, which is made finite by presence of interactions. Such broadened features may develop from corresponding δ-like peaks exhibited in the spectrum of the related noninteracting system (from which the physical interacting system is obtained by adiabatically switching on the Coulomb interaction v); the δ-like shape of such peaks is connected to the fact the QPs are infinitely-lived in the absence of interactions. The established connection between the fully interacting and the noninteracting systems allows one to interpret a number of experimental observations obtained in spectroscopies in the framework of an effectively single-particle picture. This is well in the spirit of the original Landau's concept of quasiparticle [174, 175].

3.1.2 Linear Response and Hedin's Equations

The self-energy incorporates all the dynamical correlation and relaxation effects beyond the HF description; such effects can be equivalently understood on the basis of the dynamical response of the system to addition and removal of one particle from it.[3] In linear response, one defines the total screened potential acting on the system,

$$v_{\mathrm{tot}}(\mathbf{r}, \omega) = \int d^3r'\, \epsilon^{-1}(\mathbf{r}, \mathbf{r}', \omega) v_0(\mathbf{r}') \tag{3.11}$$

as the potential resulting from reduction of the bare external potential v_0 due to dynamical relaxation of the charge density in the system; such reduction is represented through the inverse of the dielectric function ϵ. The irreducible polarizability P is defined through

$$\epsilon(\mathbf{r}, \mathbf{r}', \omega) = \delta(\mathbf{r} - \mathbf{r}') - \int d^3r''\, v(\mathbf{r} - \mathbf{r}'') P(\mathbf{r}'', \mathbf{r}', \omega) \tag{3.12}$$

where v is the bare Coulomb potential. The polarizability operator P describes the dynamical response of the charge density to the presence of the screened potential v_{tot}, i.e. $P = \delta n / \delta v_{\mathrm{tot}}$. The reducible

[3] Such response effects are completely neglected within the HF approximation.

polarizability χ, describing charge density response to the bare external potential v_0, $\chi = \delta n/\delta v_0$, is related to P through the Dyson's equation (written in compact form),

$$\chi = P + Pv\chi. \tag{3.13}$$

Through Eqs. (3.12) and (3.13), one can obtain an explicit expression for the inverse dielectric function ϵ^{-1} in terms of the reducible polarizability,

$$\epsilon^{-1}(\mathbf{r}, \mathbf{r}', \omega) = \delta(\mathbf{r} - \mathbf{r}') + \int d^3r''\, v(\mathbf{r} - \mathbf{r}'')\chi(\mathbf{r}'', \mathbf{r}', \omega). \tag{3.14}$$

The screened Coulomb potential is defined in an analogous fashion as in Eq. (3.11),

$$W = \epsilon^{-1}v. \tag{3.15}$$

One can then set up a perturbative expansion of the self-energy in terms of the screened Coulomb potential W. The use of W instead of the bare Coulomb potential v is physically justified by the observation that physical quantities will be actually renormalized by presence of many-body screening and relaxation effects. Furthermore, usage of the finite-range screened Coulomb potential will be likely to remove the divergences encountered in a naive expansion in terms of the bare, infinite-range, Coulomb potential. Such a program was carried out by Hedin, who was able to recast the many-body problem as a formally closed set of integral equations, relating the Green's function G, the self-energy Σ, the irreducible polarizability P, the screened Coulomb potential W, and the vertex function Γ [126]. The celebrated Hedin's equations read[4]

$$G_0(1,2) = G_0(1,2) + \int d(34)\, G_0(1,3)\Sigma(3,4)G(4,2) \tag{3.16}$$

$$W(1,2) = v(1,2) + \int d(34)v(1,3)P(3,4)W(4,2) \tag{3.17}$$

$$\Sigma(12) = i\int d(34)\, G(1,3)\Gamma(3,2;4)W(4,1) \tag{3.18}$$

$$P(1,2) = -i\int d(34)G(1,3)G(4,1)\Gamma(3,4;2) \tag{3.19}$$

$$\begin{aligned}\Gamma(2,3;1) = {} & \delta(1,2)\delta(1,3) \\ & + \int d(4567)\frac{\delta\Sigma(1,2)}{\delta G(4,5)}G(4,6)G(7,5)\Gamma(6,7;3).\end{aligned} \tag{3.20}$$

[4]The numerical labels are shorthand notation for position, spin and time coordinates, i.e. $i \equiv \mathbf{r}_i\sigma_i t_i$, and $i^+ \equiv \mathbf{r}_i\sigma_i t_i + \eta$, with $\eta \to 0^+$.

Equation (3.16) is equivalent to Eq. (3.9); equation (3.17) is easily obtained from the Eqs. (3.15) and (3.12). Equation (3.19) shows that polarization of the electron gas is made up by the creation of electron-hole pairs (product of retarded and advanced Green's functions), and their interaction is determined by the vertex Γ, which is determined by the change in the self-energy upon electron transitions.

3.1.3 The GW Approximation

Hedin's equations constitute a set of coupled integral equations which are to be solved self-consistently for G. In practice, one seeks for an approximation that allows one to partially decouple this set of equations [127, 226]. This is equivalent to truncating the self-energy expansion to a given order in G. The lowest order approximation to Σ is obtained by neglecting vertex corrections, i.e. assuming $\Gamma(2,3;1) = \delta(1,2)\delta(1,3)$. In this approximation, the irreducible polarizability is built up by noninteracting electron-hole pairs,

$$P_{\text{RPA}}(1,2) = -iG(1,2)G(2,1^+), \tag{3.21}$$

i.e. electron and holes propagate independently throughout the interacting system. Such approximation for P is equivalent to the so-called random phase approximation (RPA) of many-body theory [88]. Hence Eq. (3.18) implies that Σ is given by

$$\Sigma_{GW}(1,2) = iG(1,3)W(3,1) \tag{3.22}$$

which represents the so-called GW approximation to self-energy. W is given by Eq. (3.19), with P obtained in the RPA, Eq. (3.21). Σ_{GW} should be in principle computed self-consistently with G, which satisfies the Dyson's equation (3.16). The noninteracting Green's function G_0 (3.8) can be constructed from the eigenfunctions and eigenvalues obtained from solution of the Schrödinger equation for a reference independent-particle system, and iterated until convergence is achieved. In practice, such self-consistent GW approach is not expected to systematically improve computed spectroscopic properties, with respect to the simplest G_0W_0 scheme, in which only the first iteration step is performed [167, 226]. This is particularly true for calculation of QP band gaps of insulating materials, for which self-consistent GW has been generally found to perform worse than G_0W_0, or GW_0 [281], in which only G is iterated to self-consistency, while W is kept constant to its original RPA expression. This tendency points out the importance of

the vertex correction when attempting a consistent treatment of G and W. The G_0W_0 method is more extensively described in Section 3.2, where its practical implementation into a plane waves framework is also discussed.

3.1.4 The COHSEX Approximation

In his original work of 1965, Hedin proposed an even simpler approximation for the self-energy, derivable in the static limit ($\omega \to 0$) from the GW self-energy,[5]

$$\Sigma_{GW}(\mathbf{r}, \mathbf{r}', \omega) = \frac{i}{2\pi} \int d\omega' \, \mathrm{e}^{i\eta\omega'} G(\mathbf{r}, \mathbf{r}', \omega + \omega') W(\mathbf{r}, \mathbf{r}', \omega'). \quad (3.23)$$

In fact, the frequency-dependent, polarizable part of W,

$$W_p(\mathbf{r}, \mathbf{r}', \omega) = W(\mathbf{r}, \mathbf{r}', \omega) - v(\mathbf{r}, \mathbf{r}') \quad (3.24)$$

can be written in the spectral representation,

$$W_p(\mathbf{r}, \mathbf{r}', \omega) = \int_0^\infty d\omega' \, \frac{2\omega' B(\mathbf{r}, \mathbf{r}', \omega')}{\omega'^2 - (\omega - i\eta)^2}. \quad (3.25)$$

One then observes that Eq. (3.23) can be rewritten as [127]

$$\begin{aligned} \Sigma_{GW}(\mathbf{r}, \mathbf{r}', \omega) = &- \sum_l \Theta_l \, W(\mathbf{r}, \mathbf{r}', \epsilon_l - \omega) f_l(\mathbf{r}) f_l^*(\mathbf{r}') \\ &+ \sum_l f_l(\mathbf{r}) f_l^*(\mathbf{r}') \int_0^\infty d\omega' \, \frac{B(\mathbf{r}, \mathbf{r}', \omega')}{\varepsilon_l^{\mathrm{QP}} - \omega - (\omega' - i\eta)} \end{aligned} \quad (3.26)$$

If $(\varepsilon_l^{\mathrm{QP}} - \omega)$ is smaller than the characteristic excitation energy of the screened interaction (the plasmon energy), then one can simplify Eq. (3.26) by assuming $\varepsilon_l^{\mathrm{QP}} - \omega \approx 0$. By performing the energy integration, one obtains the following (energy-independent) approximation to the GW self-energy, known as the Coulomb-hole and screened-exchange (COHSEX) approximation [126],

$$\Sigma_{GW}(\mathbf{r}\sigma, \mathbf{r}'\sigma', 0) = \Sigma_{\mathrm{COH}}(\mathbf{r}, \mathbf{r}', 0) + \Sigma_{\mathrm{SEX}}(\mathbf{r}, \mathbf{r}', 0) = \Sigma_{\mathrm{COHSEX}}(\mathbf{r}, \mathbf{r}') \quad (3.27)$$

where

$$\Sigma_{\mathrm{COH}}(\mathbf{r}, \mathbf{r}', 0) = \frac{1}{2} \delta(\mathbf{r} - \mathbf{r}') W_p(\mathbf{r}, \mathbf{r}', 0) \quad (3.28)$$

[5] For the sake of simplicity, spin coordinates and quantum numbers are dropped in the following.

is the Coulomb hole (COH) self-energy contribution arising from interaction of the electron with the Coulomb hole created by polarization of the rest of the system (dynamical effects in the screening are nonetheless neglected since W_p is evaluated for $\omega = 0$). In this approximation, the COH term is a local function in space. The static screened-exchange (SEX) contribution is given by

$$\Sigma_{\text{SEX}}(\mathbf{r}, \mathbf{r}', 0) = -\sum_l \Theta_l \, W(\mathbf{r}, \mathbf{r}', 0) f_l(\mathbf{r}) f_l^*(\mathbf{r}'), \tag{3.29}$$

which is the statically screened Fock interaction (the Fock potential is in fact obtained by replacing W with v). COHSEX goes beyond the HF theory, in the sense that it takes into account (static) polarization effects that are completely missing in the HF scheme. In fact, the HF approximation is equivalent to completely neglecting electronic screening ($\epsilon = 1$), which amounts at assuming $P = 0$ in Eq. (3.17); hence, the self-energy Σ reduces to the HF expression, $\Sigma_{\text{HF}} = iGv$. The Coulomb hole term, developing from the poles of the inverse dielectric function, is also disregarded in this approximation.

The COHSEX method generally improves computed spectroscopic properties with respect to HF, although with a tendency to overestimate QP band gaps of materials [141], as a result of overestimation of the self-energy with respect to its fully dynamical form. On the other hand, the computational cost of a COHSEX calculation is practically equivalent to HF, and much smaller than full GW, since all excitation processes from occupied to empty states are not taken into account in the evaluation of the dielectric function. Moreover, the energy-independent expression of the COHSEX self-energy operator is suitable to being implemented into an hybrid DFT scheme, as described in Section 3.3.

3.2 Practical GW Calculations

3.2.1 The G_0W_0 Approximation

In practical GW calculations, one typically considers the KS system as the reference noninteracting system from which G, P, and Σ are initially evaluated. One starts by constructing G from the KS orbitals ϕ_l and eigenvalues $\varepsilon_l^{\text{KS}}$ obtained from a DFT calculation,

$$G_{\text{KS}}(\mathbf{r}, \mathbf{r}', \omega) = \sum_l \frac{\phi_l(\mathbf{r}) \phi_l^*(\mathbf{r}')}{\omega - [\varepsilon_l^{\text{KS}} + i\eta \, \text{sgn}(\mu - \varepsilon_l^{\text{KS}})]}. \tag{3.30}$$

The RPA polarizability is obtained similarly, starting from Eq. (3.21) and performing a Fourier transformation [3, 325],

$$P_{\text{RPA}}(\mathbf{r},\mathbf{r}',\omega) = \sum_{l,l'}(\Theta_l - \Theta_{l'})\,\frac{\phi_l(\mathbf{r})\phi_{l'}^*(\mathbf{r})\phi_{l'}(\mathbf{r}')\phi_l^*(\mathbf{r}')}{\omega - (\varepsilon_{l'}^{\text{KS}} - \varepsilon_l^{\text{KS}}) + i\eta}. \tag{3.31}$$

The dielectric function ϵ is constructed from Eq. (3.12), and, once inverted, yields the screened interaction W; then, the self-energy Σ_{GW} is evaluated using Eq. (3.22).

In order to evaluate the QP energies, one then computes the first-order correction to KS eigenvalues, using $[\hat{\Sigma}_{GW}(\omega) - \hat{v}_{\text{xc}}]$ as perturbation operator,

$$\varepsilon_l^{\text{QP}} = \varepsilon_l^{\text{KS}} + \langle\phi_l|\,[\hat{\Sigma}_{GW}(\varepsilon_l^{\text{QP}}) - \hat{v}_{\text{xc}}]\,|\phi_l\rangle\,. \tag{3.32}$$

The self-energy should be evaluated at the unknown QP energy $\varepsilon_l^{\text{QP}}$. However, one can perform an expansion of $\Sigma_{GW}(\varepsilon_l^{\text{QP}})$ at first order around the corresponding KS eigenvalue [141],

$$\varepsilon_l^{\text{QP}} = \varepsilon_l^{\text{KS}} + Z_l\,\langle\phi_l|\,[\hat{\Sigma}_{GW}(\varepsilon_l^{\text{KS}}) - \hat{v}_{\text{xc}}]\,|\phi_l\rangle \tag{3.33}$$

where $Z_l = (1 - \langle\phi_l|\,d\hat{\Sigma}/d\omega\,(\omega = \varepsilon_l^{\text{KS}})\,|\phi_l\rangle)^{-1}$ is the QP weight. This approximation is generally able to provide results very close to the ones obtained by exactly diagonalizing the GW Hamiltonian. Equation (3.33) provides the G_0W_0 correction to the KS band structure. Evaluation of G, P, and Σ using the first-order approximation (3.33) to QP energies (and assuming, for example, single-particle wavefunctions to be well described within DFT, i.e. neglecting nondiagonal elements in the self-energy) leads to the first iteration step of the self-consistent GW scheme. Clearly the DFT starting point may substantially affect the accuracy of the obtained GW correction [35, 46]: if for example the starting KS wavefunctions are qualitatively wrong, i.e. are far from the actual form of the QP wavefunctions, then the corresponding matrix elements of Σ are expected to be inaccurate, and this may in principle result in a degradation with respect to the DFT description. In this respect, self-consistent GW should instead provide results independent of the starting approximation.

3.2.2 Plane waves implementation of the G_0W_0 scheme

In a plane waves and pseudopotential scheme, one resorts to a plane waves basis set to represent the starting KS wavefunctions, $\langle\mathbf{r}|l,\mathbf{k}\rangle = \phi_{l,\mathbf{k}}(\mathbf{r})$ given by Eq. (2.66), as well as matrix elements of the various

operators to be defined in the GW scheme. Such approach is employed in the BERKELEYGW code [70] used for GW calculations presented in this work. In reciprocal space, the static RPA polarizability reads

$$[P_{\text{RPA}}]_{\mathbf{GG}'}(\mathbf{q},0) = \sum_{l,l',\mathbf{k}} \Theta_l(1-\Theta_{l'}) \frac{M_{ll'}(\mathbf{k},\mathbf{q},\mathbf{G}) M^*_{ll'}(\mathbf{k},\mathbf{q},\mathbf{G}')}{\varepsilon_{l\mathbf{k}+\mathbf{q}} - \varepsilon_{l'\mathbf{k}} + i\eta} \quad (3.34)$$

where $M_{ll'}(\mathbf{k},\mathbf{q},\mathbf{G}) = \langle l, \mathbf{k}+\mathbf{q}| \, \mathrm{e}^{i(\mathbf{q}+\mathbf{G})\cdot\hat{\mathbf{r}}} \, |l', \mathbf{k}\rangle$, $\mathbf{q}$ is a vector of the first BZ, and the summation is over occupied (denoted l) and empty (denoted l') single-electron states. The size of the matrix $[P_{\text{RPA}}]_{\mathbf{GG}'}$ is determined by the dielectric energy cutoff $E^{\text{eps}}_{\text{cut}}$, which limits the summation (3.34) to reciprocal lattice vectors such that $(1/2)|\mathbf{G}|^2 < E^{\text{eps}}_{\text{cut}}$. The dielectric energy cutoff also determines an upper limit to the number of empty states that should be included in the summation (3.34). In practice, the static RPA polarizability converges quite rapidly with respect to the number of the included empty states.

The static dielectric matrix is evaluated using the definition (3.12), which in reciprocal space reads

$$\epsilon_{\mathbf{GG}'}(\mathbf{q},0) = \delta_{\mathbf{GG}'} - v(\mathbf{q}+\mathbf{G}) P_{\mathbf{GG}'} \quad (3.35)$$

where

$$v(\mathbf{q}+\mathbf{G}) = \frac{4\pi}{|\mathbf{q}+\mathbf{G}|^2} \quad (3.36)$$

is the bare Coulomb potential. The dielectric matrix is then inverted to yield the screened Coulomb interaction (at zero frequency),

$$W_{\mathbf{GG}'}(\mathbf{q},0) = \epsilon^{-1}_{\mathbf{GG}'}(\mathbf{q},0) v(\mathbf{q}+\mathbf{G}'). \quad (3.37)$$

In order to extend the dielectric response to finite frequencies, one can define the fully frequency-dependent retarded (r) and advanced (a) dielectric matrices,

$$\epsilon^{\text{r/a}}_{\mathbf{GG}'}(\mathbf{q},\omega) = \delta_{\mathbf{GG}'} - v(\mathbf{q}+\mathbf{G}) \sum_{l,l',\mathbf{k}} \Theta_l(1-\Theta_{l'}) M_{ll'}(\mathbf{k},\mathbf{q},\mathbf{G}) M^*_{ll'}(\mathbf{k},\mathbf{q},\mathbf{G}') \times \frac{1}{2}\left[\frac{1}{\varepsilon_{l\mathbf{k}+\mathbf{q}} - \varepsilon_{l'\mathbf{k}} - \omega \mp i\eta} + \frac{1}{\varepsilon_{l\mathbf{k}+\mathbf{q}} - \varepsilon_{l'\mathbf{k}} + \omega \pm i\eta}\right], \quad (3.38)$$

and evaluate the screened interaction on a grid of frequencies along the real axis. Alternatively, one can resort to plasmon-pole approximations: a simple analytic model for the frequency-dependent dielectric matrix

is assumed, in which free parameters are chosen so as to satisfy certain exact conditions. The Hybertsen-Louie generalized plasmon-pole model [141] implemented in the BERKELEYGW code [70], takes advantage of the fact that the imaginary part of the inverse dielectric matrix, $\mathrm{Im}\,\epsilon^{-1}_{\mathbf{GG'}}(\mathbf{q},\omega)$, is generally a peaked function of ω. For each set of indeces $(\mathbf{q},\mathbf{G},\mathbf{G'})$, $\mathrm{Im}\,\epsilon^{-1}$ is thus modeled as

$$\mathrm{Im}\,\epsilon^{-1}_{\mathbf{GG'}}(\mathbf{q},\omega) = A_{\mathbf{GG'}}(\mathbf{q})[\delta(\omega-\widetilde{\omega}_{\mathbf{GG'}}(\mathbf{q}))-\delta(\omega+\widetilde{\omega}_{\mathbf{GG'}}(\mathbf{q}))]. \quad (3.39)$$

The corresponding real part is given by

$$\mathrm{Re}\,\epsilon^{-1}_{\mathbf{GG'}}(\mathbf{q},\omega) = 1 + \frac{\Omega^2_{\mathbf{GG'}}(\mathbf{q})}{\omega^2 - \widetilde{\omega}^2_{\mathbf{GG'}}(\mathbf{q})} \quad (3.40)$$

where $\Omega_{\mathbf{GG'}}(\mathbf{q})$ is an effective plasma frequency. The model parameters $A_{\mathbf{GG'}}(\mathbf{q})$ and $\widetilde{\omega}_{\mathbf{GG'}}(\mathbf{q})$ are fixed by requiring $\epsilon^{-1}_{\mathbf{GG'}}(\mathbf{q},\omega)$ to satisfy the Kramer-Kronig relation, as well as by imposing a frequency sum rule on $\mathrm{Im}\,\epsilon^{-1}_{\mathbf{GG'}}(\mathbf{q},\omega)$ [141]. The resulting model dielectric function is thus completely determined, and in fact contains no empirically fitted parameters.

Similarly to COHSEX, the GW self-energy can be separated in frequency-dependent Coulomb hole and screened exchange contributions,

$$\begin{aligned}\langle l,\mathbf{k}|\,\hat{\Sigma}_{\mathrm{COH}}(\omega)\,|l',\mathbf{k}\rangle = \frac{i}{2\pi}\sum_{l''}\sum_{\mathbf{qGG'}} M^*_{l''l}(\mathbf{k},-\mathbf{q},-\mathbf{G})M_{l''l'}(\mathbf{k},-\mathbf{q},-\mathbf{G'}) \\ \times\int_0^\infty d\omega'\,\frac{\left[\epsilon^{\mathrm{r}}_{\mathbf{GG'}}\right]^{-1}(\mathbf{q},\omega') - \left[\epsilon^{\mathrm{a}}_{\mathbf{GG'}}\right]^{-1}(\mathbf{q},\omega')}{\omega-\varepsilon_{l,\mathbf{k}-\mathbf{q}}-\omega'+i\eta}\end{aligned} \quad (3.41)$$

$$\begin{aligned}\langle l,\mathbf{k}|\,\hat{\Sigma}_{\mathrm{SEX}}(\omega)\,|l',\mathbf{k}\rangle = -\sum_{l''}^{\mathrm{occ}}\sum_{\mathbf{qGG'}} M^*_{l''l}(\mathbf{k},-\mathbf{q},-\mathbf{G})M_{l''l'}(\mathbf{k},-\mathbf{q},-\mathbf{G'}) \\ \times\left[\epsilon^{\mathrm{r}}_{\mathbf{GG'}}\right]^{-1}(\mathbf{q},\omega-\varepsilon_{l,\mathbf{k}-\mathbf{q}})v(\mathbf{q}+\mathbf{G'}),\end{aligned} \quad (3.42)$$

which can be evaluated analytically when the generalized plasmon-pole model is used to describe the frequency dependence [70]. In any case, the Coulomb hole part of the GW self-energy contains an in principle infinite summation over empty states, which in practice has to be truncated. The number of empty states needed to converge the COH self-energy is typically very large, and its evaluation thus

constitutes another computationally intensive step in GW calculations (besides the evaluation of the polarizability (3.34), and the inversion of the dielectric matrix). Recently, numerical techniques that allow one to avoid the explicit calculation of excited states through density-functional perturbation theory approaches have been developed and implemented in GW codes [113, 306].

Finally, the QP energies are computed through Eq. (3.33). One then has to evaluate the matrix elements of the exchange-correlation potential v_{xc} over the KS wavefunctions, which are expanded in plane waves up to a cutoff energy $E_{\text{cut}}^{\text{xc}}$.

3.3 Combining MBPT and DFT: Self-Consistent Dielectric-Dependent Hybrid Functionals

The generally good performance of hybrid xc functionals in electronic structure calculations of insulating solids is justified by incorporation of part of the derivative discontinuity in the GKS gap [271]. Popular hybrid functionals such as PBE0, and especially HSE06, have been shown to systematically improve agreement of computed band gaps with respect to experiments, compared with LDA and GGA [216, 235]. However, the value of the admixed EXX fraction in those functionals was originally obtained by requiring them to provide molecular thermochemical properties with an accuracy comparable to higher level perturbation theory (Møller-Plesset) methods [244]. Such an approach is not necessarily justified when applied to extended systems, for which GW is the state-of-the-art method for computation of the electronic structure. In the zero-frequency limit, the GW self-energy can be put in correspondence with the general expression of the full-range PBE0-like hybrid xc potential obtained by differentiating the xc energy (2.36),

$$v_{\text{xc}}[n](\mathbf{r},\mathbf{r}') = \alpha v_{\text{x}}(\mathbf{r},\mathbf{r}') + (1-\alpha)v_{\text{x}}^{\text{GGA}}(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}') + v_{\text{c}}^{\text{GGA}}(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}') \tag{3.43}$$

where

$$v_{\text{x}}(\mathbf{r},\mathbf{r}') = \frac{\delta E_{\text{x}}[n]}{\delta n(\mathbf{r})} = \sum_l \Theta_l\, \phi_l(\mathbf{r})\phi_l(\mathbf{r}')v(\mathbf{r},\mathbf{r}') \tag{3.44}$$

is the nonlocal potential derived from the KS EXX energy, Eq. (2.32). The zero-frequency limit of $\Sigma_{GW}(\mathbf{r},\mathbf{r}',\omega)$ corresponds to the COHSEX

expression for self-energy,

$$\Sigma_{\text{COHSEX}}(\mathbf{r},\mathbf{r}') = -\sum_l \Theta_l\, W(\mathbf{r},\mathbf{r}',0)\phi_l(\mathbf{r})\phi_l^*(\mathbf{r}') - \frac{1}{2}\delta(\mathbf{r}-\mathbf{r}')\left[v(\mathbf{r},\mathbf{r}') - W(\mathbf{r},\mathbf{r}',0)\right]. \tag{3.45}$$

The statically screened Coulomb potential

$$W(\mathbf{r},\mathbf{r}',0) = \int d\mathbf{r}''\, \epsilon^{-1}(\mathbf{r},\mathbf{r}'',0)v(\mathbf{r}'',\mathbf{r}') \tag{3.46}$$

can be simplified by neglecting the microscopic components of the static dielectric function, i.e. by approximating

$$\epsilon^{-1}(\mathbf{r},\mathbf{r}',0) \approx \frac{1}{\epsilon_\infty}\delta(\mathbf{r}-\mathbf{r}'), \tag{3.47}$$

where ϵ_∞ is the macroscopic dielectric constant of the material [3, 325],

$$\epsilon_\infty = \lim_{\mathbf{q}\to 0} \frac{1}{\epsilon^{-1}_{\mathbf{G}=0,\mathbf{G}'=0}(\mathbf{q},0)} \tag{3.48}$$

evaluated at zero frequency, and comprising the electronic contribution only (ion-clamped dielectric constant). One thus obtains

$$W(\mathbf{r},\mathbf{r}',0) \approx \epsilon_\infty^{-1} v(\mathbf{r},\mathbf{r}'). \tag{3.49}$$

The screened-exchange part of Σ_{COHSEX} is thus identified with the nonlocal Fock exchange part in the hybrid xc potential (3.43), since both decay asymptotically as $1/|\mathbf{r}-\mathbf{r}'|$. Under the stated approximations, the EXX fraction is thus given by [6, 90, 198]

$$\alpha = \frac{1}{\epsilon_\infty}. \tag{3.50}$$

The self-energy contributions beyond the screened exchange term, i.e. the static Coulomb hole and the dynamical contributions not described within COHSEX, are modeled through the local, short-ranged parts of (3.43). In other words, the assumption is made that the full-range hybrid xc potential (3.43) correctly describes the long-range behavior of the self-energy, and that the admixed GGA xc potential accounts for the remaining, supposedly short-ranged, contributions.

3.3.1 Literature Review of Approaches to Determine the Exchange Fraction

Several approaches have been proposed and benchmarked in the literature to evaluate the exchange fraction α in the full-range hybrid functional (3.43) for electronic structure calculations. Two opposite philosophies can be identified: (i) fitting of α so as to reproduce some desired electronic feature, typically the material band gap, and (ii) *ab initio* or semi-empirical evaluation of α based on Eq. (3.50).

In the first class of approaches, the exchange fraction has been tuned to recover the band gap obtained from experiment [4, 6, 33] or computed within GW [45], and the resulting functional has been used for calculation of a different and more complex electronic features, such as defect levels positions in the bulk band gap and band edges alignment at semiconductor interfaces.

Approaches to determine α based on Eq. (3.50) differ from each other depending on how the dielectric constant ϵ_∞ is evaluated. Marques *et al.* computed it using the PBE functional [198]; Conesa evaluated α based on the experimental value of ϵ_∞ [50] and the value calculated at the hybrid DFT level [51]; other authors employed simplified models to obtain the static dielectric constant [277–279], or determined it within the independent-particle approximation[6] [158] or the RPA [257], sometimes employing a self-consistent approach. Skone *et al.* proposed a method for determining α and ϵ_∞ self-consistently, based on the relationship (3.50) and using both the PBE and the PBE0 functionals as starting approximation to compute the dielectric constant [283]. The latter is calculated going beyond the RPA, i.e. including the exchange and correlation contributions to polarization function.

3.3.2 Method Used for Calculation of the Dielectric Constant

The general approach to evaluate the dielectric function within DFT-GKS makes use of the definition of the irreducible KS polarizability $\chi_\mathrm{s}(\mathbf{r}, \mathbf{r}', \omega)$, defined as the density response function to a variation in the generalized KS potential, $\chi_\mathrm{s} = \delta n / \delta v_\mathrm{s}$, where $v_\mathrm{s}(\mathbf{r}, \mathbf{r}')$ is the generalization of the KS potential (2.21) within the GKS scheme, possibly including nonlocal xc potentials (derived, e.g., from hybrid xc functionals). The relationship between χ_s and the physical (reducible) polarizability χ of Eq. (3.13) is obtained along the lines of the reasoning in Section 3.1.2: in the presence of a perturbing external potential v_0, the

[6] This approximation corresponds to setting $\chi = \chi_\mathrm{s}$ in Eq. (3.51), i.e. to neglecting polarization effects.

electronic density redistributes in the system, inducing a change in the GKS electronic potential; the two response functions are related by

$$\chi = \chi_{\mathrm{s}} + \chi_{\mathrm{s}} \left(v + f_{\mathrm{xc}} \right) \chi, \tag{3.51}$$

where

$$f_{\mathrm{xc}} = \frac{\delta v_{\mathrm{xc}}}{\delta n} \tag{3.52}$$

is the exchange-correlation kernel.

By neglecting the xc contribution to polarizability, i.e. assuming $f_{\mathrm{xc}} = 0$, one is left with the RPA polarizability, which is equivalent to the GW approximation. In this approximation, the employed xc potential does enter only in the construction of the independent-particle polarizability χ_{s}, but not in the higher-order screening contributions. Since the aim here is to devise a fully self-consistent scheme to evaluate α based on its relation with ϵ_∞, and since α fixes the form of the xc potential to be used, it is desirable to consistently include the xc effects in determining the density response function. Such contributions can lead to appreciable modification of the computed optical properties, especially when f_{xc} is derived from nonlocal xc potentials [236]. Going beyond the RPA requires evaluating the derivative of v_{xc} with respect to density, which in the case of nonlocal hybrid xc potentials is not straightforward to compute, due to implicit dependence of v_{xc} on the density through the KS orbitals.

A suitable computational approach allowing one to directly evaluate this term is the coupled-perturbed Kohn-Sham (CPKS) method [85,227], an extension of the coupled-perturbed Hartree-Fock (CPHF) [86, 255] method including electronic correlation in the evaluation of the dielectric response. By means of this method, the eigenstates of the perturbed Hamiltonian as obtained from application of an external electric field are expressed in terms of the unperturbed ones; the unitary matrix connecting the two orbital manifolds is approximated within first-order perturbation theory and evaluated iteratively through the self-consistent coupled-perturbed procedure. This matrix fully includes contributions derived from possibly nonlocal energy terms in the Hamiltonian, such as the ones arising from the EXX part in hybrid xc functionals. Hence, the static polarizability is computed iteratively, and, after convergence is achieved, the dielectric constant ϵ_∞ can be readily evaluated. Due to usage of moderate size Gaussian-type basis sets, the CRYSTAL09 code offers an efficient implementation of the CPKS method for calculation of the dielectric properties of extended systems [87].

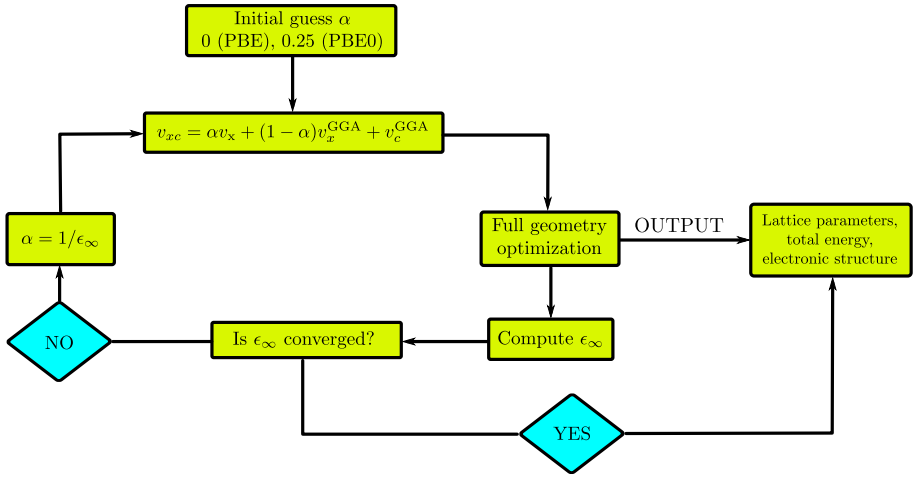

Figure 3.1: *Flowchart of the sc-PBE0α_{ϵ_∞} method.*

3.3.3 Self-Consistent Hybrid Functional (sc-PBE0α_{ϵ_∞})

Once ϵ_∞ is calculated in the first step using a suitably chosen xc approximation (which may be chosen, e.g., as PBE or PBE0), a new hybrid DFT potential can be constructed with the exchange fraction obtained from inversion of the computed dielectric constant. The procedure is then iterated until α and ϵ_∞ are consistently determined, according to the diagram in Figure 3.1. The xc functional constructed with the converged α will be referred to as "sc-PBE0α_{ϵ_∞}" in the following, as it is in fact a self-consistent (sc) re-parametrization of the full-range PBE0 functional.

Optimization of lattice parameters of the bulk solid is performed at each iteration step, so that the dielectric constant is computed at the true ground state corresponding to the xc functional at hand. In fact, structural deformations in principle affect computed electronic band gaps and dielectric constants, so that a consistent determination of structural, electronic and dielectric properties may be aimed for.

The self-consistent scheme has been shown to converge within two or three iteration steps for most semiconductors and insulators [283], and even after a single iteration the computed band gap is substantially closer to experiment with respect to the PBE or PBE0 values. Optimization of the cell geometry at each self-consistency step does not alter quantitatively how quickly this convergence is achieved [100].

CHAPTER 4

Modeling of Point Defects in Insulating Materials

4.1 Introduction: The Supercell Approximation

Atomic impurities such as substitutional atoms, atomic vacancies, or interstitials, are spontaneously formed in materials, where they are often introduced throughout the synthesis and production processes even when the latter are carried out under very well-controlled conditions. The presence of such native defects substantially alters, among others, the optical, electrical and magnetic properties of the material, and a correct interpretation of experiments requires these modifications to be properly described by theory.

When native point defects are formed into an otherwise ideal crystal, the original periodicity gets inevitably broken. However, the experimental concentration of these intrinsic defects is typically too low to be realistically modeled through a correspondingly large primitive cell representing a single impurity embedded in the ideal host lattice: the number of atoms belonging to such supercell would be of the order of several thousands; calculations on cells of this size is at or beyond the limit of applicability, in terms of

computational cost, of state-of-the-art DFT methods (e.g., hybrid functionals) using conventional implementations. In practice, one has to introduce an artificial periodicity employing much smaller supercells, which would correspond to an unrealistically large defect concentration. Consequently, the spurious interaction between image defects can substantially affect the calculated properties of the defective system, to an extent at which comparison with experiments may become meaningless. Such interaction can be quantum-mechanical (overlap of nearby defect wavefunctions), elastic, electrostatic (for nonneutral defects), or magnetic in nature, and has to be carefully accounted for when extrapolating computed quantities to very low defect concentrations (dilute limit) [93], which is typically the condition suitable for comparison with experiments for native defects. However, usage of the supercell approximation in a periodic framework is advantageous over, e.g., cluster approaches [230], since it allows one to interpret the defect as a small perturbation of the perfect periodic crystal, which is naturally described using periodic boundary conditions. In fact, it constitutes today the most popular approach to theoretical modeling of point defects in solids.

4.2 Defect Formation Energy

Within the supercell approximation, calculation of fundamental properties of the defective system is achievable. One of the most representative of these properties is the energy cost of creating a single defect in an otherwise perfect crystal. The defect formation energy is given by (in the grand canonical ensemble at zero temperature,[1] and for a neutral defect, $q = 0$)

$$E^{\mathrm{f}}_{\mathrm{D},0} = E_{\mathrm{D},0} - E_{\mathrm{H}} - \sum_i n_i \mu_i, \tag{4.1}$$

where $E_{\mathrm{D},0}$ and E_{H} are the ground-state total energies of the defective and nondefective (host) supercells, respectively. n_i is the number of atoms of species i added to ($n_i > 0$) or removed from ($n_i < 0$) the host supercell while forming the defective one. The chemical potentials $\{\mu_i\}$ of the reservoirs for the atomic species involved in the defect formation are ultimately fixed by the experimental conditions at which the defect is created. For example, one can simulate a condition in which the material

[1] For a generalization of the definition of defect formation energy to finite temperature and pressure, see Ref. [93].

is synthesized in an atmosphere predominantly enriched with a given atomic species, leading to corresponding creation of substitutional or interstitial impurities in the perfect crystal. The range of values that μ_i can span is determined by appearance of a pure (solid or gaseous) phase of the species i [309, 329] in some region of μ_i values.

In order to compute reliable defect formation energies, convergence of the supercell size should be carefully checked [93]. (i) For a neutral defect, one should take into account the quantum-mechanical overlap between defect wavefunctions in nearby supercells, possibly leading to dispersion of the corresponding defect state along high-symmetry lines in the BZ; in contrast, in the case of an ideally isolated defect, such dispersion is suppressed, as the defect state is atomic in nature. (ii) The relaxation of the host lattice atomic positions can produce elastic interactions between defects, if the host lattice structure is strongly perturbed by presence of the defect, and/or the supercell is chosen to be too small. (iii) Interactions between defects with a nonzero total spin may also alter the formation energy. Contributions (ii) and (iii) decay as $1/L^3$, where L is the representative linear supercell size; the defect formation energy has been found to converge rapidly with supercell size when only these interaction energy terms have to be cared of (this is the case for neutral defects) [159].

Charged defects (charge $q = \pm 1, \dots$) may also be stabilized in semiconductors and insulators, where the neutral defect can donate or trap charge carriers (electron or holes) to or from the host crystal; as a consequence, the expression for the formation energy features an additional term, taking into account possibility of charge carriers exchange with an electron (or hole) reservoir of chemical potential μ_e [329],

$$E^\mathrm{f}_{\mathrm{D},q} = E_{\mathrm{D},q} - E_\mathrm{H} - \sum_i n_i \mu_i + q\mu_\mathrm{e} + E^\mathrm{corr}_{\mathrm{D},q}, \tag{4.2}$$

where the term $E^\mathrm{corr}_{\mathrm{D},q}$ corrects the formation energy for the electrostatic interaction between image charged defects, and μ_e is referred to the top of the valence band (VB) of the bulk supercell. Thus $E^\mathrm{f}_{\mathrm{D},q}$ depends on the asymptotic value of the average electrostatic potential in the bulk crystal. Due to the long-range nature of the Coulomb potential, such average potential cannot be uniquely defined for an infinite system with no surface termination [153]. Hence, the average electrostatic potential of the defective and the bulk supercells has to be aligned in order to define an absolute energy scale to which the chemical potential μ_e may be referred. In principle, the average potential should converge to its

bulk value if the supercell is large enough to represent a large enough portion of the bulk crystal; while quite rapid for neutral defects, such convergence is much slower for charged defects, due to the asymptotic q/r decay of the electrostatic potential associated with the defect charge q [159]. In practice, calculations for charged defects are performed by adding an homogeneous charge background compensating the defect charge, hence making the cell neutral; the electrostatic energy, which is diverging for a system with a net charge [182, 194], is thus made finite through charge compensation. Correspondingly, the average electrostatic potential in the bulk crystal is set to zero.

One possible approach to perform electrostatic potential alignment of pristine and defective supercells, is to choose atomic core levels (e.g., the $1s$ state of Si atom in the silicon crystal) as the energy reference for the computed KS eigenvalues (and, thus, the VB edge). This method has also proved capable to yield accurate band offsets at semiconductor interfaces [251].

4.2.1 Correction Term for Electrostatic Interactions: the Makov-Payne Scheme

The electrostatic correction term $E^{\text{corr}}_{\text{D},q}$ in Eq. (4.2) may be in principle evaluated by performing separate calculations on increasingly large supercells, and by extrapolating $E^{\text{f}}_{\text{D},q}$ to its value for a supercell of infinite size. In practice, besides this procedure being computationally cumbersome, one should be aware of how $E^{\text{corr}}_{\text{D},q}$ asymptotically decays as a function of the supercell size L, in order to properly perform the extrapolation.

Various models for $E^{\text{corr}}_{\text{D},q}$ have been proposed in the literature; a comprehensive analysis of their features are presented in Ref. [159]. Here the model due to Makov and Payne [194] is reviewed. Consider a model localized charge distribution $n_{\text{D},q}(\mathbf{r})$ (integrating to the total defect charge, q) periodically repeated and embedded in a neutralizing homogeneous charge background of density $n_0 = -q/\Omega$, where Ω is the supercell volume. The electrostatic energy of a lattice of point charges q (separated by distance L) is given by the Madelung energy,

$$E_{\text{M}} = \frac{\alpha_{\text{M}} q^2}{2L}, \tag{4.3}$$

where α_{M} is the Madelung constant, which depends on the Bravais lattice, and can be evaluated using the Ewald method [116]. The electrostatic interaction of the model defect charge distribution $n_{\text{D},q}(\mathbf{r})$

with the homogeneous background charge n_0, for a cubic cell of side L, is given by

$$E_{\mathrm{D,bg}} = -\frac{2\pi q Q_r}{3L^3}, \tag{4.4}$$

where Q_r is the quadrupole moment of the localized charge distribution $n_{\mathrm{D},q}(\mathbf{r})$,

$$Q_r = \int_\Omega d^3r\, r^2 n_{\mathrm{D},q}(\mathbf{r}). \tag{4.5}$$

Taking contributions (4.3) and (4.4) into account, the formation energy (4.2) can be corrected accordingly, resulting in the Makov-Payne (MP) electrostatic correction energy [194],

$$E_{\mathrm{D},q}^{\mathrm{MP}} = E_{\mathrm{D},q}^{\mathrm{MP1}} + E_{\mathrm{D},q}^{\mathrm{MP2}} = \frac{\alpha_{\mathrm{M}} q^2}{2\epsilon L} - \frac{2\pi q Q_r}{3\epsilon L^3}, \tag{4.6}$$

where the screening of the defect charge by the host crystal has been effectively taken into account through scaling of the electrostatic energy terms by the macroscopic dielectric constant of the embedding medium, ϵ. Such screening effects may correspond to (i) rearrangement of the electronic density in response to presence of the defect charge, or (ii) rearrangement of the electronic and ionic charge. In case (i) the defect formation energy is evaluated fixing the atomic structure at the unrelaxed (e.g., nondefective) one, and ϵ should be identified with the high-frequency (ion-clamped) dielectric constant ϵ_∞; in case (ii) atoms around the defect are free to relax, and the low-frequency (relaxed-ion) dielectric constant ϵ_0 should be used instead.

Lany and Zunger proposed to calculate the quadrupole term in Eq. (4.6) using the difference in the total charge density of the charged and the neutral defective systems [178]. This has been found to be predominantly contributed by a delocalized screening charge of density $n_{\mathrm{s}} = (1 - 1/\epsilon)q/\Omega$, produced by rearrangement of the electronic density in response to presence of the defect. By performing the spatial integration in Eq. (4.5), one obtains (for a cubic cell)

$$E_{\mathrm{D},q}^{\mathrm{MP2}} = -\frac{\pi q^2}{6\epsilon L}\left(1 - \frac{1}{\epsilon}\right), \tag{4.7}$$

which is observed to decay as $1/L$, in the same way as the MP1 term; using the definition in Eq. (4.6), this term can be rewritten

$$E_{\mathrm{D},q}^{\mathrm{MP2}} = -E_{\mathrm{D},q}^{\mathrm{MP1}} c_{\mathrm{sh}}\left(1 - \frac{1}{\epsilon}\right), \tag{4.8}$$

where c_{sh} is a geometric prefactor depending on the supercell shape; for a cubic cell, $c_{\text{sh}} = \pi/(3\alpha_{\text{M}}) \approx 0.369$ [179]. The MP correction term, with the quadrupole contribution evaluated according to Lany and Zunger, reads

$$E^{\text{MP}}_{\text{D},q} = \left[1 - c_{\text{sh}}\left(1 - \frac{1}{\epsilon}\right)\right] \frac{\alpha_{\text{M}} q^2}{2\epsilon L}. \tag{4.9}$$

Hence the term proportional to the Madelung energy is effectively scaled (at order $1/\epsilon$) by a factor $(1+f)$ which, under quite general assumptions and for different supercell shapes, has been found to be ≈ 0.65 ($f \approx -0.35$) [178]. It has also been pointed out that electrostatic corrections are to be applied in conjunction with potential alignment of the bulk and defective systems, in order to obtain properly corrected defect formation energies [159].

Shift of the Kohn-Sham Potential

The correction to supercell energy discussed above determines a corresponding shift in the KS potential [159, 296]. This is readily obtained from the definition of KS potential,

$$v_{\text{s}}[n](\mathbf{r}) = \frac{\delta(E[n] - T[n])}{\delta n(\mathbf{r})} + \frac{\delta E^{\text{corr}}_{\text{D},q}}{\delta n(\mathbf{r})}. \tag{4.10}$$

The charge density distribution associated with the defect can be written

$$n_{\text{D},q}(\mathbf{r}) = q|\phi_{\text{D}}(\mathbf{r})|^2 = \sum_I z_I\, \delta(\mathbf{r} - \mathbf{R}_I) - n(\mathbf{r}), \tag{4.11}$$

where ϕ_{D} is the KS defect wavefunction, z_I and $\mathbf{R}_I$ are the nuclei charge and position within the cell. The derivative of $E^{\text{corr}}_{\text{D},q}$ is evaluated by using the chain rule; if $E^{\text{corr}}_{\text{D},q}$ is calculated using the MP formula (4.6), the shift to be applied to the KS potential reads [45, 159]

$$v^{\text{MP}}_{\text{s},q}(\mathbf{r}) = \frac{\delta E^{\text{MP}}_{\text{D},q}}{\delta n} = -\frac{\alpha_{\text{M}} q}{\epsilon L} + \frac{2\pi Q_r}{3\epsilon L^3} + \frac{2\pi}{3\epsilon L^3} q r^2. \tag{4.12}$$

4.3 Optical and Electrical Properties

The presence of even small concentrations of point defects in insulators and semiconductors may substantially modify their electrical and optical properties. This fact is commonly exploited in applications, since it allows one to engineer the electronic properties of technologically

important materials. Various spectroscopic techniques can be used to identify and characterize defects in solids. From the theory side, it is necessary to conceive a computational approach by means of which quantities directly comparable with experiments can be calculated.

The presence of point defects in insulating crystals often results in the formation of defect electronic states in the forbidden energy gap of the material. Such defect states may be populated by excess electrons or holes, according to the donor or acceptor nature of the defect.

Excess charges may be trapped at the defect site, or delocalized over the whole crystal, as a result of thermal ionization of the defect at room temperature. This causes the observed n-type (electron excitation into the conduction band) or p-type (hole excitation into the VB) conductivity of some defective semiconductors. The number of charge carriers thermally transferred from the defect to the host crystal depends on the relative stability of different defect charge states, and can be analyzed on the basis of the corresponding thermodynamic (or adiabatic) transition levels (see Section 4.3.1). Shallow levels, i.e. defect levels that largely preserve the nature of electronic states of bulk band edges, are often involved in thermally assisted ionization processes. Shallow defect levels can be described within the hydrogenic effective-mass theory of Kohn and Luttinger [156]. Conversely, deep levels are usually found stable in their neutral charge state, for which the associated excess charge remains trapped at the defect site or in its vicinity, and hence does not contribute to room-temperature conductivity. Their theoretical description is more involved and cannot be pursued on the basis of perturbative approaches, like for shallow levels [16].

Defects for which the neutral charge state is stable at room temperature may also be ionized by transfer of the excess charge into the host crystal, following absorption of some external (e.g., optical) perturbation. This gives rise to electron or hole excitation from the defect state into the conduction or valence band, depending on the defect donor or acceptor nature. A number of spectroscopic techniques can characterize defect-related optical transition mechanisms in defective semiconductors and insulators. From the theoretical point of view, optical excitation and emission processes are to be modeled on the basis of optical transition levels (see Section 4.3.1). In principle, computed DFT-KS defect levels do not provide any insight into the nature of electronic excitations in defective materials. Since LDA and GGA severely underestimate the bulk band gap of insulators, positioning of defect states may turn out to be even qualitatively wrong, predicting

them to be resonant with the bulk band edges [178]. Usage of hybrid functionals or many-body GW methods leads to improvements, as a consequence of the bulk band structure being correctly described, both in terms of band gap size and of band edge positions with respect to the defect levels [8,45,64]. Thermodynamic and optical excitation/emission energies can then be determined by comparing defect formation energies for charge states differing by one unit, following charge carriers exchange between the defect and the host crystal. Such charge transition levels [178, 267, 309] are apt to interpretation of optical spectroscopy, electrical and electron paramagnetic resonance (EPR) experiments.

4.3.1 Charge Transition Levels

Since a point defect can be viewed as an atomic-like system embedded in the crystalline medium, ionization potentials and charge affinities can be determined based on an approach similar to the delta-self-consistent-field (ΔSCF) method [147]. Within ΔSCF, excitation energies in finite systems (atoms or molecules) can be computed based on ground-state energy differences between the N-electron and the $(N \pm 1)$-electron systems (see also Section 2.3).

Considering a variation of the defect charge state from q' to q, and not considering structural relaxation at the new charge state (i.e., retaining the minimum energy configuration of the system at the initial charge state q'), the optical transition level $\mu^{\rm opt}(q/q')$ is defined as the electron chemical potential at which $E^{\rm f}_{{\rm D},q'}$ equals $E^{\rm f}_{{\rm D},q}$,

$$\mu_{\rm e} = \mu^{\rm opt}(q/q') \quad \Longrightarrow \quad E^{\rm f}_{{\rm D},q'} = E^{\rm f}_{{\rm D},q}. \tag{4.13}$$

Since the external conditions under which the defect is formed are assumed to be fixed, the term proportional to the chemical potentials $\{\mu_i\}$ cancels, and one is left with the final expression for the optical transition level,

$$\mu^{\rm opt}(q/q') = \frac{E'_{{\rm D},q'} - E'_{{\rm D},q}}{q - q'}, \tag{4.14}$$

whose value is referred to the top of the bulk VB. Equation (4.14) is expressed in terms of the ground-state energy of the defective supercells for different charge states q' and q; here $E'_{{\rm D},q}$ is the total energy corrected for the electrostatic interaction between image defects which is finite when $q \neq 0$. One can also recast the energy difference appearing in Eq. (4.14) in terms of KS defect eigenvalues. If the defect charge state varies by one unit, e.g. $q = q' + 1$ due to excitation of an excess electron

from the donor defect level into the bulk crystal,[2] using Janak's theorem [143] one has

$$E'_{\mathrm{D},q'} - E'_{\mathrm{D},q} = \int_0^1 d\eta\, \varepsilon_{N+1}(N+\eta; q'+1-\eta), \tag{4.15}$$

where the occupation of the $(N+1)$-th KS spin-orbital is varied from $\eta = 0$ (ε_{N+1} corresponding to the lowest unoccupied spin-orbital of the N-electron system, defect charge state $q = q'+1$) to $\eta = 1$ (ε_{N+1} corresponding to the highest occupied spin-orbital of the $(N+1)$-electron system, defect charge state q'). In the spirit of Slater's transition-state theory [285], the total energy difference (4.15) can be approximated as [7, 97]

$$E_{\mathrm{D},q'} - E_{\mathrm{D},q} \approx \frac{1}{2}\left[\varepsilon_{N+1}(N+1; q') + \varepsilon_{N+1}(N; q = q'+1)\right]. \tag{4.16}$$

Inserted in Eq. (4.14), this approximation allows one to compute the optical transition level (q/q') by keeping track of the shift in the KS frontier eigenvalue upon removal of an electron from the corresponding orbital, not taking into account subsequent structural relaxations. This approximate approach has been demonstrated to provide transition energies in good agreement with experiments [97], as well as with transition levels computed on the basis of a direct evaluation of total energy differences [7]. Analogously, optical transition levels corresponding to trapping of an electron by a defect with initial charge state q can be computed; the charge trapping induces a change in the charge state from q to $q' = q-1$, and

$$E'_{\mathrm{D},q'} - E'_{\mathrm{D},q} \approx \frac{1}{2}\left[\varepsilon_{N+1}(N; q) + \varepsilon_{N+1}(N+1; q' = q-1)\right] \tag{4.17}$$

Thermodynamic transition levels $\mu^{\mathrm{therm}}(q/q')$ are to be related to adiabatic charge excitations or decays $q_{\mathrm{i}} \to q_{\mathrm{f}}$, in which the transition develops over timescales allowing for structural relaxation around the defect in its final charge state, q_{f}. The thermodynamic level is obtained from the corresponding optical level, by adding or subtracting the relaxation energy

$$E_{\mathrm{rel}} = E_{\mathrm{D},q_{\mathrm{f}}}\big|_{q_{\mathrm{i}}\ \mathrm{geom}} - E_{\mathrm{D},q_{\mathrm{f}}}\big|_{q_{\mathrm{f}}\ \mathrm{geom}}, \tag{4.18}$$

[2] To fix the ideas, consider for example the optical ionization of an initially neutral donor defect ($q' = 0$): its final charge state will be $q = +1$.

according to whether $q \equiv q_\mathrm{f} = q_\mathrm{i} + 1 \equiv q' + 1$ (electron excitation), or $q' \equiv q_\mathrm{f} = q_\mathrm{i} - 1 \equiv q' + 1$ (electron trapping) [97],

$$\mu^{\mathrm{therm}}(q/q') = \mu^{\mathrm{opt}}(q/q') \pm E_{\mathrm{rel}}. \tag{4.19}$$

The formalism presented above for electron transitions can equally well be applied to study of optical or thermodynamic transitions of excess holes from an acceptor-like point defect level into the host VB.

In order to correctly evaluate the KS eigenvalues to be used in Eq. (4.16), one has to take into account the finite-size issues already discussed for calculation of defect formation energies. In fact, the KS eigenvalues obtained from the two bulk calculations corresponding to different defect charge states q and q' have to be referenced to some common energy scale, which may be identified with some suitably chosen atomic core level of species in the bulk crystal. Moreover, for charged defects, the shift in the KS potential due to electrostatic interaction between images (computed within the MP correction scheme, Eq. (4.12)) induces a corresponding shift in the KS defect eigenvalue [45, 159],

$$\begin{aligned} \varepsilon^{\mathrm{MP}}_{\mathrm{D},q} &= -\frac{\alpha_\mathrm{M} q}{\epsilon L} + \frac{2\pi}{3\epsilon L^3}\left[\int d^3r\, r^2 n_{\mathrm{D},q}(\mathbf{r}) + q\,\langle\phi_\mathrm{D}|\,\hat{r^2}\,|\phi_\mathrm{D}\rangle\right] \\ &= -\frac{\alpha_\mathrm{M} q}{\epsilon L} + 2\times\frac{2\pi Q_r}{3\epsilon L^3}. \end{aligned} \tag{4.20}$$

Hence, the electrostatic correction for the KS defect level, which is connected to the total energy correction through $\varepsilon^{\mathrm{MP}}_{\mathrm{D},q} = -(2/q)E^{\mathrm{MP}}_{\mathrm{D},q}$, can be approximated using the simplified formula (4.9), finally yielding

$$\varepsilon^{\mathrm{MP}}_{\mathrm{D},q} = -2(1+f)\left(\frac{\alpha_M}{2L}\right)\frac{q}{\epsilon_\infty}. \tag{4.21}$$

Notice that the high-frequency dielectric constant ϵ_∞ has been used in Eq. (4.21). In fact, when optical transitions are computed, the screening of the final defect charge is due to polarization of the electron gas only, while ions are assumed to be frozen in their initial equilibrium configuration. This justifies usage of the electronic dielectric constant instead of the total (combined electronic and ionic) one. No electrostatic corrections should be applied when computing the relaxation energy (4.18), since total energies are evaluated for the same charge state q_f, and thus errors in total energy due to finite-size effects are expected to cancel against each other [62].

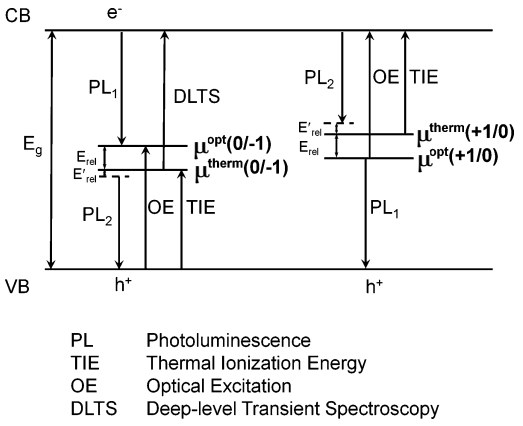

Figure 4.1: *Schematic representation of electronic transitions involving a neutral defect ($q = 0$) and a positively ($q = +1$) and negatively ($q = -1$) charged one. Representative optical and thermodynamic charge transition levels are reported. Spectroscopic techniques suitable for probing each transition are indicated. Reprinted and adapted with permission from [97]. Copyright 2010, AIP Publishing LLC.*

Connection to Experiments

Transition levels can be related to measured excitation and emission energies probed through spectroscopies. Figure 4.1 schematically shows possible transition paths involving defect levels which can occur in defective semiconductors and insulators. Theoretical studies can play a fundamental role in elucidating the complex optical transition mechanisms observed in photoluminiscence and absorption spectroscopy experiments. Photoluminiscence can be produced by decay of one electron from the conduction band (CB) into the defect state (emission energy $\varepsilon_{\mathrm{CB}} - \mu^{\mathrm{opt}}(0/-1)$), or from the defect level into the VB (emission energy $\mu^{\mathrm{opt}}(+1/0) - \varepsilon_{\mathrm{VB}}$). More complex photoluminiscence paths can also occur (labeled PL_2 in Figure 4.1): for example, an electron can decay nonradiatively from the CB into the defect level, forming a negatively charged defect which reaches its equilibrium geometry; a subsequent radiative transition from the defect level into the VB gives rise to the observed photoluminiscence.

Optical transitions involving defect levels can also be observed in optical absorption spectra. For example, an initially neutral defect can be ionized by absorption of a photon, as an excess electron is promoted into the CB (excitation energy $\varepsilon_{\mathrm{CB}} - \mu^{\mathrm{opt}}(+1/0)$).

Charge transitions involving thermodynamic levels result from

thermal ionization of the defect, and the magnitude of the transition energy (or, equivalently, the distance of the thermodynamic level from the bulk band edges) is indicative of the stability of a given defect charge state at room-temperature conditions. Deep thermodynamic levels can also be probed using spectroscopic techniques that allow for relaxation in the final charge configuration, such as deep-level transient spectroscopy.

More on excitation mechanisms occurring in defective materials will be discussed in connection to investigation of defective oxides in Chapters 6 and 7.

Part II

Investigation of Bulk and Defective Oxides

CHAPTER 5

Bulk Oxides: Band Gaps and Phase Stabilities

In the present chapter a systematic investigation of band gaps and ground-state properties is presented for the following set of wide-gap metal oxides in several bulk polymorphic structures: zinc oxide (ZnO), titanium dioxide (TiO_2), zirconium dioxide or zirconia (ZrO_2), and tungsten trioxide (WO_3). Results are also reported for MgO, which is studied as a simple test system, although being an insulator with a gap as large as 8 eV.

DFT calculations are performed using several local/semilocal and hybrid xc functionals, and results obtained from PW-PP and LCAO-based calculations are critically compared. GW band gaps for selected polymorphs of the above set of oxides are also computed to provide benchmark of the different analyzed DFT methods, along with comparison with available experimental results. The relative stability of several polymorphic structures at zero temperature is investigated at different levels of theory. Finally, chemical reduction of TiO_2 to Ti_2O_3 (bulk phases), involving a change in the oxide stoichiometry, is analyzed. For the last two points, the accuracy of DFT methods in describing the energetics of phase or stoichiometry transformations in

bulk oxide materials is investigated. For all of the above properties, a systematic comparison is reported of at least one approach belonging to the classes of local/semilocal (LDA, GGA), standard hybrid (PBE0, HSE06, and B3LYP), and dielectric-dependent hybrid functionals.

5.1 Band Gap: Local/Semilocal, Hybrid Functionals and GW

In this section, the following materials are considered: cubic (c) rocksalt MgO, wurtzite (wz) ZnO, anatase (a) TiO_2, tetragonal (t) ZrO_2, and WO_3 in the room-temperature γ-monoclinic (γ-m) structure. Band gaps are computed within DFT (within both PW-PP and LCAO schemes) using several xc functionals, as well as within GW; results are reported and compared in Table 5.1, along with selected experimental results. For a more meaningful comparison of the different computational approaches, and due to the exceedingly large cost of geometry optimizations using hybrid functionals with the PW code QUANTUM ESPRESSO, calculations in this section are performed at the experimental geometry.

It should be first emphasized that comparison between theoretical calculations and experimental data is typically far from straightforward when it comes to spectroscopic properties of semiconductors and insulators. From the experimental point of view, different techniques give access to different physical observables. In (direct and inverse) photoemission experiments, electrons are removed from or added to the solid (charged excitations), probing hence the energy difference between the N- and the $N \pm 1$-electron systems. From the theoretical side, such total energy differences are described by the poles of the single-particle Green's function G, i.e. the QP energies (see Section 3.1.1). Thus, the measured QP energies or their difference (defining the fundamental gap) can be put in direct comparison with values obtained by solving the QP eigenvalue problem [141], after devising an approximation to the self-energy operator Σ, for example within the GW scheme, or by approximating it with a DFT xc potential (in which case the QP energies reduce to the KS eigenvalues). In contrast, optical measurements involve neutral excitations, which can be seen as creation of electron-hole pairs. The binding energy of electron-hole pairs is hence inherently included. The theoretical description of neutral excitations in principle requires working with the two-particle (electron-hole) Green's function, and such an analysis lies beyond the scope of the present work. Nonetheless, by

knowing or guessing the exciton binding energy, optical measurements can be useful for extrapolating QP gaps when direct and inverse photoemission measurements are not available for the material at hand. This situation is not uncommon, as discussed in the following for some of the materials investigated here.

As a second source of possible disagreement with experiment, DFT calculations performed at frozen ion positions cannot not take into account electron-phonon interaction, which generally leads to renormalization of the band gap, reducing it with respect to the zero-temperature QP gap [37, 110]. Given the above considerations, it is worthwhile to give an overview of the experimental scenario (selected experimental results are reported in Table 5.1).

5.1.1 Experimental Data

The fundamental gap for MgO can be extrapolated from thermoreflectance measurements [323] following estimation of the exciton binding energy.

In the case of ZnO, different optical experiments [185, 258] agree on the value of the reported QP gap, which has been obtained from optical experiments by detailed analysis of the excitonic levels.

For anatase TiO_2, the situation is more complicated, since, to the author's knowledge, no experimental data are available to date for the photoemission gap. However, analysis of the temperature dependence of the absorption coefficient allows one to estimate the indirect optical gap to be 3.42 eV [295]. For rutile TiO_2, measurements of both the QP and the optical band gap are available. An optical band gap of ~ 3.0 eV has been reported in several investigations [238, 295], while combined photoemission and inverse photoemission experiments yielded for the direct fundamental band gap the values of 3.3 ± 0.5 eV [297] and 3.6 ± 0.2 eV [256]. Assuming a similar relationship between fundamental and optical gap holds for anatase, one could tentatively estimate the anatase QP gap to be in the range 3.7 – 4.0 eV, as also suggested in Refs. [149, 176].

The only available data for the pure phases of ZrO_2 are from optical absorption measurements [92], from which a direct optical gap of 5.78 eV for the tetragonal phase was inferred. Photoemission studies yielded smaller measured gaps of 5.68 eV [266] and 5.5 eV [22]; however these values cannot be attributed to any of the three ZrO_2 phases, as the probed samples were amorphous or polyscrystalline.

For WO_3, several ultraviolet direct/inverse photoemission studies

report a band gap in the range 3.28 – 3.39 eV [165, 208, 320]. Instead, absorption measurements typically show large data dispersion, with an optical band gap spanning the range 2.6 – 3.2 eV [32, 134, 155].

5.1.2 DFT LDA, GGA and Standard Hybrids: Comparison Between PW and LCAO

Results of the present calculations confirm that, as expected, hybrid functionals provide substantial improvement over the well known LDA/GGA gap underestimation. In particular, band gaps computed at the HSE06 level are in good agreement with experiment for a-TiO_2, t-ZrO_2 and γ-m-WO_3, as also confirmed by previous calculations on titanium [62, 144, 176] and tungsten [315] oxides. For the same materials, PBE0 overestimates the band gap with respect to experiment, while B3LYP generally gives better results; this findings are in agreement with several previous investigations [21, 36, 96, 315, 330]. Quantitative disagreement with the literature, when present, may be attributed to different choices of geometry (DFT optimized vs. experimental, the latter being considered in the present part of the study). Independently of the adopted theoretical treatment, an indirect gap is obtained for a-TiO_2 and t-ZrO_2, and a direct one for the other materials. Both values are reported in Table 5.1.

In the case of MgO and ZnO, it is observed that hybrid functionals perform differently from what previously outlined. MgO is an insulator with a wide band gap of nearly 8 eV; its computed gap, even using hybrid functionals, is obtained smaller than the experimental value. In the case of ZnO, due to the strong gap underestimation at the LDA/GGA level, and contrary to what it is found for the other semiconductors, PBE0 seems to perform well, similarly to B3LYP, while HSE06 is not sufficient to correct the PBE underestimation. Recent studies using hybrid functionals [198, 283, 303, 314, 326] confirm the observed trends for both MgO and ZnO.

Finally, a comment is due on the comparison between results obtained within PW-PP and LCAO schemes. From Table 5.1, it is found that discrepancies in computed band gaps range from ~ 0.05 to ~ 0.30 eV, most of them being within 0.20 eV, depending on the material and the functional used. The closest agreement is observed at the LDA/GGA level, while with hybrid functionals discrepancies become more substantial. Different results are the consequence of different kinds of approximations characterizing the two computational schemes. Generation of reliable PPs on the one hand and construction

of accurate Gaussian-type basis sets for solid-state calculations on the other, are certainly two of the main critical points in this respect. For example, in the case of ZnO, using a Zn PP with only $3d$ and $4s$ electrons treated in the valence, and adopting an inadequate basis set for Zn in LCAO calculations, result in computed gaps differing by as much as $\sim$ 0.5 eV within LDA/GGA and $\sim$ 1 eV at the hybrid functional level. Adoption of a better PP and basis set indeed partially reduce the disagreement. Inclusion of $3s$ and $3p$ semicore electrons in Zn is also found to be crucial for obtaining reliable G_0W_0 corrections of the PBE gap,[1] similarly to what has been reported by Gori *et al.* [112]. In conclusion given the above discussion, the agreement between the computed PW-PP and LCAO band gaps should be deemed satisfactory.

5.1.3 Many-Body Perturbation Theory: G_0W_0

Table 5.1 also reports G_0W_0 band gaps computed on top of the DFT-PBE band structure. A generally good agreement is obtained with experiment for all the materials. The results presented here are in line with previous investigations, although numerical values may differ as a consequence of the different computational setup adopted, e.g. the starting DFT approximation to QP energies and wavefunctions, and the treatment of the frequency-dependent dielectric function [149].

For MgO, the computed gap of 7.88 eV is very close to experiment and to the results of the G_0W_0 investigation of Fuchs *et al.* adopting a HSE03 starting point [95]. Calculations starting from DFT-PBE using the PAW method gave smaller QP gaps of 7.25 eV [281] and 7.41 eV [44].

For TiO_2, the obtained indirect QP gap of 3.73 eV is identical to the full-frequency result of Landmann *et al.* [176], and in close agreement with the ones reported in other investigations [47, 149, 239, 300].

Quasiparticle band structure calculations of ZrO_2 are scarce. The only G_0W_0 studies the author is aware of reported a fundamental band gap of 6.40 eV [162] and 5.56 eV [146] for tetragonal zirconia. The result reported here of 6.06 eV for the direct gap is in between these two values, and compatible with the measured optical gap of 5.87 eV.

As for γ-m-WO_3, recent studies of Galli and coworkers yield QP gaps of 3.26 eV [254] and 3.30 eV [251], which are very close to the result reported in Table 5.1, and in excellent agreement with the measured photoemission gap.

[1] The computed G_0W_0 gap without explicitly treating semicore $3s$ and $3p$ electrons is $\sim$ 1 eV.

Table 5.1: *Fundamental band gaps (eV) for the various materials computed within local/semilocal, standard hybrid, and G_0W_0 schemes, comparing PW-PP and LCAO schemes. All calculations are performed at the experimental geometry (see Table 5.7). When the fundamental gap is indirect, direct band gaps are reported in parenthesis. Estimated errors are 20 meV and 50 meV for the DFT and GW results, respectively.*

	Type	Method	LDA	PBE	PBE0	HSE06	B3LYP	G_0W_0@PBE	Expt.
MgO	rs	PW	4.63	4.79	7.18	6.44	7.04	7.88	7.83[a]
		LCAO	4.86	4.93	7.35	6.67	7.07		
ZnO	wz	PW	0.78	0.81	3.18	2.46	2.90	3.06	3.44[b]
		LCAO	0.92	1.02	3.35	2.75	3.05		
TiO_2	a	PW	2.03 (2.16)	2.13 (2.26)	4.40 (4.47)	3.67 (3.74)	3.94 (4.03)	3.73 (3.81)	3.42[c]
		LCAO	2.09 (2.18)	2.18 (2.28)	4.23 (4.33)	3.59 (3.70)	3.79 (3.89)		
ZrO_2	t	PW	4.01 (4.18)	4.12 (4.29)	6.52 (6.69)	5.78 (5.95)	6.09 (6.25)	5.87 (6.06)	5.78[d]
		LCAO	3.94 (4.09)	4.00 (4.16)	6.33 (6.45)	5.61 (5.75)	5.90 (6.02)		
WO_3	γ-m	PW	1.84	1.94	3.90	3.16	3.53	3.34	3.38[e]
		LCAO	1.86	1.91	3.74	3.09	3.39		

[a] Thermoreflectance at 85 K, Ref. [323].
[b] Transmission spectroscopy at 4.2 K, Ref. [185]; photoluminiscence at 2 K, Ref. [258].
[c] Transmission spectroscopy, Ref. [295].
[d] Reflectance spectroscopy (VUV) in Y-doped ZrO_2, Ref. [92].
[e] Photoemission (UPS-IPES), Ref. [208].

Table 5.2: *Band gap energy (eV) computed using dielectric-dependent hybrid functionals defined through the exchange fractions reported in Table 5.3. Comparison with results obtained within PBE0 and G_0W_0 is provided. Only the direct gap is reported for ZrO_2, since the measured gap is extrapolated via a model for a direct gap. MAE and MARE are the mean absolute and mean absolute relative error, respectively, with respect to the computed G_0W_0 band gap. All DFT calculations are performed within the LCAO scheme at the experimental geometry. Estimated errors are 20 meV and 50 meV for the DFT and GW results, respectively.*

	Type	PBE0α_{PBE}	PBE0α_{PBE0}	PBE0	G_0W_0@PBE	Expt.
MgO	rs	8.06	8.33	7.38	7.88	7.83
ZnO	wz	3.18	3.94	3.35	3.06	3.44
TiO_2	a	3.38	3.72	4.24	3.73	3.42
ZrO_2	t	5.81	6.12	6.45	6.06	5.78
WO_3	γ-m	3.23	3.50	3.74	3.34	3.38
MAE (eV)		0.20	0.31	0.42		
MARE (%)		4.6	8.1	9.6		

Finally, ZnO stands out as a particularly critical case, for which considerable disagreement is found in the theoretical literature concerning its QP properties. Several studies provide QP gaps ranging from $\sim$ 2.1 eV [95, 281] to 3.4 eV, obtained by Louie and coworkers [276] within G_0W_0@LDA. In the latter study, it was suggested that smaller computed gaps may be the result of a false convergence behavior of the dielectric function and self-energy with respect to the number of included empty states. Stankovski *et al.* [290] studied the effect of different plasmon-pole models on the band gap, while Friedrich *et al.* [94] carried out G_0W_0 on top of all-electron DFT calculations to investigate the effect of the PP approximation. At present there is no general consensus on which would be the most correct approach to describe the QP band structure of this material. Although the above convergence issues have been carefully considered here,[2] G_0W_0 calculations gave a gap of 3.06 eV, which is still significantly smaller than the experimental one; nonetheless, it agrees with the values reported in several previous investigations using a G_0W_0@LDA/GGA approach.

[2] See Computational Details.

Table 5.3: *Electronic dielectric constant ϵ_∞ (estimated error within 0.02) and corresponding exchange fraction α (%) evaluated at the PBE and PBE0 levels. Calculations are performed within the LCAO scheme at the experimental geometry. Measured dielectric constants are reported for comparison.*

		PBE		PBE0		Expt.[a]
	Type	ϵ_∞	α	ϵ_∞	α	ϵ_∞
MgO	rs	3.10	32.2	2.87	34.9	2.96
ZnO	wz	4.31	23.2	3.24	30.8	3.74
TiO_2	a	6.52	15.3	5.16	19.4	5.62
ZrO_2	t	5.45	18.3	4.62	21.6	4.9
WO_3	γ-m	5.43	18.4	4.56	21.9	4.81

[a] Measured static electronic dielectric constants are taken from the following references: MgO, Ref. [187]; ZnO, Ref. [12]; TiO_2, Ref. [322]; ZrO_2, Ref. [92]; WO_3, Ref. [140].

5.1.4 DFT Dielectric-Dependent Hybrid Functionals

The performance of different kinds of dielectric-dependent hybrid functionals is now analyzed. The EXX fraction is defined as $\alpha = 1/\epsilon_\infty$; the macroscopic dielectric tensor is computed within both PBE and PBE0, without iterating to self-consistency (first step of the sc-PBE0α_{ϵ_∞} method, see Section 3.3, except for usage of the experimental geometry); for materials with a noncubic primitive cell, ϵ_∞ is defined as the average of the diagonal elements of the diagonalized dielectric tensor written in real space.

Table 5.3 reports the computed dielectric constants along with the corresponding exchange fractions α_{PBE} and α_{PBE0}, defining the modified PBE0 functionals, PBE0α_{PBE} and PBE0α_{PBE0}. Notice that, as expected, PBE and PBE0 systematically overestimates and underestimates, respectively, the dielectric constant, as a consequence of the opposite tendency in estimation of the band gap.

Table 5.2 shows that, even without self-consistent treatment, the dielectric-dependent hybrid approach remarkably improves the computed gap over the PBE0 functional. Agreement with experiment and G_0W_0 calculations is overall satisfactory with both modified hybrids, with the exception of ZnO, for which PBE0 already performs well. Since reliable experimental photoemission gaps are not available for all the materials, mean errors in Table 5.2 are defined with respect to the computed G_0W_0 gaps. Error analysis shows that the PBE0α_{PBE}

functional achieves the closest agreement with G_0W_0 for the class of materials under investigation. It is thus taken as the starting point for evaluating α in most of the calculations presented in this chapter.

5.2 Oxide Polymorphs: Band Gaps, Crystal Structures and Phase Stability

Several other structural modifications of the materials considered above are investigated in this section. This includes analysis of the their relative phase stability, electronic and dielectric properties, and equilibrium crystal structure, by using different DFT xc functionals.

Phase stability is evaluated on the basis of zero temperature DFT calculations, thus not taking into account thermal and entropic contributions. In particular, the thermal contribution mainly comes from lattice vibrations (phonons), and may be relevant when considering temperature-induced structural phase transitions. Here it is argued that an analysis of only the zero-temperature electronic contribution is still able to provide an at least qualitative picture on the stability issue. In fact, although in this case quantitative comparison with thermochemical data may be of limited significance, it can be interesting to assess whether theory is able to correctly reproduce the experimentally found phase stability sequence (at ambient pressure) for a given material.

Since typically tiny energy differences are involved in transformations between phases with similar stabilities, properly reproducing the stability order is an extremely challenging task, pushing to the limits of accuracy of DFT methods. Different functionals may predict different stability sequences, as a consequence of changes in relative total energies of the order of some meV per atom, going from one level of theory to another. Even though quantitative conclusions are hard to be drawn, it is however valuable to check whether, for example, different hybrid functionals (with different α) provide similar results in terms of phase stability.

In view of this, the performance of the PBE, PBE0, and sc-PBE0α_{ϵ_∞}(obtained by using the PBE to compute α as a first guess) functionals is tested in the following. Equilibrium geometries are also investigated, as it is expected that a good xc functional is able to yield optimized cell parameters within few percents from experiment. Moreover, accurate geometry optimizations are needed for obtaining meaningful total energies to be compared for analysis of phase stability. For this reason, in the present section all the reported quantities

Table 5.4: *Electronic dielectric constant ϵ_∞ (estimated error within 0.02) and corresponding exchange fraction $\alpha = 1/\epsilon_\infty$ (%) evaluated within PBE and dielectric-dependent PBE0 (PBE0α_{PBE} and sc-PBE0α_{ϵ_∞}) for various polymorphs of the studied materials. Calculations are performed within the LCAO scheme at the optimized geometry.*

		PBE		PBE0α_{PBE}		sc-PBE0α_{ϵ_∞}	
	Type	ϵ_∞	α	ϵ_∞	α	ϵ_∞	α
MgO	rs	3.13	31.9	2.80	35.7	2.78	36.0
ZnO	wz	4.30	23.2	3.28	30.8	3.14	31.8
	zb	4.67	21.4	3.37	29.6	3.19	31.3
TiO_2	r	7.98	12.5	6.76	14.8	6.59	15.2
	a	6.58	15.2	5.59	17.9	5.45	18.3
	b	6.95	14.4	5.99	16.7	5.87	17.0
ZrO_2	m	5.29	18.9	4.62	21.6	4.55	22.0
	t	5.56	18.0	4.89	20.5	4.82	20.8
	c	5.86	17.1	5.09	19.7	4.99	20.0
WO_3	ε-m	5.58	17.9	4.74	21.1	4.65	21.5
	tr	5.54	18.1	4.68	21.4	4.56	21.9
	γ-m	5.55	18.0	4.68	21.4	4.57	21.9
	or	5.43	18.4	4.55	22.0	4.45	22.5
	t	6.47	15.4	5.33	18.7	5.17	19.4
	c	10.09	9.90	8.27	12.1	7.98	12.5

(including dielectric constants, which in turn affect the amount of exact exchange entering the definition of the modified PBE0 functional) are computed at the minimum energy configuration of the relevant xc functional. The CRYSTAL09 code, which is very efficient for structural optimizations with hybrid functionals, is employed throughout all the calculations performed here. Due to the always considerable cost of geometry optimizations for the largest cells, the present study is limited to a single self-consistency step on α (the resulting functional is denoted sc-PBE0α_{ϵ_∞}); however, as shown in the insets of Figure 5.1, the dielectric constant converges rapidly for the materials considered here, so that even one step of self-consistency gives a functional very close to the fully converged one.

In Table 5.4 the electronic dielectric constant and the corresponding exchange fraction calculated at different levels of theory are reported for the crystallographic phases under investigation. Comparison of Tables 5.3 and 5.4 evidences the effect of geometry optimization on the computed dielectric constants (at the PBE level) for the commonly studied polymorphs.

Table 5.5: *Differences in total energy per formula unit (meV, estimated error 1 meV) with respect to a chosen crystallographic phase (redefining the zero of the energy for each functional individually), for various polymorphs of the studied materials, computed at different levels of theory.*

	Type	PBE	PBE0	PBE0α_{PBE}	sc-PBE0α_{ϵ_∞}
ZnO	wz	0	0	0	0
	zb	49	54	102	80
TiO_2	r	20	62	18	22
	a	0	0	0	0
	b	-5	32	1	2
ZrO_2	m	-20	17	33	25
	t	0	0	0	0
	c	74	45	40	36
WO_3	ε-m	5	1	-2	-8
	tr	0	0	0	0
	γ-m	3	6	2	6
	or	6	10	36	1
	t	30	70	-153	-139
	c	201	301	-403	-462

5.2.1 Polymorphs Structures and Relative Stability

Apart from wurtzite (wz), which is the most thermodynamically stable phase of ZnO at ambient conditions, the zinc-blende (zb) modification can be stabilized upon growth in specific conditions [229]. The rocksalt structure exists only at high pressures, and it is not addressed here. From Table 5.5 it is seen that the better stability of the wurtzite phase is correctly predicted at all levels of theory, with the sc-PBE0α_{ϵ_∞}confirming the result of the PBE0α_{PBE}. These findings are in agreement with previous calculations using LDA, GGA and screened-exchange hybrid functionals [305].

The three naturally occurring polymorphs of TiO_2 are rutile (r), anatase (a) and brookite (b). Their relative thermodynamic stability critically depends on crystal size. At ambient conditions, the thermodynamically most stable macrocrystalline (crystal size exceeding 35 nm) phase of TiO_2 is rutile. From calorimetric measurements of the phase transformation of anatase and brookite to rutile it is suggested that the phase stability sequence is rutile $>$ brookite $>$ anatase [148]. It is well-known that GGA is not able to reproduce the correct sequence, predicting anatase to be more stable than rutile [215]. The present results confirm such tendency and suggest that, independently of the amount of

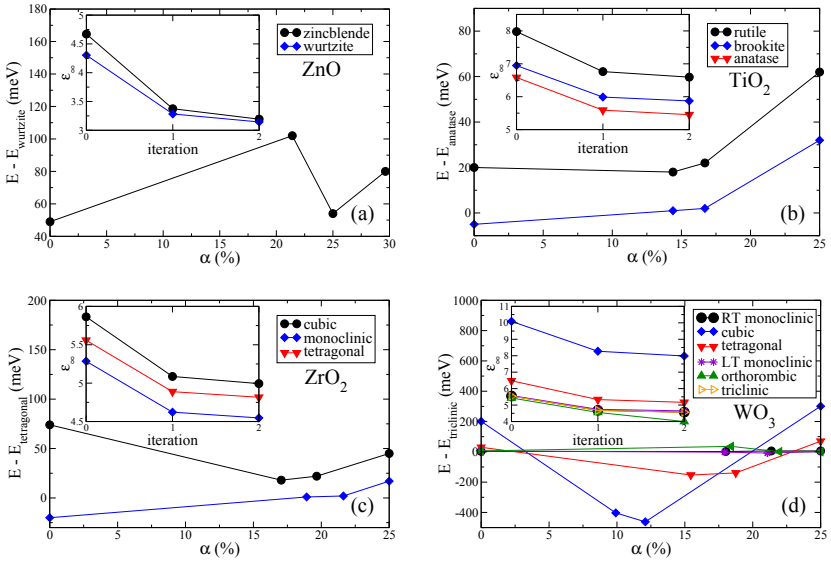

Figure 5.1: *Variation of total energy as a function of the exchange fraction α entering the definition of the dielectric-dependent hybrid functional. Total energies per formula unit are given with reference to that of a chosen polymorph for each material: (a) wurtzite for ZnO, (b) anatase for TiO_2, (c) tetragonal for ZrO_2, (d) γ-monoclinic for WO_3. In the insets convergence of the electronic dielectric constant ϵ_∞ is shown.*

EXX introduced, hybrid functionals are similarly not able to capture the correct sequence, in agreement with the previous investigation of Labat *et al.* [170]. Recent studies have demonstrated that dispersion van der Waals interactions should be taken into account in order to correctly predict phase stablity in TiO_2 [49, 212]. This lies beyond the scope of any hybrid functional approach.

Zirconia features three different structural modifications at ambient pressure; the baddeleyite structure, with a monoclinic (m) unit cell, is stable at room temperature, and transforms into the tetragonal (t) phase at 1480 K, which is stable until 2650 K when it is converted into the cubic (c) fluorite phase. In fact, the tetragonal and monoclinic phases can be viewed as distorted cubic structures [96]. From Table 5.5 it is inferred that while PBE predicts the correct stability sequence as a function of increasing temperature, PBE0 inverts the monoclinic and tetragonal phases, in fact predicting the latter to be more stable than the former, in agreement with previous studies [96] but in contrast with the experimental evidence. Varying the exchange fraction within the

dielectric-dependent hybrid scheme does not yield qualitatively different results, similarly to what is found for TiO_2.

Finally, the case of WO_3 deserves special attention, as the phase diagram at ambient pressure of this material is exceptionally rich, exhibiting five different polymorphs transforming into each other upon temperature variation. The simple cubic (c) structure, only recently successfully stabilized at ambient conditions [56], is constituted by a lattice of regular corner-sharing WO_6 octahedra, forming a strongly ionic compound. Several structures are successively obtained by decreasing the temperature from the melting point at 1700 K: tetragonal (t, stable above $740\,^{\circ}$C), orthorhombic (or, from $330\,^{\circ}$C to $740\,^{\circ}$C), room-temperature γ-monoclinic (γ-m, from $17\,^{\circ}$C to $330\,^{\circ}$C), triclinic (tr, from $-50\,^{\circ}$C to $17\,^{\circ}$C), and low-temperature ε-monoclinic (ε-m, from $-140\,^{\circ}$C to $-50\,^{\circ}$C) [315].

Octahedra become more and more distorted going from the tetragonal to the low-temperature monoclinic phase. Previous works have investigated how such distortions affect the electronic band structure, as a consequence of energy-lowering rearrangement of the WO_6 units [61, 251, 315]. A recent theoretical study suggested that tilting of the WO_6 octahedra along different crystalline axes is the main responsible for band gap opening going from the cubic and tetragonal phases (where no tilting is present) to the other, less symmetric structures [251]. This is apparent from Table 5.6, in which the reported band gap increases by as much as ~ 1 eV or more passing from the tetragonal and cubic structures to the more distorted ones. A similar relationship between structural and electronic properties was found in a recent DFT study on doped WO_3 [302], in which intercalated atoms drive the structural deformation responsible for the band gap reduction.

As expected, the electronic dielectric constant follows an opposite trend with respect to band gap, being larger in the cubic and tetragonal phases and decreasing with increasing band gap, as shown in Table 5.4. This in turn affects the corresponding exchange fractions, whose strong dependence upon the considered phase is at the origin of the wrong prediction of the stability order of WO_3 within the dielectric-dependent hybrid scheme. In fact, Table 5.5 shows that, while using the PBE0 functional the cubic and tetragonal phases are predicted to be the least stable, in agreement with the experimental order as a function of temperature, they turn out to be the most stable ones when α is obtained from the computed dielectric constant for each phase separately. This behavior is not observed in the other materials, where the electronic and

Table 5.6: *Band gap energy (eV, estimated error 20 meV) computed within PBE and dielectric-dependent PBE0 for various polymorphs of the studied materials. Fundamental and direct gaps are reported outside and inside parenthesis, respectively. Calculations are performed within the LCAO scheme at the optimized geometry.*

	Type	PBE	PBE0	PBE0α_{PBE}	sc-PBE0α_{ϵ_∞}	Expt.[a]
MgO	rs	4.63	7.38	8.24	8.67	7.83
ZnO	wz	1.07	3.50	3.32	4.07	3.44
	zb	0.91	3.29	2.93	3.76	
TiO_2	r	1.71	3.90	2.76	2.96	3.3
	a	2.14 (2.23)	4.36 (4.36)	3.42 (3.46)	3.66 (3.70)	3.42
	b	2.36	4.46	3.53	3.71	
ZrO_2	m	3.38 (3.78)	5.67 (6.10)	5.09 (5.59)	5.34 (5.77)	5.83
	t	3.80 (3.86)	6.07 (6.20)	5.43 (5.53)	5.66 (5.76)	5.78
	c	3.12 (3.63)	5.43 (6.04)	4.67 (5.25)	4.91 (5.50)	6.1
WO_3	ε-m	1.80	3.85	3.28	3.52	
	tr	1.67	3.68	3.06	3.36	
	γ-m	1.62	3.67	3.07	3.34	2.6
	or	1.39	3.34	2.81	3.09	
	t	0.53	2.23	1.53	1.76	1.8
	c	0.55 (1.58)	2.19 (3.21)	1.15 (2.17)	1.29 (2.31)	

[a] Measured photoemission gap of r-TiO_2 from Ref. [297]; direct optical gap of m-, t-, c-ZrO_2 from Ref. [92]; optical gap of t- and γ-m-WO_3 from Ref. [251]. The remaining references for experimental band gaps are those in Table 5.1.

dielectric properties vary more smoothly upon the structural transition, or, equivalently, the electronic structure correlates more weakly with lattice distortions. In those cases, the dielectric-dependent hybrid approach proved to qualitatively reproduce the results of standard PBE0. This is clearly seen in Figure 5.1, where relative total energy is plotted as a function of the exchange fraction. Noncrossing lines in the positive energy half-plane indicate that no phase stability inversion occurs going from PBE0 to PBE0α_{PBE} and sc-PBE0α_{ϵ_∞}, contrary to what happens in the case of WO_3.

5.2.2 Band Gap Dependence on Polymorphic Structure

The interesting case of WO_3 shows that the electronic structure can be significantly affected by the polymorphic structure. Having assessed the performance of the dielectric-dependent hybrid method in band gaps computation of some well-characterized oxide polymorphs (see Table 5.2), it is worthwhile to investigate the electronic structure of other

less common phases.

Table 5.6 collects computed band gaps for all the phases. Here, all calculations are performed at the optimized geometries. Due to the experimental difficulty in obtaining samples with a well-defined phase, measured gaps are not available for all the structures or are very disperse, as in the case of brookite TiO_2 [71]. Hence, comparison with previous hybrid and GW results provides an alternative source of benchmark.

In general, the PBE0$\alpha_{\rm PBE}$ confirms itself as an accurate method for computing electronic structures. Notice however that self-consistency on the dielectric constant does not always improve the computed gap, as happens in the case of MgO and ZnO. Analysis of Table 5.4 and of the insets in Figure 5.1 shows that convergence on the dielectric constant is rapidly achieved, as also reported in Ref. [283].

In the following, results from the present calculations are compared with the existing theoretical and, when available, experimental data.

For ZnO, previous hybrid calculations predict the band gap in the zinc-blende phase to be ~ 0.2 eV smaller than in wurtzite [305], which is also supported by the experimental evidence [181], confirming the PBE0$\alpha_{\rm PBE}$ results.

For TiO_2, the obtained direct gap for rutile is smaller by ~ 0.7 eV than for anatase. The computed gap for rutile is smaller than previous G_0W_0@LDA/GGA results spanning the range from 3.3 to 3.6 eV [47, 149, 176, 332], but close to the value of 2.85 eV obtained within G_0W_0@LDA+U [239], and compatible with photoemission measurements, giving 3.3 ± 0.5 eV [297] and ~ 3.1 eV [270] for the QP gap. For the brookite phase, a recent G_0W_0@PBE investigation yielded a direct band gap of 3.86 eV [332], larger than in rutile and anatase, and in good agreement with the sc-PBE0α_{ϵ_∞} result.

For ZrO_2, the predicted band gap ordering for the three phases is in line with previous investigations using hybrid functionals [96] and GW, [146, 192] and numerical values agree very well with the GW_0@LDA results of Ref. [146] reporting an indirect band gap of 5.34 eV, 5.92 eV and 4.97 eV for the monoclinic, tetragonal and cubic structures, respectively.

Only recently GW calculations were performed for all the ambient pressure polymorphs of WO_3 [251]. Computed band gaps within G_0W_0@LDA are in good agreement with the sc-PBE0α_{ϵ_∞} results, and confirm the obtained band gap ordering.

Table 5.7: *Optimized cell parameters for selected polymorphs of the studied materials, computed at different levels of theory. Calculations are performed within the LCAO scheme using* CRYSTAL09.

	Type	Parameter	PBE	PBE0	sc-PBE0α_{ϵ_∞}	Expt.[a]
MgO	rs	a (Å)	4.258	4.207	4.185	4.212
ZnO	wz	a (Å)	3.292	3.264	3.257	3.250
		c (Å)	5.119	5.077	5.071	5.207
TiO_2	a	a (Å)	3.821	3.765	3.789	3.781
		c (Å)	9.672	9.656	9.631	9.515
ZrO_2	t	a (Å)	3.640	3.608	3.613	3.571
		c (Å)	5.288	5.209	5.224	5.182
WO_3	γ-m	a (Å)	7.444	7.334	7.348	7.306
		b (Å)	7.672	7.605	7.616	7.540
		c (Å)	7.885	7.801	7.817	7.692
		β (°)	90.65	90.61	90.62	90.88

[a] Experimental lattice constants are taken from the following references: Ref. [186] for MgO, Ref. [58] for ZnO, Ref. [138] for TiO_2, Ref. [291] for ZrO_2, Ref. [190] for WO_3.

5.2.3 Equilibrium Geometries

Finally, in Table 5.7 the optimized cell parameters computed within PBE, standard PBE0 and dielectric-dependent PBE0 are reported for some selected polymorphs. Full optimization of both cell parameters and atomic positions is carried out. Results indicate that sc-PBE0α_{ϵ_∞} generally performs as well as standard PBE0 for calculation of structural properties, the error with respect to experimental lattice constants being in most cases within $\sim 2\%$.

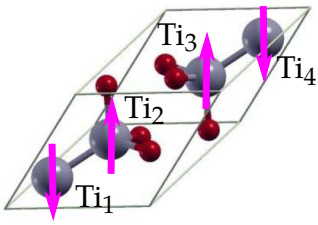

Figure 5.2: *Bulk primitive cell of corundum Ti_2O_3. Grey and red spheres represent Ti and O atoms, respectively. The arrows schematically represent the spin configuration of the Ti 3d electrons corresponding to the computed sc-PBE0α_{ϵ_∞} ground state.*

Table 5.8: *Relative total energies E_{tot} (meV, estimated error 1 meV) per unit Ti_2O_3, band gap E_g (eV, estimated error 20 meV) and lattice constants a and c (Å), computed at the sc-PBE0α_{ϵ_∞} level. AFM1 and AFM2 labels two different antiferromagnetic (AFM) configurations; FM refers to the ferromagnetic one. $\alpha = 12.8\%$ is used in the dielectric-dependent PBE0; this value is obtained from the electronic dielectric constant $\epsilon_\infty = 7.81$ computed within PBE0α_{PBE0} in the AFM1 configuration. Total energies are reported relative to that of the lowest-energy AFM1 phase. For spin configurations the notation (Ti_1, Ti_2, Ti_3, Ti_4), with $Ti_n = +/-$, indicates spin up/down for the 3d electron of the Ti_n atom in the unit cell (see Figure 5.2). Calculations were performed within the LCAO scheme with* CRYSTAL09 *and for the optimized structure.*

Magnetic configuration	E_{tot}	E_{g}	a	c
AFM1 $(-, +, +, -)$	0	0.59	5.167	13.671
AFM2 $(-, +, -, +)$	6			
Nonmagnetic	13			
FM	13			
Expt.[a]		0.11	5.157	13.610

[a] Experimental band gap from combined conductivity and thermoelectric coefficient measurements [280]. Experimental lattice parameters from Ref. [260].

5.3 Stoichiometry Transformations: Chemical Reduction of Bulk TiO_2

Finally, the performance of the dielectric-dependent hybrid method in calculating reaction energies for the chemical reduction of TiO_2 to Ti_2O_3 in their bulk phases is analyzed. Differences in total energy involved in chemical reactions are typically of the order of few eV, compared to few meV or tens of meV, normally observed in crystallographic phase transitions. This allows one to meaningfully address the issue of whether the method is able to quantitatively improve over standard hybrid functionals in terms of computed ground-state total energies.

Two reduction pathways for bulk TiO_2 are considered as a case study of reduction mechanisms typically occurring in oxide materials:

$$2\,TiO_2 + H_2 \rightarrow Ti_2O_3 + H_2O \tag{5.1}$$

$$2\,TiO_2 \rightarrow Ti_2O_3 + {}^1\!/_2\,O_2. \tag{5.2}$$

Ti_2O_3 is a small gap insulator featuring corundum structure. Its bulk unit cell is rhombohedral and contains two Ti_2O_3 units, with atoms positioned as shown in Figure 5.2. The magnetic properties of Ti_2O_3

at low temperature have been subject of intensive investigation both at the experimental [1, 150] and theoretical level [38, 117, 139, 142]. However, no general consensus has been achieved on whether the low-temperature ground state of Ti_2O_3 is antiferromagnetic (AFM) or nonmagnetic. Here, the issue is re-investigated using both standard and dielectric-dependent PBE0. Local/semilocal DFT predicts the material to be metallic [202], at odds with the experimental evidence. Spin-polarized PBE0 calculations yield an AFM ground state with AFM1 $(-, +, +, -)$ spin configuration of Ti $3d$ electrons (see Figure 5.2). Previous studies within HF [38], hybrid DFT, and DFT+U [139] reported the same result. Ferromagnetic (FM) and nonmagnetic solutions are found to be much higher in energy at the PBE0 level. Interestingly, different investigations using screened-exchange hybrid functionals predicted the nonmagnetic configuration to be the most stable [117, 142]. Calculations using the dielectric-dependent PBE0α_{PBE0} functional confirm the magnetic configuration to be the AFM1, as reported in Table 5.8. The corresponding computed band gap of 0.59 eV is larger than the experimental value of 0.11 eV obtained by combined conductivity and thermoelectric coefficient measurements [280], but still closer than the PBE0 result of 2.63 eV, and in agreement with screened-exchange hybrid calculations [142] reporting a value of 0.57 eV.

For TiO_2, the rutile phase is considered in the present analysis, since experimental data are available for the reaction enthalpy of the transformation to Ti_2O_3. Full geometry optimizations are carried out within PBE0, PBE0α_{PBE0}, and sc-PBE0α_{ϵ_∞}. A band gap of 3.03 eV is obtained within sc-PBE0α_{ϵ_∞}, which is very close to the previously reported value of 2.96 eV (see Table 5.6), obtained using as initial guess for α the value computed within PBE instead of PBE0.

The computed energies of reactions (5.1) and (5.2) are summarized in Table 5.9. The dielectric-dependent PBE0 yields results in excellent agreement with experiment, with a clear improvement over PBE0. Other first-principles investigations within GGA + U found that choosing the value of the Hubbard parameter U between 2 eV and 3 eV results in reaction energies close to experiments [139, 191]. However, the chosen value of U strongly affects the computed electronic band structure. Ti_2O_3 has been reported to be metallic when using $U = 2$ eV, and $U = 3$ eV is needed to open a gap of comparable size with the experimental one. For TiO_2 it has been demonstrated that U should be set to ~ 10 eV in order to obtain a band gap of 3 eV [249], close to the experimental

Table 5.9: *Reaction energies (eV, estimated error 10 meV) for the reduction of bulk TiO_2, computed with different hybrid functionals, and compared with experimental heats of reaction. The exchange fraction $\alpha = 1/\epsilon_\infty$ (%) used in calculations on TiO_2 and Ti_2O_3 is reported. The computed dielectric constants ϵ_∞ within PBE0 and PBE0α_{PBE0}, respectively, are: 6.10 and 6.65 for TiO_2, and 5.63 and 7.81 for Ti_2O_3 (estimated errors within 0.02). Calculations were performed within the LCAO scheme using* CRYSTAL09.

		PBE0	PBE0α_{PBE0}	sc-PBE0α_{ϵ_∞}	Expt.[a]
	$\alpha(TiO_2)$	25	16.4	15.0	
Reaction	$\alpha(Ti_2O_3)$	25	17.8	12.8	
$2\,TiO_2 + H_2 \rightarrow Ti_2O_3 + H_2O$		0.96	1.22	1.26	1.30
$2\,TiO_2 \rightarrow Ti_2O_3 + {}^1\!/_2\,O_2$		3.55	3.81	3.85	3.81

[a] See Ref. [191] and references therein.

one. Another deficiency of the DFT+U approach lies in the failure to predict accurate lattice parameters for Ti_2O_3 [139]. Both of the above issues are solved using the sc-PBE0α_{ϵ_∞}method, which has been found out to be able to accurately characterize reaction energies, electronic structures and structural parameters within a single, parameter-free, and fully *ab initio* theoretical approach.

5.4 Summary and Conclusions

The dielectric-dependent hybrid paradigm has been tested for calculation of the electronic structure and the ground-state total energy of bulk oxide materials. Concerning calculation of the band gap, this approach has generally proved superior to any other density-functional method, either at the local/semilocal or fixed-α hybrid functional level. Its predictive accuracy is comparable to that achieved by the more expensive many-body GW methods. An exception is represented by MgO and ZnO, for which standard hybrid methods perform better. In this respect, the dielectric-dependent approach seems to be most accurate for materials with a dielectric constant higher than $\sim$ 4. Geometry optimization at each self-consistency step makes this issue more critical.

Concerning calculation of ground-state total energies, two test cases have been considered: prediction of the crystallographic phase stability order (as determined by the zero temperature electronic contribution), and determination of reaction energies associated to model reduction pathways for a prototypical bulk oxide material, titanium dioxide. In the former case, the sc-PBE0α_{ϵ_∞}approach proves to perform similarly to PBE0, except in the case of WO_3, in which a peculiar correlation between structural and electronic properties leads to inversion of the (qualitatively correct) phase stability order predicted within PBE0. Instead, computed reaction energies for the $TiO_2 \rightarrow Ti_2O_3$ transformation clearly improve when passing from PBE0 to sc-PBE0α_{ϵ_∞}, which is able to provide at the same time a good description of the electronic and structural properties for both compounds.

For structural properties (equilibrium cell parameters) the sc-PBE0α_{ϵ_∞}proves essentially as good as PBE0 for all the analyzed materials.

In conclusion, the method qualifies for accurate determination of both excited- and ground-state properties of bulk oxides, constituting

a useful tool for describing those systems in which a good description of both is called for. However, in its present form, the method can only be applied to homogeneous or quasi-homogeneous systems. The study of point defects in bulk materials, for which at low defect concentrations the dielectric properties are essentially determined by the host pristine material, but the electronic and optical properties are substantially altered, constitutes the first step towards treatment of weakly inhomogeneous systems. Due to the weak deviation from a perfect crystal, the sc-PBE0α_{ϵ_∞} method is still applicable in its present formulation, and turns out to be very promising in description of the optical features in defective semiconductors and insulators. Analysis of its performance in these systems is the central topic of the following chapters.

Computational details

PW-PP calculations PW-PP DFT calculations were performed using the QUANTUM ESPRESSO package [103]. A norm-conserving Troullier-Martins PP [304] with $2s$ and $2p$ electrons in the valence was used for oxygen, while for metals, norm-conserving Hartwigsen-Goedecker-Hutter PPs [111, 123] explicitly including semicore electrons in the valence were employed. A careful treatment of core-valence partitioning in metal atoms has been proved to be determinant for obtaining reliable GW corrections [197], especially in II-IV semiconductors, as analyzed in detail in Ref. [262]. It turns out that explicit treatment of semicore electrons of the metal atom in oxides is always needed, as it was experienced in the case of ZnO (see discussion in Section 5.1.3). Since here PW-PP DFT calculations served as starting point for subsequent GW calculations, choice of the PPs was guided by the above considerations. Hence, $(n-1)\,s$ and $(n-1)\,p$ electrons of the metal atom were always explicitly treated in the valence, where n is the main quantum number of the outermost occupied electronic shell. PPs were tested by checking that computed optimized lattice parameters and electronic band structures were in agreement with the reference literature, as well as that the former showed the expected trends at the LDA and GGA levels of theory when compared to experiment.

Table 5.10: *Computational parameters of DFT PW calculations:* $\mathbf{k}$*- and* $\mathbf{q}$*-point BZ sampling and PW kinetic energy cutoff* E_{cut} *(Ry).*

	$\mathbf{k}$-point grid	$\mathbf{q}$-point grid	E_{cut}
MgO	4x4x4	4x4x4	500
ZnO	4x4x2	4x4x2	300
TiO_2	4x4x2	4x4x2	150
ZrO_2	4x4x4	4x4x4	120
WO_3	2x2x2	2x2x2	100

Including semicore electrons in the self-consistent calculation of charge density clearly requires using a large number of PW components in the expansion of the KS orbitals, as semicore electrons are tightly bound to the nucleus. This dramatically affects the resulting computational cost, especially in hybrid functional calculations, in which evaluation of the nonlocal EXX using a PW basis set under periodic boundary conditions is particularly demanding.

BZ sampling is also critical for obtaining well-converged results. Monkhorst-Pack [213] $\mathbf{k}$-point grids were chosen in such a way that estimated errors in computed band gaps and total energies were within 20 meV. The $\mathbf{q}$-point grids needed for evaluation of the EXX contribution were chosen equal to the $\mathbf{k}$-point grids (see Table 5.10).

Full structural optimizations were carried out only at the LDA and GGA levels, as optimizations with hybrid functionals are very expensive with QUANTUM ESPRESSO. A quasi-Newton Broyden-Fletcher-Goldfarb-Shanno (BFGS) scheme was adopted for the search of the energy minimum. Convergence thresholds on atomic forces and pressure were set to 10^{-3} a.u. and 0.5 kbar, respectively.

LCAO Calculations The LCAO method, as implemented in the CRYSTAL09 package [78], was employed for DFT calculations using localized basis sets. All-electron calculations were performed for MgO, ZnO and TiO_2, while for ZrO_2 and WO_3 the inner electrons of Zr and W atoms were described through effective core PPs (ECPs) generated with relativistic atomic calculations at the HF level.

The adopted all-electron and valence basis sets were previously tested in analogous solid state calculations. The following Gaussian-type all-electron basis sets were employed in the calculations performed in this work: 8-511(d1) for Mg from Ref. [307], pob-TZVP for Zn from Ref. [240], 8-4611(d41) for Ti from Ref. [333]. For O the 8-411(d1) basis set of Ref. [263] was used, except for the case of ZrO_2, in which the basis set from Ref. [220] was adopted, following the choice of a previous investigation on this material [96]. Small-core Hay-Wadt ECPs [125] were employed for the heavy Zr and W atoms. For Zr the $4s$, $4p$, $5s$ and $4d$ electrons were explicitly treated, using the 311(d31) valence electron basis set of Ref. [30]. For W, electrons belonging to shells $5s$, $5p$, $6s$ and $5d$ were included in the valence, and a modified Hay-Wadt double-ζ basis set [315] was employed to describe them.

In the analysis of the chemical reduction of TiO_2 to Ti_2O_3, for both Ti and O atoms, different basis sets were adopted, following previous HF and DFT + U studies [38, 139] on the same materials. In molecular calculations basis sets for oxygen [263] and hydrogen [99] were chosen such that experimental atomization energies for H_2, O_2 and H_2O molecules were well-reproduced at the PBE0 level.

As a high accuracy is needed when studying the energetics, BZ samplings were performed so as to ensure convergence within 1 meV per formula unit on total energies for the different polymorphs of the materials under investigation. This corresponds to Monkhorst-Pack grids including a number of $\mathbf{k}$-points in the irreducible BZ at least as large as in PW-PP calculations.

Different cutoff thresholds for the evaluation of Coulomb and exchange integrals written in terms of Gaussian-type orbitals were set to their standard values in CRYSTAL09.[3] The self-consistent field was considered converged when total energy

[3]The following thresholds were used: 10^{-7} for Coulomb overlap tolerance, 10^{-7} for Coulomb

Table 5.11: *Computational cutoff parameters used in GW calculations: cutoff energies E_{cut}^{eps} and E_{cut}^{xc} (Ry) controlling the size of the dielectric matrix in reciprocal space and the number of plane waves in the expansion of xc potential, respectively; number of empty states included in the evaluation of the polarizability χ_s and of the Coulomb hole (COH) term.*

			Empty states	
	$E_{\text{cut}}^{\text{eps}}$	$E_{\text{cut}}^{\text{xc}}$	χ_s	Σ_{COH}
MgO	90	80	300	900
ZnO	70	200	200	2750
TiO_2	14	50	500	1300
ZrO_2	25	50	500	2100
WO_3	16	60	800	2300

difference between two subsequent cycles become lower than 10^{-6} a.u.

Geometry optimizations were performed at both local/semilocal and hybrid functional level by fully optimizing both lattice parameters and atomic coordinates. A quasi-Newton algorithm with a BFGS Hessian updating scheme was adopted for the search of the energy minimum. Geometry optimizations were terminated when the maximum component and root-mean-square (rms) of energy gradients become lower than 0.00045 and 0.00030 a.u., respectively, and the maximum and rms atomic displacements were below 0.00180 and 0.00120 a.u., respectively.

***GW* calculations** *GW* calculations were performed using the BERKELEYGW code [70, 141]. The non self-consistent G_0W_0 scheme was adopted, in which both the electronic Green's function and the screened Coulomb potential were constructed starting from DFT-KS eigenfunctions and eigenvalues computed at the PBE level of theory (G_0W_0@PBE). No further iteration was carried out on either the Green's function or the self-energy (see Section 3.2). DFT calculations were performed at the experimental geometry using QUANTUM ESPRESSO. The frequency dependence of the dielectric function was evaluated through the plasmon-pole approximation of Hybertsen and Louie [141].

Concerning cutoff parameters discussed in Section 3.2, careful convergence studies were carried out for all the materials, ensuring that accuracy within 50 meV on QP band gaps was reached (see the procedure described in Ref. [196]). Cutoff parameters used in the presented calculations are reported in Table 5.11. BZ samplings were performed as in corresponding PW-PP DFT calculations.

DFT xc functionals In local/semilocal calculations, the Perdew-Zunger expression for the correlation energy [248] for the LDA, and the PBE parametrization [242] for the GGA were employed. Hybrid functionals considered were the full-range PBE0 [244] and the screened-exchange HSE06 [132,133].[4] The B3LYP hybrid functional [20,292] was also tested.

penetration tolerance, 10^{-7} for exchange overlap tolerance, 10^{-7} for exchange pseudo-overlap in direct space, 10^{-14} for exchange pseudo-overlap in reciprocal space.

[4]The CRYSTAL14 code is used for HSE06 calculations, as screened-exchange hybrid functionals are not implemented in CRYSTAL09.

In hybrid calculations with QUANTUM ESPRESSO, PPs were generated at the closest GGA level, i.e. PBE for PBE0 and HSE06, and BLYP for B3LYP. All PPs were obtained from the QUANTUM ESPRESSO PP library, except for the Zn atom, for which the PP was converted from the `cpmd` format [111, 123, 161].

For defining dielectric-dependent hybrid functionals, electronic dielectric constants were computed within the CPKS method as implemented in the CRYSTAL09 code (see Section 3.3.2).

CHAPTER 6

Oxygen Vacancies in Wide-Gap Metal Oxides

Oxygen vacancies (O vacancies) represent a prototypical native point defect in oxide materials, where they are inevitably formed throughout the synthesis process, as a result of exposition of the samples to high temperatures and strongly reducing atmospheres. The presence of O vacancies has substantial impact on the observed optical, magnetic, and electrical properties of substoichiometric oxides [98,231], and a correct interpretation of, e.g., optical processes observed in O-deficient wide-gap oxides requires fundamental understanding of these modifications.

Removal of an O atom from a stoichiometric oxide results in the presence of two unpaired electrons, previously shared in a chemical bond between the removed O atom and a metal cation, which typically redistribute in the neighborhood of the vacancy. In transition metal oxides (TMOs), the unpaired (or excess) electrons occupy an empty or partially filled d orbital belonging to metal ions in the vicinity of the vacancy. However, the degree of localization of the excess charge is actually strongly influenced by the nature of the chemical bonding in the host material. In the extreme case of a highly ionic oxide like

MgO, the strong electrostatic potential between ions (the Madelung potential) stabilizes the excess charge in the vacancy void, forming a so-called color center (F center), which gives rise to characteristic optical transitions which, provided the vacancy is located at the MgO surface, fall in the visible range of the electromagnetic spectrum. A similar phenomenology can sometimes be encountered in reduced TMOs, although in this case a residual charge localization on d orbitals of the nearby undercoordinated cations can still be present. This situation is common in nonreducible oxides like ZrO_2: the two excess electrons pair up in a closed-shell singlet state, and are confined by the Madelung potential to stay in the vacancy void. The vacancy formation energy is higher than in reducible oxides, such as TiO_2, WO_3 or NiO: here the Madelung field is less intense, and the excess charge localizes on cations not necessarily in the immediate vicinity of the vacancy [98,233]. The accompanying lattice distortion involves to certain degree some fully coordinated atoms of the bulk crystal, leading to formation of a polaron. In reducible oxides, formation of a magnetic ground state, in which the unpaired electrons occupy different cation d orbitals with the same spin polarization (realizing a triplet spin configuration), is often favored. EPR techniques are very useful for characterizing these magnetically active defect centers in substoichiometric oxides [102].

In wide-gap oxides, O vacancies lead to building up of deep occupied defect levels in the band gap of the material. Electronic transitions between defect states and the bulk VB and CB are the cause of the peculiar optical properties observed, e.g., in absorption and photoluminiscence (PL) spectroscopies. For example, absorption of a photon in the reduced oxide may excite an excess electron from the defect level to the bulk CB: if the vacancy is initially neutral ($q = 0$, with both excess electrons occupying the defect level), then the final defect state will be characterized by a formal charge $q = +1$. An extra electron may also decay into the VB, and the resulting electron-hole recombination gives rise to a luminiscence feature. Ionization of the vacancy may also be the consequence of thermal excitation of the excess electrons in the CB: in this case the vacancy is stable, at room temperature, in a charged state $q = +1$ or $q = +2$. Electrons excited in the CB contribute to the n-type conductivity which is sometimes observed in O-deficient TMOs.

Theoretical modeling of defects in oxides ideally aims to account for the complex electronic, optical, and magnetic features exhibited in these systems within one single method. In TMOs, correctly describing

localization of the defect-related excess charge on transition metal cations d orbitals is challenging: local or semilocal DFT typically fails in providing the correct picture, due to the inherent self-interaction error which tends to delocalize the charge over a too large portion of the crystal (see, e.g., the case of reduced TiO_2, Ref. [73]). Correspondingly, the electronic states introduced by the defect are wrongly predicted to merge with the bulk band edges, even when experiments indicate a deep donor or acceptor behavior; hence, the observed defect-related optical features cannot be suitably accounted for by standard DFT methods [45].

Hybrid xc functionals, when used in conjunction with the charge-transition levels (CTLs) formalism for computation of electronic transitions (see Section 4.3.1), constitute an excellent tool for accurately characterizing the excess charge distribution and the related localization of defect states in the band gap [72]. In fact, on the one hand they are expected to be less affected by the self-interaction error, which instead is present in local/semilocal approximations, and thus they are more suitable to correctly predict charge localization around the defect (i.e., to yield the correct ground state). On the other hand, as discussed in Chapter 5, even the electronic structure of nondefective materials, when computed at the hybrid functional level, is quantitatively closer to experiments; this clearly helps in correctly locating defect electronic levels in the band gap, thus giving a suitable description of the related optical features.

The sc-PBE0α_{ϵ_∞}hybrid functional is a promising candidate for addressing all the above issues, being a nonempirical hybrid xc potential derived as an approximation of the exactly self-interaction free COHSEX self-energy (see also Chapter 7), generally performing better than standard hybrid functionals in the calculation of electronic structures. The *ab initio* evaluation of the admixed portion of EXX makes it potentially superior to other semiempirical DFT methods, in terms of predictive power, when applied to defective materials. Clearly, since magnetic ground states can often be realized in substoichiometric oxides (and this in turn also affects the description of the excess charge localization and, hence, of the electronic levels in the band gap), spin-polarized DFT calculations should be performed.

Given the above considerations, the present chapter is devoted to an assessment of the efficiency of the sc-PBE0α_{ϵ_∞}functional in describing defective wide-gap oxide semiconductors, with particular concern to O vacancies. Three different O-deficient TMOs are considered: rutile and anatase TiO_2, WO_3 in the room-temperature γ-monoclinic crystal

structure, and tetragonal ZrO_2.

6.1 Reduced Titanium Dioxide: TiO_{2-x}

The optoelectronic properties of O-deficient TiO_2 (TiO_{2-x}) have been extensively investigated both experimentally [75] and theoretically [63, 72, 89, 144, 195, 203, 204], making this system an ideal model for testing the performance of the dielectric-dependent hybrid method in describing the electronic properties of defective oxides.

The observed n-type conductivity of TiO_2 samples is usually rationalized with the presence of O vacancies, acting as intrinsic donors, formed under strongly reducing synthesis conditions [75]. The larger intrinsic conductivity of the anatase phase, which also makes it more appealing for applications, has been explained with the different behavior of O vacancies in the two polymorphs [63, 203].

On the other hand, optical experiments found the associated defect levels to be located deep in the TiO_2 band gap. Infrared absorption (IR) [57] ultraviolet photoelectron [221] and electron energy-loss experiments [128] on reduced rutile samples all agree in attributing a feature at about 1 eV below the CB to the presence of O vacancies. Theoretical calculations have revealed that the apparent contradiction between the observed n-type conductivity and the deep nature of optical levels in TiO_{2-x} is understood on the basis of the different character (shallow vs. deep) of the O-vacancy thermodynamic and optical transition levels [63, 203].

Furthermore, the excess charge introduced by the O vacancy has been shown to redistribute differently in the host crystal, according to whether the rutile or anatase polymorph is considered, resulting in localization of the extra electrons at topologically different Ti atoms [272].

It is the goal of the present section to re-investigate these issues within the sc-PBE0α_{ϵ_∞}approach, which by construction is sensitive to the different electronic screening features of the rutile and anatase phases.

6.1.1 Rutile

The ground state of rutile TiO_2 with a neutral O vacancy is found to be a triplet, with the two excess electrons redistributing over the whole (110) plane containing the vacancy [Figure 6.1(a)]. Most of the excess charge is contributed by the fully-coordinated Ti_2 (spin density $n_s = n_\uparrow - n_\downarrow = 0.70$), Ti_3 (two equivalent atoms, $n_s = 0.15$ each)

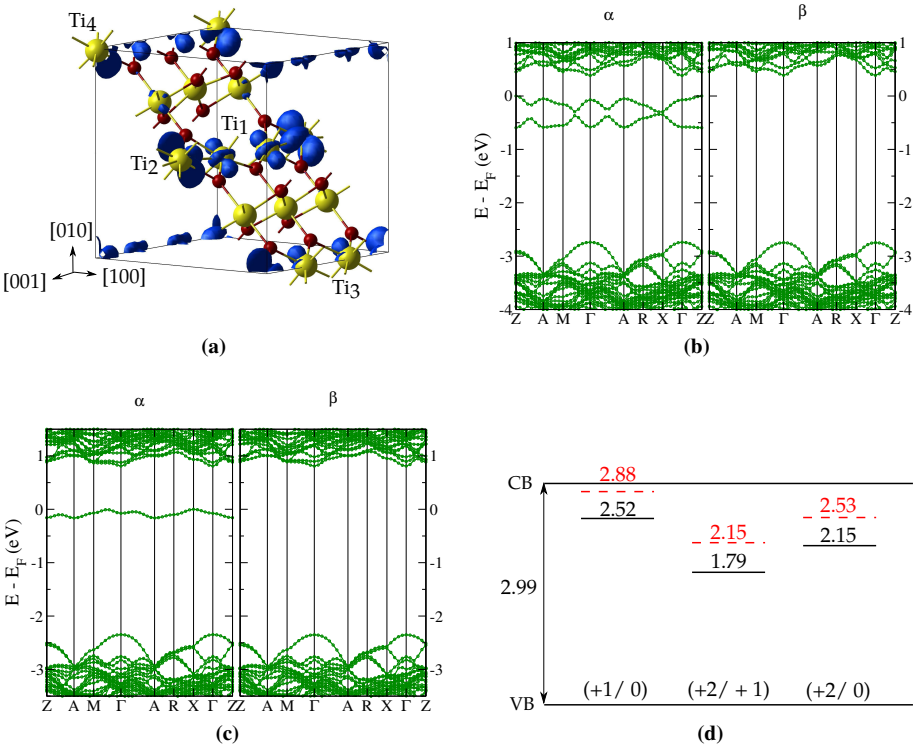

Figure 6.1: *O vacancy in rutile TiO_2. (a) Isosurface plot of the charge density associated with the excess electrons for the neutral vacancy. The yellow (large) and red (small) spheres represent Ti and O atoms, respectively. Only the atomic plane containing the O vacancy is shown. Corresponding electronic band structure (shown for both spin majority, α, and spin minority, β) computed at the equilibrium geometry, (b) for the neutral, and (c) for the singly-charged O vacancy. (d) Computed optical (solid, black) and thermodynamic (dashed, red) charge-transition levels. Positions (eV) are given with respect to the top of the valence band (VB).*

and Ti_4 ($n_s = 0.65$) atoms. The two under-coordinated Ti_1 atoms instead accommodate a smaller fraction of the charge ($n_s = 0.13$ each).[1] This picture is in qualitative agreement with recent scanning tunneling microscopy (STM) measurements on the rutile (110) surface [272] suggesting localization of the extra electrons at Ti atoms far from the vacancy[2] with the formation of a polaron. EPR measurements also

[1] A triplet broken-symmetry solution is found higher in energy by $\sim$ 60 meV.

[2] In this case, actually, due to presence of the surface, the extra electrons are preferably localized at subsurface Ti atoms, but molecular dynamics simulations indicate that a significant contribution is also given by a fivefold coordinated Ti atom at the (110) surface far from the vacancy (see Ref. [272]). Notice

confirmed the presence of Ti^{3+} ionic species formed as a consequence of electron trapping at Ti^{4+} centers in TiO_{2-x} crystals [28, 328].

Notice that localization of the unpaired electrons at Ti^{3+} centers is correctly described only within the hybrid functional approach [73] at variance with the case of local/semilocal functionals which commonly fail in giving a correct picture, due to incomplete cancellation of the self-interaction error. Notice also that the above physical picture is correctly captured only at the level of spin-polarized DFT, while neglecting spin polarization leads to a completely different description, with the excess charge trapped in the vacancy void in a configuration typical of a color center [144, 195]. The corresponding closed-shell ground state singlet is found to be highly unstable, being above the triplet solution by nearly 2 eV in total energy. In the spin-compensated case, the two undercoordinated Ti atoms relax towards the vacancy, shortening their distance to 2.90 Å (compared with an unrelaxed distance of 2.99 Å); instead, in the true, spin-polarized ground state (triplet), they move away from the vacancy, resulting in an equilibrium distance of 3.37 Å.

The obtained magnetic ground state is compatible with the evidence of triplet and doublet signals observed in EPR measurements on rutile TiO_{2-x} crystals, and attributed to neutral and singly-charged O vacancies, respectively [28, 328].

A triplet configuration can be also stabilized in which nonetheless the excess charge is confined at the vacancy void, a situation typically encountered in nonreducible O-deficient oxides [232].[3] However, the corresponding ground-state energy is found to be ~ 0.8 eV higher than that of the triplet polaron-like solution represented in Figure 6.1(a). This finding definitely rules out the possibility of an F-center-like behavior of O vacancies in rutile, in favor of the polaron picture, which is in agreement with all the available experimental evidence. This conclusion is also corroborated by other theoretical investigations using larger supercells [63].

The bulk band gap of rutile TiO_2, as computed using the self-consistent hybrid functional, is found to be 2.99 eV, a value compatible with the available data from photoemission/inverse photoemission experiments, yielding an electronic gap of 3.3 ± 0.5 eV

also that agreement with the STM findings can be only qualitative in nature, since the calculations presented here are performed for the O vacancy in bulk TiO_2.

[3] Such solution was obtained by augmenting the basis set with Gaussian functions corresponding to a ghost O atom positioned in the vacancy; furthermore, a polaron-like configuration was stabilized using the very same basis set, which allowed for comparison of the corresponding total energies.

[297] and ~ 3.1 eV [270]. Upon removal of an O atom, the two resulting excess electrons occupy two triplet defect states located in the band gap at about $0.5 - 1$ eV below the CB [Figure 6.1(b)]. As the vacancy is ionized, the remaining excess electron localizes at the three Ti atoms in the row of the O vacancy, again with a dominant contribution from the fully-coordinated Ti_2 atom ($n_s = 0.68$). The associated occupied defect level is still localized at about 1 eV below the CB, and turns out to be quite flat, as shown in Figure 6.1(c).

Figure 6.1(d) shows the computed CTLs. The optical levels lie deep in the band gap, resulting in excitation energies from the defect states to the CB of 0.47 eV, 1.20 eV and 0.84 eV for the $(+1/0)$, $(+2/+1)$ and $(+2/0)$ transitions, respectively. Notice that, as a result of the finite supercell size, the residual defect-defect interaction (i.e. overlap between polarons in neighboring cells) [273] leads to a small dispersion of the highest-occupied defect band (about 0.2 eV). This should be considered as indicative of the numerical accuracy of the computed level. Also taking into account this limitation, the computed optical levels can be considered consistent with the features, found at about 1 eV below the bottom of the CB, observed in experimental spectra [57, 128, 221]. In particular, IR measurements [57] revealed two absorption peaks at 0.75 eV and 1.18 eV for the single and double ionization, which correlate well with the computed first and second optical ionization energies for the O vacancy.[4]

Thermodynamic transition levels are significantly higher in energy, reflecting the sizable structural relaxation associated with the polaronic distortion. The $(+1/0)$ adiabatic transition is located about 0.1 eV below the CB, indicating that the neutral O vacancy may be stable at sufficiently low temperature. The second transition $(+2/+1)$ is at 0.81 eV, while the double ionization [$(+2/0)$ transition] requires 0.46 eV. These values should be compared with measured thermal ionization energies observed in a range between 0.3 and 0.6 eV [101].

6.1.2 Anatase

The nature of the neutral O vacancy in anatase TiO_2 has been extensively studied in the theoretical literature [60, 63, 74, 89, 203, 204, 214]. However, no consensus has been achieved so far concerning the

[4]The computed $(+1/0)$ optical transition level is inherently less accurate than $(+2/+1)$ one (this is noticed *a posteriori* by comparison with the experimental absorption energies); in fact, the defect KS levels associated with the neutral O vacancy exhibit a dispersion of about 0.2 eV, whereas the KS defect level for the singly-charged O vacancy shows no dispersion; thus, the $(+1/0)$ level should be attributed an accuracy of ~ 0.2 eV, while the $(+2/+1)$ level is not affected by the band dispersion error.

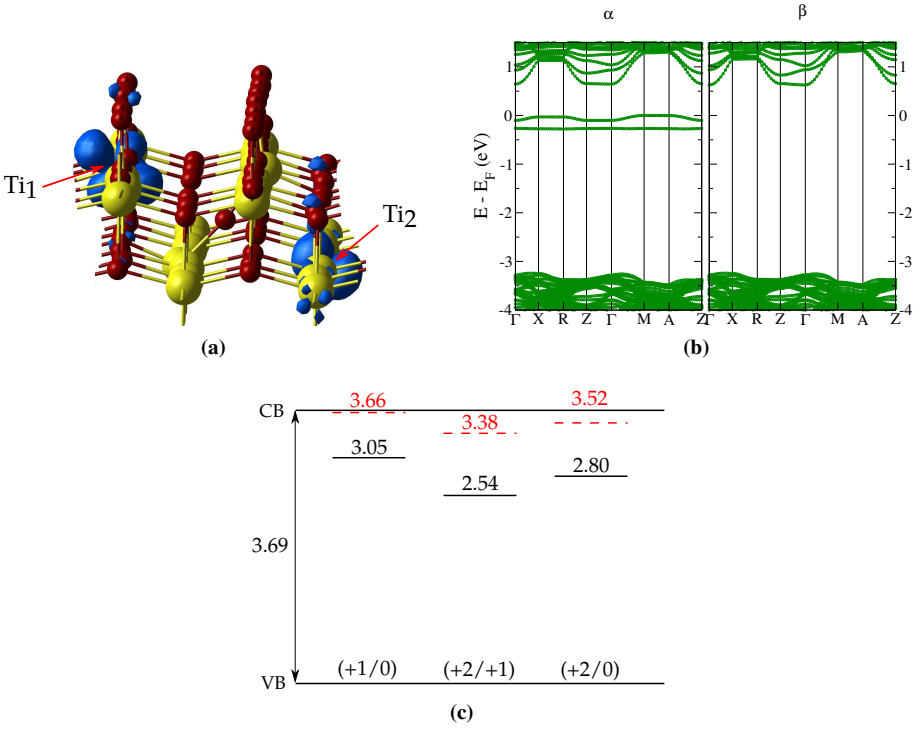

Figure 6.2: *O vacancy in anatase TiO_2. (a) Isosurface plot of the charge density associated with the excess electrons for the neutral vacancy. The yellow (large) and red (small) spheres represent Ti and O atoms, respectively. (b) Corresponding electronic band structure (shown for both spin majority, α, and spin minority, β) computed at the equilibrium geometry. (c) Computed optical (solid, black) and thermodynamic (dashed, red) charge-transition levels. Positions (eV) are given with respect to the top of the valence band (VB).*

redistribution of the associated excess electrons in the host crystal. This difficulty may be rationalized with the presence of several minima on the relevant adiabatic potential energy surface when the structure is allowed to relax. The realization of a specific configuration is strongly related to the description of structural relaxation around the vacancy, which is affected, for example, by the employed DFT xc approximation.

The performed sc-PBE0α_{ϵ_∞}calculations yield a triplet ground state for the system with a neutral O vacancy, with a corresponding substantial rearrangement of the atomic positions around the vacancy. As shown in Figure 6.2(a), this implies a reduction of the local symmetry around the vacancy, with a nearest-neighbor O atom breaking its bond with a bulk Ti atom (Ti_2) and considerably relaxing towards the vacancy

void.[5] The same behavior was observed in other computational studies in which symmetry-breaking relaxation was allowed for [89, 204, 214]. The two excess electrons are found to be localized at two Ti atoms (the under-coordinated Ti_1, with $n_s = 0.92$, and the bulk Ti_2, with $n_s = 0.84$), similarly to a previous DFT + U investigation [214]. This finding is again in qualitative agreement with the STM investigation of Diebold and coworkers [272] which suggests that the excess charge preferably stays close to the vacancy in anatase, as opposed to the case of rutile in which the major contribution is given by the fully-coordinated Ti atoms (see Section 6.1.1).

For the system with a neutral vacancy, the computed electronic structure shows two triplet defect states at $\sim 0.7 - 1$ eV below the CB [Figure 6.2(b)]. The computed band gap is 3.69 eV, larger than the value of 3.42 eV obtained from optical absorption experiments at low temperature [295]. As discussed in Chapter 5, no data for the photoemission gap of stoichiometric anatase TiO_2 are available in the experimental literature. However, the computed gap is consistent with the results of GW calculations [100, 149, 176].

When the vacancy is ionized, a doublet occupied defect state locates at ~ 0.7 eV below the CB. The extra electron remains localized ($n_s = 0.92$) at the under-coordinated Ti_1 atom.

The computed thermodynamic transition levels are significantly shallower than in rutile [compare Figure 6.1(d) and Figure 6.2(c)]. In particular, the $(+1/0)$ level is practically resonant with the CB, suggesting that an electron can be thermally excited into the CB, thus accounting for the experimentally observed larger conductivity of anatase samples.

Similarly to the case of rutile, optical levels are substantially deeper than the corresponding thermodynamic ones. Computed excitation energies amount at 0.64 eV, 1.15 eV and 0.89 eV for the $(+1/0)$, $(+2/+1)$ and $(+2/0)$ transitions. Experimentally, a feature at $\sim 1.0 - 1.1$ eV below the CB has been reported in different experiments (combined x-ray photoemission/absorption spectroscopy [298] as well as scanning tunneling spectroscopy [272]) on reduced anatase surfaces, and has been attributed to O vacancies. Notice that, since the stable charge state of the O vacancy is $q = +1$ at room temperature, it seems likely that the $(+2/+1)$ transition is mainly probed in optical experiments; accordingly, the computed excitation energy of 1.15 eV

[5] A different triplet solution can be also stabilized, in which the point-group symmetry is preserved; it is however computed to be 0.23 eV higher in energy than the symmetry-broken one.

matches well with the measured spectroscopic features.

In conclusion, the sc-PBE0α_{ϵ_∞}approach proves capable of explaining several experimental signatures related to the different behavior of O vacancies in the two most common TiO_2 polymorphs, rutile and anatase. In particular, the larger conductivity of anatase can be understood by analysis of the thermodynamic transition levels, whereas the deep nature of the optical features observed in spectroscopies is rationalized on the basis of the computed optical transition levels. The presented results are in qualitative, and sometimes quantitative, agreement with previous investigations at the hybrid-functional [63, 72] and DFT + U [203] level.

6.2 Reduced Tungsten Trioxide: WO_{3-x}

The optical and electrical properties of substoichiometric WO_3 have been extensively investigated experimentally, particularly in relationship with the electrochromic properties exhibited in this material [65]. However, the microscopic origin of this effect, which is intimately connected with the typical blue coloration of WO_{3-x} films [69, 137] has been longly debated, and several models have been put forward for rationalizing experimental observations [65, 66, 83].

On the theory side, a few first-principle investigations have been performed on O-deficient γ-monoclinic WO_3 [43, 173, 211], but among them only one employed state-of-the-art methods yielding a correct description of the bulk electronic structure [316]. The latter study of Wang *et al.* revealed a delicate interplay between the concentration of O vacancies and the metallic or insulating nature of WO_{3-x} [316]. Furthermore, different behaviors were observed according to the orientation of the W-O-W chain along which the O atom is removed. For the sake of simplicity, the present discussion is however limited to the case of an O vacancy created along the a crystallographic axis of the monoclinic cell,[6] which will be shown to give rise to the observed optical features of WO_{3-x}, namely its blue coloration.

sc-PBE0α_{ϵ_∞}calculations indicate the closed-shell singlet solution to be more stable (by 0.17 eV) than the open-shell triplet one for the system with a neutral vacancy. Correspondingly, the two extra electrons occupy a defect state located at about $1.0-1.5$ eV below the CB [Figure 6.3(a)];

[6] Along the W-O-W a-chain, W-O bonds throughout have practically the same lengths, while for chains along the *b* and *c* axes, alternating long and short bonds are present; this peculiar structural feature determines a different behavior of the O vacancy, according to previous calculations [173, 316].

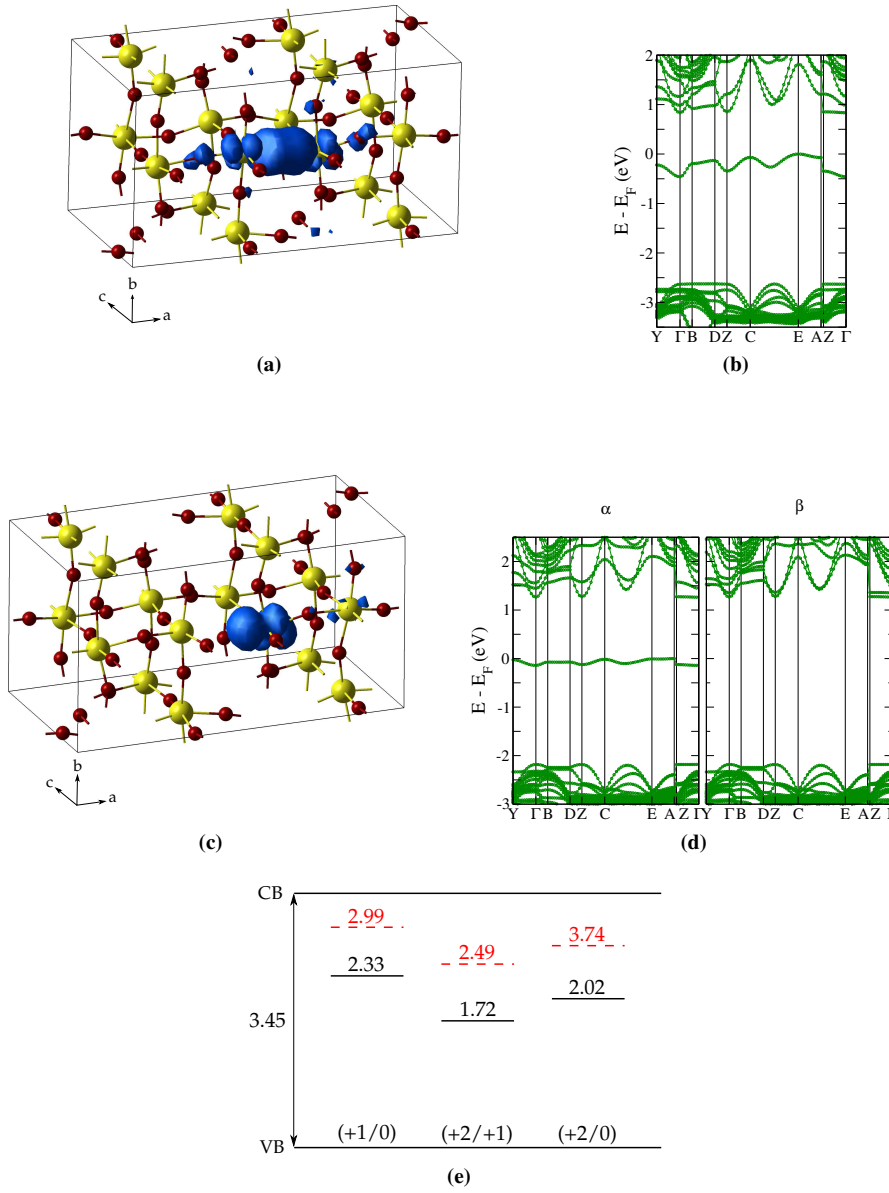

Figure 6.3: *O vacancy in γ-monoclinic WO_3. Isosurface plot of the charge density associated with the excess electrons for the (a) neutral and (c) singly-charged vacancy. The yellow (large) and red (small) spheres represent W and O atoms, respectively. Corresponding electronic band structure computed at the equilibrium geometry, for the (b) neutral and (d) singly-charged vacancy (shown for both spin majority, α, and spin minority, β). (e) Computed optical (solid, black) and thermodynamic (dashed, red) charge-transition levels. Positions (eV) are given with respect to the top of the valence band (VB).*

they localize mainly at the vacancy void, with a contribution from the two under-coordinated W atoms [Figure 6.3(b)].

As the vacancy is ionized, the extra electron localizes at one of the under-coordinated W atoms [Figure 6.3(c)], leading to formation of a reduced W^{5+} ionic state. In fact, one of the proposed models for the chromic mechanism in WO_{3-x} suggests that the presence of W^{5+} centers results from localization of excess electrons at W^{6+} sites upon light absorption [83]. As shown in Figure 6.3(d), this $5d$ electron occupies a defect state ~ 1.3 eV below the CB (~ 2 eV if the geometry of the neutral vacancy is retained).

Several experiments have been performed trying to elucidate the spectroscopic properties of substoichiometric WO_3 samples. The coloration efficiency of WO_{3-x} ($x < 0.4$) films has been reported to increase with increasing O deficiency [17]. Their typical blue coloration has been correlated with a broad absorption band with a maximum at 900 nm (~ 1.38 eV) [69, 137], although this feature has been observed in amorphous films. PL spectra exhibit an emission peak at 550 nm (~ 2.26 eV) [237], and photoelectron measurements confirm the presence of a defect state attributed to O vacancies at ~ 2 eV above the VB [31, 137].

The computed optical transition levels reported in Figure 6.3(e) are consistent with these observations, being positioned in the band gap in a range from 1.72 eV to 2.33 eV above the VB. The dispersion of the neutral vacancy defect state limits the accuracy of the calculated transition levels to ~ 0.2 eV. Notice that the computed band gap is larger than the reported values from absorption measurements (see Ref. [315] and references therein), but is in excellent agreement with ultraviolet direct/inverse photoemission experiments (3.45 eV from Ref. [208]), as well as with GW calculations [100, 254]. Thus, the excitation energy pertaining to, e.g., the $(+1/0)$ transition, which is obtained to be 1.12 eV, is compatible with the experimentally observed features in the absorption spectra.

The corresponding thermodynamic levels are predicted to be well-detached from the bottom of the CB, with a minimum excitation energy of 0.46 eV computed for the $(+1/0)$ transition. This implies that the specific kind of O vacancy considered here cannot contribute to the observed n-type conductivity of WO_{3-x} films, which has been attributed to O vacancies [106]. Thus, although the present work clarifies the origin of the spectroscopic features evidenced in reduced WO_3, in agreement with the previous investigation of Wang *et al.* [316], further

Table 6.1: *Formation energy (eV) for the O vacancy in the different materials.*

	x	Formation energy
TiO_{2-x} (rutile)	1/24	5.2
TiO_{2-x} (anatase)	1/32	4.7
WO_{3-x}	1/16	5.3
ZrO_{2-x}	1/36	6.6

study is necessary to elucidate the possibly different behavior of other inequivalent O vacancy sites.

Results of ongoing work by the author (cited as [C4] in Section 1.4) evidence that thermodynamic CTLs associated to O vacancies created along the b and c crystallographic axes of the monoclinic cell are much shallower than those computed for the vacancy along a, and thus may account for room-temperature conductivity of reduced WO_3 crystals. For the O vacancy along b and c, the most stable ground state is a triplet, and the excess electrons localize along the whole W-O-W chain interrupted by the vacancy. This charge redistribution comes with a strong lattice relaxation, a fingerprint of the polaronic nature of these defect states. Hence, the strong structural anisotropy of γ-monoclinic phase of WO_3 reflects on the features of the W-O bonds along the different crystallographic axes and, thus, on the nature of the O vacancy.

6.3 Reduced Zirconium Dioxide: ZrO_{2-x}

The O vacancy in reduced zirconia has been reported to exhibit features similar to those of F centers typically observed in nonreducible oxides [232], the extra charge being trapped in the vacancy void [108]. EPR studies [108] indicate that ZrO_2 is generally less prone to lose oxygen than other reducible oxides, such as TiO_2. This is confirmed by the sc-PBE0α_{ϵ_∞}calculations, which give a formation energy for the O vacancy in ZrO_{2-x} larger by more than 1 eV with respect to the other oxides (see Table 6.1), in agreement with previous investigations [98, 316].

As shown in Figure 6.4(a), upon removal of an O atom in ZrO_2, the two excess electrons stabilize in the vacancy void, with a contribution from the $4d$ orbitals of two nearest-neighbor Zr atoms (Zr_1 and Zr_2); the corresponding ground state is a closed-shell singlet. As expected, the neighboring Zr atoms relax towards the vacancy, resulting in a shrinking of the distance Zr_1-Zr_2 from 3.61 Å to 3.48 Å, and of Zr_1-Zr_3 from

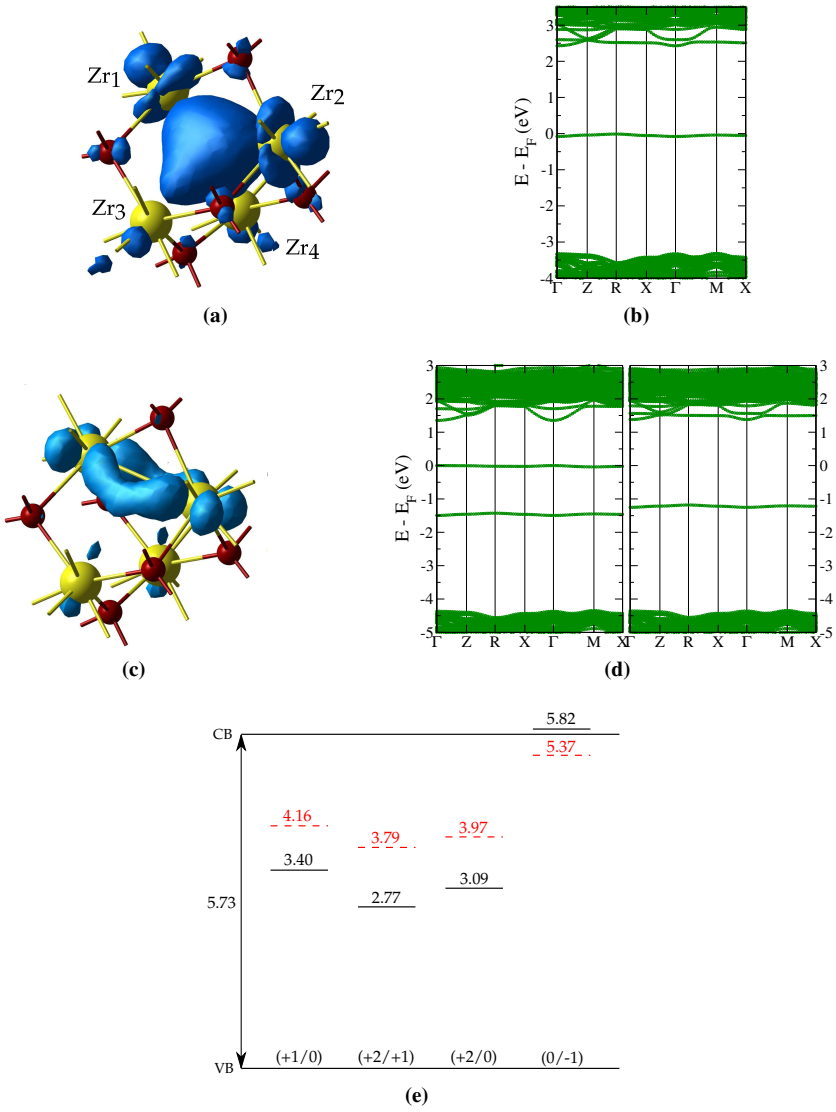

Figure 6.4: *O vacancy in tetragonal ZrO_2. Isosurface plot of the charge density associated with (a) the two excess electrons for the neutral vacancy, and (c) the trapped electron for the negatively-charged vacancy. The yellow (large) and red (small) spheres represent Zr and O atoms, respectively. Corresponding electronic band structures, computed at the equilibrium geometry, for (b) the neutral, and (d) the negatively-charged vacancy (shown for both spin majority, α, and spin minority, β). (e) Optical (solid, black) and thermodynamic (dashed, red) charge-transition levels. Positions (eV) are given with respect to the top of the VB.*

3.65 Å to 3.61 Å. The corresponding band structure [Figure 6.4(b)] exhibits a flat defect state positioned at ~ 3.3 eV above the VB, in close agreement with a previous B3LYP study [108].

As the vacancy is ionized, the remaining extra electron distributes at the two Zr_1 and Zr_2 atoms, as well as in the space between them, in a bonding-like configuration. The extra electron occupies a doublet state in the band gap at ~ 3.6 eV above the VB.

Finally, a negatively-charged O vacancy can be realized following trapping of an electron from the bulk CB. Such trapping phenomena can affect, for example, the performance of electronic devices, for which high-κ oxide materials such as zirconia or hafnia (HfO_2) have been proposed as gate electrodes [261]. The trapped electron occupies an additional defect state in the band gap at $\sim$ 1.2 eV above the (now spin-resolved, with spin majority shown) states associated with the neutral vacancy [Figure 6.4(d)], in a bonding-like configuration between the Zr_1 and the Zr_2 atoms [Figure 6.4(c)].

The computed CTLs corresponding to both vacancy ionization and electron trapping are shown in Figure 6.4(e). A band gap of 5.73 eV is calculated, to be compared with the experimental value of 5.78 eV deduced from optical absorption measurements [92]; typically GW calculations give a band gap larger by $\sim 0.1 - 0.2$ eV (see Ref. [100] and references therein).

Much experimental work is available in the literature investigating the PL properties of ZrO_{2-x}. However, a clear understanding of the underlying mechanisms seems to be lacking. Several emission features in the region between 2.0 eV and 3.5 eV are typically observed, and have been related to O deficiency [286]. The large uncertainty in experimental results may be the consequence of the practical difficulty in obtaining high-purity single crystals of zirconia with a well-defined phase.

PL measurements on tetragonal nanocrystalline ZrO_{2-x} showed an emission peak at 350 nm ($\sim$ 3.54 eV), which was attributed to electron-hole recombination involving F-center states in the band gap. Analysis of the PL spectrum evolution with annealing allowed to assign this feature to O vacancies [52]. The computed position of the $(+1/0)$ optical transition level (3.40 eV above the top of the VB) may account for this observation.

Another study on undoped tetragonal zirconia nanocrystals evidenced the presence of an emission peak at $\sim$ 2.8 eV [287]. This may be related to a PL process involving the $(+2/+1)$ optical level, which is computed to be at 2.77 eV above the top of the VB.

The thermodynamic levels are found below the CB by at least 1.5 eV, indicating that all the O-vacancy charged states are stable at room temperature. An exception is represented by the $(0/-1)$ level, which is separated from the CB by ~ 0.3 eV. This level is associated to electron trapping processes causing the O vacancy to become negatively charged. The sc-PBE0α_{ϵ_∞} calculations suggest that such trapping center is stable at room temperature, and the same conclusion was obtained from a similar investigation on hafnia [34] in which, due to analogous chemical behavior of the Hf and Zr cations, O vacancies are expected to induce similar effects as in zirconia [331]. In light of this, trapping/detrapping experiments on HfO_2 found the activation energy for the trap level to be 0.35 eV [259] which is in agreement with the sc-PBE0α_{ϵ_∞} results reported for ZrO_2.

6.4 Summary and Conclusions

The behavior of O vacancies in prototypical reducible (TiO_2, WO_3) and nonreducible (ZrO_2) wide-gap oxide materials has been investigated through the sc-PBE0α_{ϵ_∞} method as defined for the corresponding pristine material. Its capability of yielding accurate electronic band structures, together with structural properties and total energies, makes it an ideal tool for investigating defect levels in insulating materials.

The computed optical transition levels are generally found in agreement with various spectroscopy experiments, providing support for their interpretation. Instead, thermodynamic levels allow one to infer about the intrinsic degree of n-type conductivity of substoichiometric oxide materials through analysis of the stability of the different O vacancy charge states. Residual discrepancies between theory and experiment may arise from two main sources: (i) the always present approximate treatment of xc terms, and (ii) usage of finite supercells. While the former is intrinsically related to the adopted functional, the latter may appear also as consequence of the evaluation of CTLs via Janak's theorem (see Section 4.3.1) instead of by direct computation of total energy differences,[7] due to the defect band dispersion introduced by the residual polaron overlapping in neighboring cells.

Fundamental differences are evidenced in the behavior of reducible and nonreducible oxides. Upon creation of an O vacancy, the former tend to exhibit polaron-like features: the excess charge mainly localizes at the metallic ions, reducing their formal charge, and the accompanying

[7] Such differences cannot be computed with the CRYSTAL09 code [77].

relaxation of the atomic environment is not restricted to atoms in the immediate vicinity of the defect. Instead, in nonreducible oxides such as ZrO_2, the excess charge is stabilized in the vacancy void, and the relaxation is more local; furthermore, transition levels are deeper, being situated near the center of the band gap. These are the fingerprints of the F-center-like behavior of O vacancies in these materials.

In conclusion, the sc-PBE0α_{ϵ_∞} approach is able to characterize defect levels in semiconductors in a way that makes it valuable for predictive studies also on doped semiconductors. Notice however that, while the method is straightforwardly applied to point (zero-dimensional) defects,[8] its extension to treatment of higher-dimensional defect structures, such as surfaces, requires facing unprecedented difficulties. For example, modeling the interface between two different semiconductors implies the exchange fraction to become a space-dependent quantity. Local hybrid functionals, in which a portion of position-dependent HF exchange is admixed to semilocal [145] or global hybrid functionals [124], may constitute a useful starting point. However, the extension of the dielectric-constant dependent mixing approach to non-homogeneous systems will eventually call for a phenomenological model of the dielectric properties at the interface. Existing models of the surface and interface optical response [67, 68, 135] may represent a possible starting point for future developments of the approach presented in this work, although, to the best of the author's knowledge, the issue has not yet been explicitly considered in the literature. An extension of the present approach in this sense would be highly desirable, making it possible to quantitatively investigate the electronic structure at interfaces.

Notice also that for treatment of defective systems, the self-interaction error should be largely corrected for (otherwise the excess charge would tend to delocalize over an exceedingly large portion of the crystal). In this respect, the sc-PBE0α_{ϵ_∞} functional proves to behave correctly in the case of O vacancies, for which also standard hybrid functionals have proved to perform well [8, 64, 72]. An *ad hoc* test for the sc-PBE0α_{ϵ_∞} functional on a defective system where even standard hybrids fail in correctly describing charge localization (evidencing that they are still suffering, to some extent, from the self-interaction error) is considered in Chapter 7.

[8] In the dilute limit, the dielectric constant is vanishingly affected by presence of the defects.

Computational Details

All the calculations were performed using the CRYSTAL09 code [77,78]. In the case of TiO_2, calculations were performed in the all-electron scheme, using the basis set from Ref. [333] for Ti. Small-core effective-core PPs were employed for modeling core electrons for W and Zr atoms [125] in WO_3 and ZrO_2, while valence electrons were described using the basis sets defined in Ref. [315] ($5s$, $5p$, $6s$ and $5d$ W electrons in the valence) and Ref. [30] ($4s$, $4p$, $5s$ and $4d$ Zr electrons in the valence), respectively. The O atom was always treated at the all-electron level, using the basis set from Ref. [263] for TiO_2 and WO_3, and from Ref. [96] for ZrO_2.

Bulk 72-atom $2 \times 2 \times 3$ and 96-atom $2\sqrt{2} \times 2\sqrt{2} \times 2$ supercells were considered for rutile and anatase TiO_2, respectively. For γ-monoclinic WO_3, a model supercell comprising 64 atoms was adopted, obtained by doubling the primitive cell along the a crystallographic axis. For tetragonal ZrO_2, a 108-atom $2 \times 2 \times 3$ supercell was employed. The O vacancy was modeled by removing one O atom from the corresponding supercells. The BZ was sampled using $2 \times 2 \times 2$ Γ-centered Monkhorst-Pack grids [213], corresponding to 6 or 8 k points in the IBZ.

Structural optimizations were performed by allowing all the atoms in the cell to relax their positions, keeping the lattice parameters fixed to that of the optimized bulk cell. Convergence thresholds in geometry optimizations were set at their standard values in the CRYSTAL09 code: the defined thresholds for the maximum and the root-mean-square of the energy gradients (atomic displacements) are 0.00045 a.u. (0.00180 a.u.) and 0.00030 a.u. (0.00120 a.u.), respectively [77].

The self-consistent procedure defining the sc-PBE0α_{ϵ_∞} functional yielded the following values for the exchange fraction: 15.2% for rutile and 18.4% for anatase TiO_2, 21.9% for RT monoclinic WO_3, 20.8% for tetragonal ZrO_2.

The $1s$ KS eigenvalue of the O atom was considered as reference for aligning band structures in bulk calculations. The spurious electrostatic interaction between image charged defects was accounted for by correcting the KS eigenvalues according to the procedure outlined in Section 4.2.1. In particular, the purely geometrical prefactor $(\alpha_{\mathrm{M}}/2L)$ in Eq. (4.21) was evaluated by computing the nuclear-nuclear interaction energy of a periodic system consisting of hydrogen atoms positioned at the defect sites [14] (for noncubic cells, L is thus the average distance between defects). For the dielectric constant ϵ_∞ in the same formula, the self-consistent value obtained within sc-PBE0α_{ϵ_∞} was used.

The O vacancy formation energy was computed with respect to $(1/2)E[\mathrm{O}_2]$, where $E[\mathrm{O}_2]$ is the ground-state total energy of the triplet O_2 molecule. This corresponds to considering the formation of the vacancy under O-rich conditions.

CHAPTER 7

Hole Localization in Al-Doped SiO_2

One of the most important requirements to be met for a DFT functional is that the self-interaction (SI) error is largely compensated. In fact, SI leads to prediction of even qualitatively wrong ground states in systems in which charge density would tend to be highly localized, favoring instead a ground state in which the charge is spread over a region as large as possible, in an attempt to minimize the SI present in the Hartree energy. Point defects in insulators constitute one paradigmatic condensed-matter system in which this happens. In particular, aluminum-doped silicon dioxide (Al-doped SiO_2) is known to be an especially difficult system to be treated, since even hybrid functionals, which, to a large extent, are expected not to suffer from SI, have proved to fail in properly describing the ground state of this system.

The Al impurity is one of the most commonly observed defects in irradiated quartz SiO_2, by which a tetravalent Si cation is replaced with a trivalent Al atom. The Al/Si substitution results in an unpaired electron (or, in an equivalent description, a hole) which, based on early experimental observations, has been identified to be trapped in a nonbonding $2p$ orbital of an O atom surrounding the substitutional Al

atom [115, 225]. The corresponding neutral Al defect center (hereafter also denoted $[AlO_4]^0$) is magnetically active, and has been the subject of extensive characterization by EPR spectroscopy [223, 224, 269], also in combination with absorption spectroscopy [207, 218, 268].

From the theory side, it has been recognized that reproduction of the experimentally observed features of the $[AlO_4]^0$ center is subordinate to a correct description of the hole localization properties. Since early DFT calculations using local or semilocal density functionals yielded a wrong picture, predicting the hole to be delocalized over all the four Al nearest-neighbor O atoms [53,171,193], this problem has been identified as a challenging testing ground for novel density-functional methods [172].

The wrong description provided by local/semilocal DFT functionals has been ascribed to incomplete cancellation of the SI brought in by the Hartree term [171, 234], as calculations using exactly SI-free Hamiltonians, such as unrestricted HF (UHF) and self-interaction-corrected (SIC) DFT, yield the correct hole localization. Pacchioni *et al.* also performed unrestricted second-order Møller-Plesset perturbation theory (UMP2) calculations [234], concluding that no appreciable role is played by correlation in the specific problem at hand.

Furthermore, it has been found out that popular hybrid functionals, prescribing 20% or 25% of EXX, such as B3LYP [20, 292] PBE0 [244] and HSE06 [132, 133] also fail in describing charge localization in Al-doped silica [105, 172, 230, 288, 334]. Thus, it has been argued that a large amount of EXX would be necessary to obtain agreement with experiments [172, 288], much in the same way as a large enough Hubbard U parameter is needed within DFT + U to solve the same problem [222]. For example, To *et al.* found that a semiempirical hybrid functional including 42% of EXX (called BB1K functional) yielded the correct picture [301].

In light of the preceding work, it may be concluded that some amount of empiricism is inevitably required to tackle the Al impurity problem within DFT, casting doubts on its actual predictive power. In view of this, the sc-PBE0α_{ϵ_∞}approach is very promising: in fact, it can be regarded as a self-consistent DFT implementation of the many-body COHSEX self-energy. COHSEX treats the exchange term exactly, and hence it is SI-free; it thus constitutes the ideal candidate for modeling systems whose ground state is dominated by the classical Hartree and electron exchange Fock interactions. It is worthwhile, and the ultimate goal of the present investigation, to assess whether the directly derived

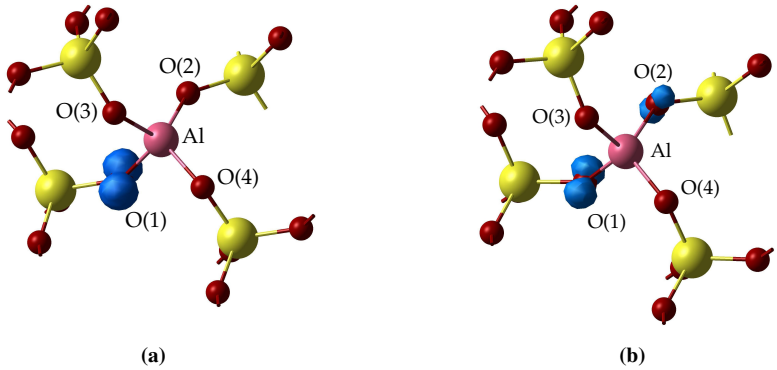

Figure 7.1: *Ball-and-stick representation (Si, O and Al atoms are shown as yellow, red, and pink spheres, respectively) of the local atomic structure around the Al impurity, as obtained from geometry optimization at the (a) sc-PBE0α_{ϵ_∞}, and (b) B3LYP level of theory. Isosurface of the spin density associated with the unpaired electron introduced by the $[AlO_4]^0$ center is shown. For bond distances see Table 7.1.*

density-functional method exhibits the same property.

7.1 sc-PBE0α_{ϵ_∞} Results for Bulk SiO_2

Within the sc-PBE0α_{ϵ_∞} approach, the dielectric constant of quartz SiO_2 is computed to be $\epsilon_\infty = 2.15$ (the experimental value is 2.38, see Ref. [42] and references therein), which corresponds to an exchange fraction $\alpha = 46.5\%$. The method predicts a somewhat overestimated band gap of 11.6 eV, whereas various experiments measured it in the quite broad range of $8 - 10$ eV [319]. The failure of the sc-PBE0α_{ϵ_∞} functional in computing reliable band gaps for some insulators with very low dielectric constants (and correspondingly large band gaps) was already reported in Ref. [283], being particularly serious when the geometry is re-optimized at each self-consistency step (see for example the case of MgO, which has a dielectric constant of ~ 3, see Chapter 5).

However, here the main concern is to pursue the correct description of the hole localization at the $[AlO_4]^0$ center, which is a ground state property; the related spectroscopic features will be deferred to separate discussion in Section 7.4.

Table 7.1: *Nearest-neighbor Si-O and Al-O distances (Å) for pristine bulk and Al-doped SiO_2 ($[AlO_4]^0$), respectively. The O atoms are labeled according to Figure 7.1.*

	Pure SiO_2	$[AlO_4]^0$		
Functional	sc-PBE0α_{ϵ_∞}	sc-PBE0α_{ϵ_∞}	B3LYP[a]	B3LYP[b]
O(1)	1.620	1.910	1.809	1.826
O(2)	1.620	1.699	1.758	1.749
O(3)	1.616	1.687	1.705	1.705
O(4)	1.616	1.691	1.700	1.700

[a] Starting geometry for optimization: pure SiO_2 structure.
[b] Starting geometry for optimization: sc-PBE0α_{ϵ_∞}-optimized $[AlO_4]^0$ structure.

7.2 Structural Deformation and Hole Localization

Figure 7.1(a) shows the local atomic structure of SiO_2 around the Al impurity, as found by minimizing the total electronic energy using the sc-PBE0α_{ϵ_∞} functional and allowing for symmetry-breaking atomic relaxations, which amounts at independently optimizing the positions of each atom of the (O_1, O_2) and (O_3, O_4) oxygen pairs (where the two O atoms within each pair are equivalent to each other in the bulk SiO_2 structure). The emerging picture agrees very well with that obtained within other rigorously SI-free approaches, such as UHF [172, 230] and SIC-DFT [59]. In particular, the hole introduced by the substitutional Al atom is found to be trapped at the O(1) atom, and its wavefunction exhibits purely $2p$ character, with the corresponding orbital lying almost perpendicularly to the Al-O-Si plane. As a consequence of charge localization, the local atomic structure distorts considerably: the O(1) atom moves away from the Al atom, resulting in an average equilibrium Al-O distance 13% larger compared to the other Al-O distances (see also Table 7.1).

7.3 Analysis of EPR Parameters

In order to further confirm the better performance of the sc-PBE0α_{ϵ_∞} approach with respect to B3LYP, hyperfine parameters relative to the hole-bearing O(1) atom are computed. The hyperfine coupling matrix $\mathbf{A}$, describing the magnetic interaction of the spin of an unpaired electron with the spin of the neighboring nuclei (^{17}O and ^{27}Al isotopes), is conveniently divided into an isotropic (spherically

Table 7.2: *Spin population of the O atoms belonging to the $[AlO_4]^0$ center, and EPR hyperfine parameters of the hole-bearing ^{17}O and of the ^{27}Al. The optimized structures are obtained starting from ideal bulk SiO_2.*

Functional Geometry	sc-PBE0α_{ϵ_∞} sc-PBE0α_{ϵ_∞}	B3LYP B3LYP	B3LYP sc-PBE0α_{ϵ_∞}	Expt. (Refs. [223, 224])
		Spin population		
O(1)	0.95	0.58	0.81	
O(2)	0.02	0.28	0.07	
O(3)	< 0.01	0.01	0.03	
O(4)	< 0.01	0.04	< 0.01	
		^{17}O(1) hyperfine matrix (G)[a]		
A_{iso}	-42.6	-26.1	-34.7	-26.0
B_1	-94.3	-61.7	-83.7	-85.0
B_2	47.1	30.7	41.8	41.2
B_3	47.2	31.0	41.9	43.8
		^{27}Al hyperfine matrix (G)[a]		
A_{iso}	-5.0	-8.1	-5.4	-5.8
B_1	-0.4	-0.2	-0.4	-0.4
B_2	-0.4	-0.1	-0.3	-0.3
B_3	0.8	0.3	0.6	0.7

[a] The principal values of the anisotropic hyperfine matrix are listed so that $B_1 < B_2 < B_3$

symmetric) and an anisotropic (dipolar) part [318],

$$\mathbf{A} = A_{\text{iso}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} B_1 & 0 & 0 \\ 0 & B_2 & 0 \\ 0 & 0 & B_3 \end{pmatrix}. \tag{7.1}$$

The isotropic part A_{iso} is proportional to the electron spin density at the nucleus, and, as such, the dominant contribution to it is caused by spin-polarization of the s electrons. Instead, the anisotropic part is related to the spin population of orbitals with higher angular momentum components; this contribution is commonly expressed in terms of a traceless matrix with principal values B_1, B_2 and B_3 reported in Table 7.2. Since the hole wavefunction has purely $2p$ character, the dipolar part gives direct access to the corresponding spin distribution. Instead, the isotropic contribution is notably harder to be reproduced, being extremely sensitive to details of the calculation in general, and to

choice of the basis set in particular [15].

In Table 7.2 the hyperfine parameters are reported for the ^{17}O(1) and ^{27}Al nuclei. Concerning the anisotropic parameters of ^{17}O, sc-PBE0α_{ϵ_∞} calculations nicely capture the experimentally observed strong anisotropy along the three axes, and numerical values are in good agreement with both experiments and the results of previous investigations based on SI-free approaches (UHF, UMP2, SIC-DFT) [59, 172, 230]. In contrast, B3LYP yields a quantitatively wrong picture, the computed parameters being substantially smaller in absolute value than experimental ones. The situation is quite the opposite for the isotropic part, for which sc-PBE0α_{ϵ_∞} overestimates the absolute value of A_{iso}, while B3LYP yields it exceptionally close to experiment. However, the latter result can be considered as fortuitous, in the sense that it is not concomitant with a correspondingly more accurate description of the hole localization. Firstly, as already discussed, at the B3LYP level the O(2) atom carries a substantial part of the hole-related spin density (see Table 7.2), in disagreement with experiment, and consequently the ^{17}O(2) EPR parameters are of the same order of magnitude as for ^{17}O(1);[1] instead, when the hole localization is correctly captured, such as at the sc-PBE0α_{ϵ_∞} level, the former are at least one order of magnitude smaller than the latter (see also Ref. [301]). Secondly, the improvement of the isotropic part does not come along with a similar improvement of the anisotropic one, which indeed is related to the proper description of the $2p$ hole wavefunction.

The above conclusion is also supported by the computed superhyperfine matrix of ^{27}Al: the sc-PBE0α_{ϵ_∞} functional gives results in quantitative agreement with both experiment and UHF calculations [230], whereas this is not true for B3LYP. Notice that the superhyperfine interaction with a dopant element like Al that introduces a hole in the structure is usually the only accessible information. In fact, in order to measure the O hyperfine constants, ^{17}O-enriched samples have to be prepared with complex and costly procedures [107].

7.4 Optical properties of the Al impurity

The Al impurity in quartz silica has been observed to act as a color center, endowing this material with the typical smoky coloration. However, considerable controversy has arisen as to which absorption

[1] The ^{17}O(2) EPR parameters, computed at the sc-PBE0α_{ϵ_∞}(B3LYP) level are (in G): $A_{\text{iso}} = -3.6\,(-15.3)$, $B_1 = -2.9\,(-31.7)$, $B_2 = 1.4\,(15.7)$, $B_3 = 1.5\,(16.0)$.

Table 7.3: *Vertical excitation energies (eV) associated to the $[AlO_4]^0$ center computed with the sc-PBE0α_{ϵ_∞} and B3LYP functionals at different optimized geometries. Comparison with theoretical results from the literature obtained within time-dependent DFT (TDDFT) using the B3LYP and BB1K exchange-correlation approximations, as well as within the many-body outer-valence Green's function (OVGF) approach. The experimental position of the main absorption peak is also reported.*

	This work					
Functional	sc-PBE0α_{ϵ_∞}	B3LYP	TDDFT-BB1K	TDDFT-B3LYP	OVGF	Expt.
Geometry	sc-PBE0α_{ϵ_∞}	sc-PBE0α_{ϵ_∞}	(Ref. [301])	(Ref. [334])	(Ref. [334])	
	4.73	2.91	3.03	1.72	2.74	2.9 (Ref. [218]), 2.85 (Ref. [207])

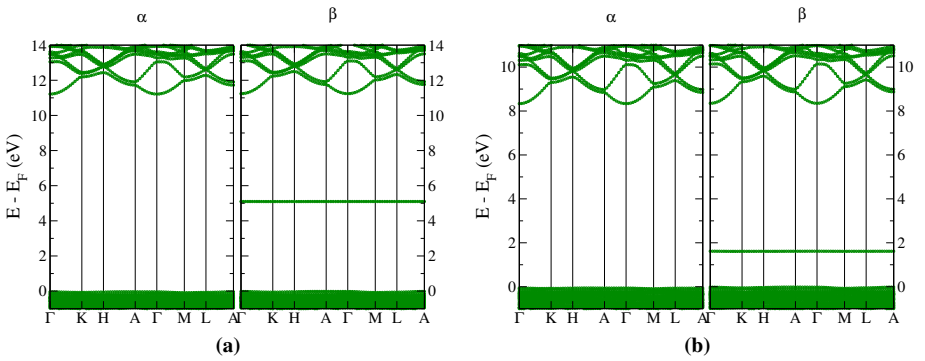

Figure 7.2: *Electronic band structure of SiO_2 with a neutral Al impurity, computed at the (a) sc-PBE0α_{ϵ_∞}, and (b) B3LYP level.*

feature had to be correlated with such observation [218, 268]. It was finally concluded that an absorption peak at about 2.9 eV is to be associated with the presence of Al centers and, thus, with the smoky coloring [207].

The observed optical transition should be related to excitation of the hole trapped at the neutral $[AlO_4]^0$ center ($q = 0$) into the VB, leading to a negatively charged defect ($q' = -1$, $[AlO_4]^{-1}$ center); accordingly, from the theory side, the optical transition level $(0/-1)$ is the relevant quantity to be compared with experiments. The sc-PBE0α_{ϵ_∞} predicts the computed level to be ~ 4.7 eV above the VB maximum, nearly 2 eV higher in the band gap with respect to experiments. This large error can be attributed to the already mentioned overestimation of the electronic band gap of bulk quartz provided by sc-PBE0α_{ϵ_∞}. Consequently, the Al-related defect level is wrongly positioned with respect to the VB [~ 5 eV above its edge, see Fig. 7.3(a)] and this eventually gives rise to the observed overestimation of the optical transition energy.

As a partial workaround, the electronic structure is computed using the B3LYP functional, which yields a band gap of 8.6 eV for bulk quartz [see Fig. 7.2(b)], falling in the range of the experimental values; the $[AlO_4]^0$ geometry obtained within sc-PBE0α_{ϵ_∞} is instead retained. The spin population and EPR parameters computed following this approach are reported in Table 7.2. The hole is again localized on the O(1) atom, although with a Mulliken density lower than that obtained by performing calculations fully within the sc-PBE0α_{ϵ_∞} scheme. Surprisingly, the computed hyperfine parameters are in even better quantitative agreement with experiments, as the lower spin density counterbalances the

overestimation yielded by sc-PBE0α_{ϵ_∞}; a similar trend was noticed in the previous hybrid-functional investigation of To *et al.* [301]. Charge density distribution analysis for the negatively-charged center leads to conclude that the same qualitative picture is obtained at the sc-PBE0$\alpha_{\epsilon_\infty}$and B3LYP levels, provided that the geometry is kept fixed at the one optimized within sc-PBE0α_{ϵ_∞}: the one-particle state corresponding to vertical excitation of the hole to the VB is still contributed by the $2p$ orbitals of the O(1) atom. The computed optical level for such transition is positioned at $\sim$ 2.9 eV above the top of the VB, in excellent agreement with experiment. For comparison, Table 7.3 reports representative results from previous theoretical studies for the vertical transition energy correlating with the experimentally found absorption band with a maximum at ~ 2.9 eV.

7.5 Summary and Conclusions

The long-standing problem of hole localization in Al-doped quartz SiO_2 is an ideal testing ground for novel DFT functionals, since the obtained ground state for this system turns out to be extremely sensitive to the SI error: local/semilocal and standard hybrid DFT functionals fail in capturing the experimentally evidenced hole localization at one of the Al-coordinated O atoms [172, 230]. In particular, the failure of popular hybrid functionals, such as B3LYP, has been attributed to the insufficient amount of EXX admixed. In the sc-PBE0α_{ϵ_∞}functional, the exchange fraction is consistently determined based on the analogy with the many-body COHSEX approximation to the electron self-energy. The COHSEX self-energy is rigorously SI-free, and thus constitutes the ideal starting point for studying systems in which incomplete cancellation of SI leads to a qualitatively wrong ground state.

For quartz silica, the sc-PBE0α_{ϵ_∞}approach yields an exchange fraction of $\sim 46\%$. The resulting hybrid functional correctly reproduces the hole localization at a single O atom nearest-neighbor to the Al impurity, also giving accurate description of the structural distortion around it, and allowing to compute EPR parameters in agreement with previous SI-free calculations [172, 230, 301] and experiment [224]. However, the defect-related optical spectroscopic features are not well-reproduced by the sc-PBE0α_{ϵ_∞}method. This failure is attributed to overestimation of the pristine quartz silica band gap. Using B3LYP on top of the sc-PBE0α_{ϵ_∞}optimized geometry yields a band structure in better agreement with experiment, and thus corroborates this hypothesis.

In conclusion, the sc-PBE0α_{ϵ_∞} approach proves capable of correcting most of the SI error inherent to local/semilocal, as well as more popular hybrid, DFT functionals. This feature is crucial for adequately describing the ground state of defective oxide materials. As far as defect-related excitation energies are concerned, their determination is subordinate to an accurate calculation of the electronic structure of the bulk material. This is not always the case for low-dielectric-constant insulators in general, and for quartz SiO_2 in particular. However, as shown in Chapter 6, the sc-PBE0α_{ϵ_∞} method is able to reproduce the whole experimental scenario when point defects in moderate gap metal-oxide semiconductors (dielectric constants $\sim 4-6$) are addressed.

Computational Details

All-electron DFT calculations were performed using the CRYSTAL09 code [77, 78]. The following all-electron basis sets were employed: 66-21G* for Si (Ref. [217]), 8-411(d1) for O (Ref. [263]), 85-11G* for Al (Ref. [39]).

The Al center was modeled in an embedding $2 \times 2 \times 2$ quartz SiO_2 supercell (72 atoms) with the atomic positions and lattice parameters fully relaxed for the bulk cell using the sc-PBE0α_{ϵ_∞} functional. For the defective supercell, further optimization of the atomic positions was carried out at fixed lattice parameters. The standard thresholds defined in CRYSTAL09 were adopted in all geometry optimizations: the defined thresholds for the maximum and the root-mean-square of the energy gradients (atomic displacements) are 0.00045 a.u. (0.00180 a.u.) and 0.00030 a.u. (0.00120 a.u.), respectively [77]. The BZ was sampled using $2 \times 2 \times 2$ Γ-centered Monkhorst-Pack grids [213], corresponding to 8 k-points in the IBZ.

The $1s$ KS eigenvalue of Si is taken as reference for aligning band structures in defect and bulk calculations. The spurious electrostatic interaction between image charged defects was accounted for by correcting the KS eigenvalues according to the procedure illustrated in Section 4.2.1 and in the Computational Details of Chapter 6.

CHAPTER 8

Conclusions

In this thesis, a nonempirical hybrid DFT method is tested for calculation of both ground-state and excited-state properties of pristine and defective oxide materials. The method is based on a material-dependent re-parametrization of the full-range PBE0 functional [244], in the sense that the exchange fraction α is obtained as the inverse of the high-frequency dielectric constant of the material [6, 198]. This identification arises straightforwardly from comparison of the expressions for the nonlocal hybrid DFT potential and the electronic self-energy, as given in the static limit of the GW approximation (known as COHSEX self-energy) [126]. Actually, in this framework, the full dielectric response function, which in principle is a spatial-dependent quantity, has to be averaged over the simulation cell (i.e. the macroscopic limit has to be taken), in order to derive an explicit expression for the constant α in terms of the macroscopic dielectric constant ϵ_∞. As far as ϵ_∞ is computed with *ab initio* approaches, the obtained functional can be considered as nonempirical, in the sense that the exchange fraction built in it is not set *a priori* or on the basis of fitting to experiments, at odds with standard hybrid functional implementations. Moreover, the method can be made

self-consistent [283], as explained in Section 3.3, thus defining the sc-PBE0α_{ϵ_∞} functional.

In Chapter 5, the sc-PBE0α_{ϵ_∞} method has been applied to computation of band gaps, dielectric constants, and ground-state properties of a class of wide-gap oxide materials (MgO, ZnO, TiO_2, ZrO_2, WO_3). For band gaps, comparison has been provided with results from GW calculations, as well as experiments. In this respect, the sc-PBE0α_{ϵ_∞} turns out to be roughly as accurate as GW, also being less expensive than the latter. The self-consistent procedure is shown to converge within two or three iteration steps for most materials, and yields dielectric constants closer to experiments with respect to the starting PBE functional. Then, the phase stability and related electronic properties of different crystallographic phases for the same set of oxides have been analyzed. The self-consistent exchange fraction α has been evaluated for each polymorph separately. The phase stability has been studied on the basis of differences in ground-state total energy, thus neglecting thermal contributions. It has been concluded that generally sc-PBE0α_{ϵ_∞} does not produce a clear improvement over standard hybrid functionals like PBE0 concerning the phase stability issue. Actually, it performs worse than PBE0 when the structural phase transition comes along with a drastic change of the electronic structure, as in the case of WO_3. Here, the low-temperature cubic and tetragonal phases exhibit a much lower band gap than the high-temperature ones. Correspondingly, the dielectric constant varies by a large amount passing, e.g., from the cubic to the monoclinic structure. This produces a large variation of the exchange fraction upon the phase transition, and ultimately leads to prediction of a wrong phase stability order. Another test on total energies is performed, calculating the reaction energy associated to a stoichiometry transformation of bulk TiO_2 into Ti_2O_3. In this case, more sizable energy differences of the order of few eV are involved. Hence, the sc-PBE0α_{ϵ_∞} method can be better tested in its capability to give accurate total energies, irrespective of numerical errors inevitably introduced by convergence of the basis set, of the BZ sampling, etc. (which in fact can produce errors on the scale of meV, which is comparable to typical differences in ground-state energies between two crystallographic phases). The computed reaction energy using sc-PBE0α_{ϵ_∞} results in better agreement with experiments with respect to standard PBE0. Furthermore, the method consistently provides a good description of electronic structures and structural properties of both the TiO_2 and Ti_2O_3 compounds, for which, e.g., DFT + U methodologies

proved to be less accurate [139].

Passing from pristine to defective oxides, Chapter 6 has been devoted to study of oxygen vacancies in rutile and anatase TiO_2, γ-monoclinic WO_3, and tetragonal ZrO_2. This set of materials comprises both reducible (TiO_2, WO_3), and nonreducible (ZrO_2) oxides, which in general exhibit qualitatively different electrical and optical behavior when becoming substoichiometric [233]. The sc-PBE0α_{ϵ_∞} functional is a promising method for *ab initio* modeling of point defects, for the following reasons: (i) exact exchange helps selecting ground states with the correct degree of charge localization, at odds with local/semilocal functionals which instead tend to artificially delocalize it; (ii) already the electronic structure and energetics of pristine wide-gap oxides are well described, as discussed in Chapter 5, and thus defect levels are expected to be correctly positioned in the band gap even within the KS scheme. Defect levels have been computed using the charge-transition levels formalism, which goes beyond a simple KS description, as discussed in Section 4.3. The sc-PBE0α_{ϵ_∞} method proves able to capture the different behavior of O vacancies in rutile and anatase TiO_2, accounting for the higher n-type conductivity found in the latter (explained by the shallower thermodynamic transition levels), as well as the nature of the optical transitions observed in both phases (deep optical transition levels). Spatial distribution of the excess charge is also found in qualitative agreement with STM measurements on both phases. Optical transitions computed for WO_3 and ZrO_2 are also found to be consistent with experiments. ZrO_2 is correctly predicted to be less prone to form O vacancies with respect to the other oxides [98], as the related formation energy is found larger by $\sim$ 1 eV. Thermodynamic levels are found to be much deeper in this material, suggesting that the O vacancy is stable in its neutral state even at high temperature. Overall the sc-PBE0α_{ϵ_∞}, employed in conjunction with the charge-transition level scheme, proves able to describe O-deficient wide-gap metal oxides in terms of their electrical and optical properties, as well as of spatial distribution of the excess charge and O-vacancy formation energy.

In Chapter 7 the description of Al impurities in quartz SiO_2 has been tackled. The choice of this peculiar defective system has been guided by the demonstrated inability of popular hybrid functionals like B3LYP to correctly capture localization of the unpaired electron introduced by substitutional aluminum [172, 234]. Since HF instead has been reported to give correct results, the failure of B3LYP has been ascribed to incomplete cancellation of the self-interaction error, and later studies

have suggested that an exchange fraction as large as $40 - 50\%$ is needed to construct a self-interaction-free hybrid functional [301]. For the sc-PBE0α_{ϵ_∞} functional, obtained for pristine SiO_2, the exchange fraction is about 46%; when applied to defective SiO_2, this functional predicts the hole to be localized on one of the four O atoms nearest neighbor to the Al impurity; localization is accompanied by considerable structural distortion of the local structure, in agreement with the experimental evidence from EPR spectroscopy [224]. Calculated EPR hyperfine parameters for O and Al atoms confirm that sc-PBE0α_{ϵ_∞} is superior to B3LYP, since it better describes the symmetry-breaking structural distortion around the Al atom, which drives localization of the hole on the O atom moving farthest away from the defect. However, sc-PBE0α_{ϵ_∞} proves not suitable to calculation of optical transitions induced by presence of the defect in this material. In fact, already in pristine SiO_2, the band gap is overestimated by as much as 2 eV. This leads to wrong prediction of the defect level position in Al-doped SiO_2, and hence of the associated transition level. Overestimation of the band gap in insulators with low dielectric constants has to be regarded as an intrinsic limitation of the method, which tends to yield exchange fractions approaching 100% when $\epsilon_\infty \to 1$, resulting in band gaps close to those that would be obtained in a HF calculation.

More generally, while this work has clearly demonstrated that the sc-PBE0α_{ϵ_∞} functional represents an improvement, both from the conceptual and the practical point of view, with respect to standard hybrid DFT approaches, it is also clear that there are still open issues in electronic structure theory that this method cannot aim at solving. For example, calculation of the electronic structure of surfaces and interfaces, which holds great importance for photovoltaics and photocatalysis, could not be straightforwardly tackled. Typically, the dielectric constant varies abruptly across the surface, and this makes choice of a global exchange fraction in the hybrid functional an ambiguous one. This difficulty is not present when point defects are investigated, since for these systems the dielectric constant is expected to converge to the bulk value in the limit of dilute defect concentrations which is usually the most relevant to comparison with experiments. Instead, the dielectric response function is substantially modified in the presence of surfaces, and thus in principle a full evaluation of this quantity should be performed using more rigorous many-body perturbation methods. However, even in cases in which the sc-PBE0α_{ϵ_∞} approach cannot be directly applied, starting from an

accurate electronic structure method for the bulk material remains advantageous. In the case of defects at surfaces, for example, an even qualitatively correct description of defect wavefunctions would be crucial to obtain meaningful GW quasiparticle corrections to the KS band structure. Hence, while one should not expect sc-PBE0α_{ϵ_∞} to provide electronic properties of surfaces and interfaces in quantitative agreement with experiment, this approach could be still the ideal starting point for higher-level methods which are naturally able to treat spatially inhomogeneous systems.

List of Figures

3.1 Flowchart of the sc-PBE0α_{ϵ_∞} method. 61

4.1 Schematic representation of electronic transitions involving a neutral defect ($q = 0$) and a positively ($q = +1$) and negatively ($q = -1$) charged one. 72

5.1 Variation of total energy as a function of the exchange fraction α entering the definition of the dielectric-dependent hybrid functional. 86
5.2 Bulk primitive cell of corundum Ti_2O_3. 90

6.1 O vacancy in rutile TiO_2. 103
6.2 O vacancy in anatase TiO_2. 106
6.3 O vacancy in γ-monoclinic WO_3. 109
6.4 O vacancy in tetragonal ZrO_2. 112

7.1 Ball-and-stick representation of the local atomic structure around the Al impurity in SiO_2. 119
7.2 Electronic band structure of SiO_2 with a neutral Al impurity. 124

List of Tables

5.1 Fundamental band gaps for the various materials computed within local/semilocal, standard hybrid, and G_0W_0 schemes, comparing PW-PP and LCAO schemes. 80
5.2 Band gap energy computed using dielectric-dependent hybrid functionals. 81
5.3 Electronic dielectric constant ϵ_∞ and corresponding exchange fraction α evaluated at the PBE and PBE0 levels. 82
5.4 Electronic dielectric constant ϵ_∞ and corresponding exchange fraction $\alpha = 1/\epsilon_\infty$ evaluated within PBE and dielectric-dependent PBE0 (PBE0α_{PBE} and sc-PBE0α_{ϵ_∞}) for various polymorphs of the studied materials. 84
5.5 Differences in total energy per formula unit with respect to a chosen crystallographic phase, for various polymorphs of the studied materials, computed at different levels of theory. 85
5.6 Band gap energy computed within PBE and dielectric-dependent PBE0 for various polymorphs of the studied materials. 88
5.7 Optimized cell parameters for selected polymorphs of the studied materials, computed at different levels of theory. 90
5.8 Relative total energies E_{tot} per unit Ti_2O_3, band gap E_{g} and lattice constants a and c, computed at the sc-PBE0α_{ϵ_∞} level. 91

5.9 Reaction energies for the reduction of bulk TiO_2 , computed with different hybrid functionals, and compared with experimental heats of reaction. 93

6.1 Formation energy for the O vacancy in the different materials. 111

7.1 Nearest-neighbor Si-O and Al-O distances for pristine bulk and Al-doped SiO_2. 120
7.2 Spin population of the O atoms belonging to the $[AlO_4]^0$ center, and EPR hyperfine parameters of the hole-bearing ^{17}O and of the ^{27}Al. 121
7.3 Vertical excitation energies associated to the $[AlO_4]^0$ center computed with the sc-PBE0α_{ϵ_∞} and B3LYP functionals at different optimized geometries. . . 123

Bibliography

[1] S. C. Abrahams. *Phys. Rev. B*, 130:2230, 1963.

[2] C. Adamo and V. Barone. *J. Chem. Phys.*, 110:6158, 1999.

[3] S. L. Adler. *Phys. Rev.*, 126:413, 1962.

[4] A. Alkauskas, P. Broqvist, F. Devynck, and A. Pasquarello. *Phys. Rev. Lett.*, 101:106802, 2008.

[5] A. Alkauskas, P. Broqvist, and A. Pasquarello. *Phys. Rev. Lett.*, 101:046405, 2008.

[6] A. Alkauskas, P. Broqvist, and A. Pasquarello. *Phys. Status Solidi B*, 248:775, 2011.

[7] A. Alkauskas and A. Pasquarello. *Physica B*, 401:670, 2007.

[8] A. Alkauskas and A. Pasquarello. *Phys. Rev. B*, 84:125206, 2011.

[9] C.-O. Almbladh and U. von Barth. *Phys. Rev. B*, 31:3231, 1985.

[10] E. Antoncik. *J. Phys. Chem. Solids*, 10:314, 1959.

[11] F. Aryasetiawan and O. Gunnarsson. *Rep. Prog. Phys.*, 61:237, 1998.

[12] N. Ashkenov, B. N. Mbenkum, C. Bundesmann, V. Riede, M. Lorenz, D. Spemann, E. M. Kaidashev, A. Kasic, M. Schubert, M. Grundmann, G. Wagner, H. Neumann, V. Darakchieva, H. Arwin, and B. Monemar. *J. Appl. Phys.*, 93:126, 2003.

[13] G. B. Bachelet, D. R. Hamann, and M. Schlüter. *Phys. Rev. B*, 26:4199, 1982.

[14] C. L. Bailey, L. Liborio, G. Mallia, S. Tomić, and N. M. Harrison. *J. Phys.: Conf. Series*, 242:012004, 2010.

[15] V. Barone. In D. P. Chong, editor, *Recent Advances in Density Functional Methods, Part I*, chapter 8, pages 287–334. World Scientific Publ. Co., 1996.

[16] F. Bassani, G. Iadonisi, and B. Preziosi. *Rep. Prog. Phys.*, 37:1099, 1974.

[17] C. Bechinger, M. S. Burdis, and J.-G. Zhang. *Solid State Commun.*, 101:753, 1997.

[18] A. D. Becke. *Phys. Rev. A*, 38:3098, 1988.

[19] A. D. Becke. *J. Chem. Phys.*, 98:1372, 1993.

[20] A. D. Becke. *J. Chem. Phys.*, 98:5648, 1993.

[21] A. Beltrán, J. R. Sambrano, M. Calatayud, F. R. Sensato, and J. Andres. *Surf. Sci.*, 490:116, 2001.

[22] E. Bersch, S. Rangan, R. Bartynski, E. Garfunkel, and E. Vescovo. *Phys. Rev. B*, 78:085114, 2008.

[23] F. Bloch. *Z. Phys.*, 52:555, 1928.

[24] P. E. Blöchl. *Phys. Rev. B*, 41:5414, 1990.

[25] P. E. Blöchl. *Phys. Rev. B*, 50:17953, 1994.

[26] M. Born and J. R. Oppenheimer. *Ann. Phys.*, 84:457, 1927.

[27] S. F. Boys. *Proc. R. Soc. Lond. A*, 200:542, 1950.

[28] F. D. Brandão, M. V. B. Pinheiro, G. M. Ribeiro, G. Medeiros-Ribeiro, and K. Krambrock. *Phys. Rev. B*, 80:235204, 2009.

[29] T. Bredow and A. R. Gerson. *Phys. Rev. B*, 61:5194, 2000.

[30] T. Bredow and M. Lerch. *Z. Anorg. Allg. Chem.*, 630:2262, 2004.

[31] R. D. Brigans, H. Höchst, and H. R. Shanks. *Phys. Rev. B*, 24:3481, 1981.

[32] R. D. Bringans, H. Hochst, and H. R. Shanks. *Phys. Rev. B*, 24:3481, 1981.

[33] P. Broqvist, A. Alkauskas, and A. Pasquarello. *Phys. Status Solidi A*, 207:270, 2010.

[34] P. Broqvist and A. Pasquarello. *Appl. Phys. Lett.*, 89:262904, 2006.

[35] F. Bruneval and M. A. L. Marques. *J. Chem. Theory Comput.*, 9:324, 2013.

[36] M. Calatayud, P. Mori-Sánchez, A. Beltrán, A. Martín-Pendás, E. Francisco, J. Andrés, and J. M. Recio. *Phys. Rev. B*, 64:184113, 2001.

[37] M. Cardona and M. L. W. Thewalt. *Rev. Mod. Phys.*, 77:1173, 2005.

[38] M. Catti, G. Sandrone, and R. Dovesi. *Phys. Rev. B*, 55:16122, 1997.

[39] M. Catti, G. Valerio, R. Dovesi, and M. Causá. *Phys. Rev. B*, 49:14179, 1994.

[40] D. M. Ceperley and B. J. Alder. *Phys. Rev. Lett.*, 45:566, 1980.

[41] G. K.-L. Chan. *J. Chem. Phys.*, 110:4710, 1999.

[42] E. K. Chang, M. Rohlfing, and S. G. Louie. *Phys. Rev. Lett.*, 85:2613, 2000.

[43] R. Chatten, A. V. Chadwick, A. Rougier, and P. J. D. Lindan. *J. Phys. Chem. B*, 109:3146, 2005.

[44] W. Chen and A. Pasquarello. *Phys. Rev. B*, 86:035134, 2012.

[45] W. Chen and A. Pasquarello. *Phys. Rev. B*, 88:115104, 2013.

[46] W. Chen and A. Pasquarello. *Phys. Rev. B*, 90:165133, 2014.

[47] L. Chiodo, J. M. García-Lastra, A. Iacomino, S. Ossicini, J. Zhao, H. Petek, and A. Rubio. *Phys. Rev. B*, 82:045207, 2010.

[48] P. A. Christiansen, Y. S. Lee, and K. S. Pitzer. *J. Chem. Phys.*, 71:4445, 1979.

[49] J. C. Conesa. *J. Phys. Chem. C*, 114:22718, 2010.

[50] J. C. Conesa. *J. Phys. Chem. C*, 116:18884, 2012.

[51] J. C. Conesa. *Catal. Today*, 208:11, 2013.

[52] Y. Cong, B. Li, S. Yue, D. Fan, and X. Wang. *J. Phys. Chem. C*, 113:13974, 2009.

[53] A. Continenza and A. Di Pomponio. *Phys. Rev. B*, 54:13687, 1996.

[54] F. Corà, M. Alfredsson, G. Mallia, D. S. Middlemiss, W. C. Mackrodt, R. Dovesi, and R. Orlando. In N. Kaltsoyannis and J. E. McGrady, editors, *Principles and Applications of Density Functional Theory in Inorganic Chemistry II*, pages 171–232. Springer, 2004.

[55] P.A. Cox. *Transition Metal Oxides: An Introduction to Their Electronic Structure and Properties*. Oxford University Press, 2010.

[56] W. A. Crichton, P. Bouvier, and A. Grzechnik. *Mater. Res. Bull.*, 38:289, 2003.

[57] D. C. Cronemeyer. *Phys. Rev.*, 113:1222, 1959.

[58] D. F. Croxall, R. C. C. Ward, C. A. Wallace, and R. C. Kell. *J. Cryst. Growth*, 22:117, 1974.

[59] M. d'Avezac, M. Calandra, and F. Mauri. *Phys. Rev. B*, 71:205210, 2005.

[60] F. De Angelis, C. Di Valentin, S. Fantacci, A. Vittadini, and A. Selloni. *Chem. Rev.*, 114:9708, 2014.

[61] G. A. de Wijs, P. K. de Boer, R. A. de Groot, and G. Kresse. *Phys. Rev. B*, 59:2684, 1999.

[62] P. Deák, B. Aradi, and T. Frauenheim. *Phys. Rev. B*, 83:155207, 2011.

[63] P. Deák, B. Aradi, and T. Frauenheim. *Phys. Rev. B*, 86:195206, 2012.

[64] P. Deák, B. Aradi, T. Frauenheim, E. Janzén, and A. Gali. *Phys. Rev. B*, 81:153203, 2010.

[65] S. K. Deb. *Philos. Mag.*, 27:801, 1973.

[66] S. K. Deb. *Sol. Energy Mater. Sol. Cells*, 92:245, 2000.

[67] R. Del Sole. In P. Halevi, editor, *Photonic Probes of Surfaces*, pages 131–174. Elsevier, 1995.

[68] R. Del Sole and E. Fiorino. *Phys. Rev. B*, 29:4631, 1984.

[69] A. Deneuville and P. Gérard. *J. Electron. Mater.*, 7:559, 1978.

[70] J. Deslippe, G. Samsonidze, D. A. Strubbe, M. Jain, M. L. Cohen, and S. G. Louie. *Comput. Phys. Commun.*, 183:1269, 2012.

[71] A. Di Paola, M. Bellardita, and L. Palmisano. *Catalysts*, 3:36, 2013.

[72] C. Di Valentin and G. Pacchioni. *Acc. Chem. Res.*, 47:3233, 2014.

[73] C. Di Valentin, G. Pacchioni, and A. Selloni. *Phys. Rev. Lett.*, 97:166803, 2006.

[74] C. Di Valentin, G. Pacchioni, and A. Selloni. *J. Phys. Chem. C*, 113:20543, 2009.

[75] U. Diebold. *Surf. Sci. Rep.*, 48:53, 2003.

[76] M. Dolg. Effective core potentials. In *Modern Methods and Algorithms of Quantum Chemistry*, volume 1, page 479. John von Neumann Institute for Computing, 2000.

[77] R. Dovesi et al. *CRYSTAL09 User's Manual*. University of Torino, 2009.

[78] R. Dovesi, R. Orlando, B. Civalleri, C. Roetti, V. R. Saunders, and C. M. Zicovich-Wilson. *Z. Kristallogr.*, 220:571, 2005.

[79] E. Engel and R. M Dreizler. *Density Functional Theory: An Advanced Course*. Springer, 2011.

[80] H. Englisch and R. Englisch. *Physica*, 121A:253, 1983.

[81] M. Ernzerhof, J. P. Perdew, and K. Burke. *Int. J. Quantum Chem.*, 106:3242, 1997.

[82] M. Ernzerhof and G. E. Scuseria. *J. Chem. Phys.*, 110:5029, 1999.

[83] B. W. Faughnan, R. S. Crandall, and P. M. Heyman. *RCA Rev.*, 36:177, 1975.

[84] E. Fermi. *Z. Phys.*, 48:73, 1928.

[85] M. Ferrero, M. Rérat, R. Orlando, R. Dovesi, and I. J. Bush. *J. Phys.: Conf. Series*, 117:012016, 2008.

[86] M. Ferrero, M. Rérat, R. Orlando, R. Dovesi, and R. Dovesi. *J. Chem. Phys.*, 128:014100, 2008.

[87] M. Ferrero, M. Rérat, R. Orlando, R. Dovesi, and R. Dovesi. *J. Comput. Chem.*, 29:1450, 2008.

[88] A. L. Fetter and J. D. Walecka. *Quantum Theory of Many-Particle Systems*. Dover Publications, 2003.

[89] E. Finazzi, C. Di Valentin, G. Pacchioni, and A. Selloni. *J. Chem. Phys.*, 129:154113, 2008.

[90] V. Fiorentini and A. Baldareschi. *Phys. Rev. B*, 51:17196, 1995.

[91] V. Fock. *Z. Physik*, 61:126, 1930.

[92] R. H. French, S. J. Glass, F. S. Ohuchi, Y. N. Xu, and E. Y. Ching. *Phys. Rev. B*, 49:5133, 1994.

[93] C. Freysoldt, B. Grabowski, T. Hickel, J. Neugebauer, G. Kresse, A. Janotti, and C. G. Van de Walle. *Rev. Mod. Phys.*, 86:253, 2014.

[94] C. Friedrich, M. C. Müller, and S. Blügel. *Phys. Rev. B*, 83:081101, 2011.

[95] F. Fuchs, J. Furthmüller, F. Bechstedt, M. Shishkin, and G. Kresse. *Phys. Rev. B*, 76:115109, 2007.

[96] F. Gallino, C. Di Valentin, and G. Pacchioni. *Phys. Chem. Chem. Phys.*, 13:17667, 2011.

[97] F. Gallino, G. Pacchioni, and C. Di Valentin. *J. Chem. Phys.*, 133:144512, 2010.

[98] M. V. Ganduglia-Pirovano, A. Hofmann, and J. Sauer. *Surf. Sci. Rep.*, 62:219, 2007.

[99] C. Gatti, V. R. Saunders, and C. Roetti. *J. Chem. Phys.*, 101:10686, 1994.

[100] M. Gerosa, C. E. Bottani, L. Caramella, G. Onida, C. Di Valentin, and G. Pacchioni. *Phys. Rev. B*, 91:155201, 2015.

[101] A. K. Ghosh, F. G. Wakim, and R. R. Addiss. *Phys. Rev.*, 184:979, 1969.

[102] E. Giamello, M. Chiesa, and M. C. Paganini. In J. Jupille and G. Thornton, editors, *Defects at Oxide Surfaces*, pages 303–326. Springer, 2015.

[103] P. Giannozzi et al. *J. Phys.: Condens. Matter*, 21:395502, 2009.

[104] T. L. Gilbert. *Phys. Rev. B*, 12:2111, 1975.

[105] R. Gillen and J. Robertson. *Phys. Rev. B*, 85:014117, 2012.

[106] M. Gillet, C. Lemire, E. Gillet, and K. Aguir. *Surf. Sci.*, 532:519, 2003.

[107] C. Gionco, S. Livraghi, S. Maurelli, E. Giamello, S. Tosoni, C. Di Valentin, and G. Pacchioni. *Chem. Mater.*, 27:3936, 2015.

[108] C. Gionco, M. C. Paganini, E. Giamello, R. Burgess, C. Di Valentin, and G. Pacchioni. *Chem. Mater.*, 25:2243, 2013.

[109] G. Giuliani and G. Vignale. *Quantum Theory of the Electron Liquid*. Cambridge University Press, 2005.

[110] F. Giustino, S. G. Louie, and M. L. Cohen. *Phys. Rev. Lett.*, 105:265501, 2010.

[111] S. Goedecker, M. Teter, and J. Hutter. *Phys. Rev. B*, 54:1703, 1996.

[112] P. Gori, M. Rakel, C. Cobet, W. Richter, N. Esser, A. Hoffmann, R. Del Sole, A. Cricenti, and O. Pulci. *Phys. Rev. B*, 81:125207, 2010.

[113] M. Govoni and G. Galli. *J. Chem. Theory Comput.*, 11:2680, 2015.

[114] M. Grätzel. *Nature*, 414:338, 2001.

[115] J. H. E. Griffiths, J. Owen, and I. M. Ward. *Nature*, 173:439, 1954.

[116] G. Grosso and G. Pastori Parravicini. *Solid State Physics*. Academic Press, 2013.

[117] Y. Guo, S. J. Clark, and J. Robertson. *J. Phys.: Condens. Matter*, 24:325504, 2012.

[118] F. Gygi and A. Baldareschi. *Phys. Rev. B*, 34:4405, 1986.

[119] D. R. Hamann, M. Schlüter, and C. Chiang. *Phys. Rev. Lett.*, 43:1425, 1979.

[120] J. Harris and R. O. Jones. *J. Phys. F: Met. Phys.*, 4:1170, 1974.

[121] W. A. Harrison. *Pseudopotentials in the Theory of Metals*. W. A. Benjamin, Inc., 1966.

[122] D. R. Hartree. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 24, pages 89–110. Cambridge University Press, 1928.

[123] C. Hartwigsen, S. Goedecker, and J. Hutter. *Phys. Rev. B*, 58:3641, 1998.

[124] R. Haunschild, B. G. Janesko, and G. E. Scuseria. *J. Chem. Phys.*, 131:154112, 2009.

[125] P. J. Hay and W. R. Wadt. *J. Chem. Phys.*, 82:299, 1985.

[126] L. Hedin. *Phys. Rev.*, 139:A796, 1965.

[127] L. Hedin and S. Lundqvist. In F. Seitz, D. Turnbull, and H. Ehrenreich, editors, *Solid State Physics*, volume 23, pages 1 – 181. Academic Press, 1970.

[128] M. A. Henderson, W. S. Epling, C. H. F. Peden, and C. L. Perkins. *J. Phys. Chem. B*, 107:534, 2003.

[129] F. Herman and J. Callaway. *Phys. Rev.*, 89:518, 1953.

[130] C. Herring. *Phys. Rev.*, 57:1169, 1940.

[131] J. Heyd and G. E. Scuseria. *J. Chem. Phys.*, 120:7274, 2004.

[132] J. Heyd, G. E. Scuseria, and M. Ernzerhof. *J. Chem. Phys.*, 118:8207, 2003.

[133] J. Heyd, G. E. Scuseria, and M. Ernzerhof. *J. Chem. Phys.*, 124:219906, 2006.

[134] G. Hodes, D. Cahen, and J. Manassen. *Nature*, 260:313, 1981.

[135] C. Hogan, R. Del Sole, and G. Onida. *Phys. Rev. B*, 68:035405, 2003.

[136] P. Hohenberg and W. Kohn. *Phys. Rev.*, 136:B864, 1964.

[137] G. Hollinger, T. Minh Duc, and A. Deneuville. *Phys. Rev. Lett.*, 37:1564, 1976.

[138] M. Horn, C. F. Schwerdtfeger, and E. P. Meagher. *Z. Kristallogr.*, 136:273, 1972.

[139] Z. Hu and H. Metiu. *J. Phys. Chem. C*, 115:5841, 2011.

[140] M. Hutchins, O. Abu-Alkhair, M. El-Nahass, and K. A. El-Hady. *Mater. Chem. Phys.*, 98:401, 2006.

[141] M. S. Hybertsen and S. G. Louie. *Phys. Rev. B*, 34:5390, 1986.

[142] F. Iori, M. Gatti, and A. Rubio. *Phys. Rev. B*, 85:115129, 2012.

[143] J. F. Janak. *Phys. Rev. B*, 18:7165, 1978.

[144] A. Janotti, J. B. Varley, P. Rinke, N. Umezawa, G. Kresse, and C. G. Van de Walle. *Phys. Rev. B*, 81:085212, 2010.

[145] J. Jaramillo, G. E. Scuseria, and M. Ernzerhof. *J. Chem. Phys.*, 118:1068, 2003.

[146] H. Jiang, R. I. Gomez-Abal, P. Rinke, and M. Scheffler. *Phys. Rev. B*, 81:085119, 2010.

[147] R. O. Jones and O. Gunnarsson. *Rev. Mod. Phys.*, 61:689, 1989.

[148] T. A. Kandiel, L. Robben, A. Alkaim, and D. Bahnemann. *Photochem. & Photobiol. Sci.*, 12:602, 2013.

[149] Wei Kang and Mark S Hybertsen. *Phys. Rev. B*, 82:085203, 2010.

[150] H. Kendrick, A. Arrott, and S. A. Werner. *J. Appl. Phys.*, 39:585, 1968.

[151] G. P. Kerker. *J. Phys. C*, 13:L189, 1980.

[152] D. I. Khomskii. *Transition Metal Compounds*. Cambridge University Press, 2014.

[153] L. Kleinman. *Phys. Rev. B*, 24:7412, 1981.

[154] L. Kleinman and D. M. Bylander. *Phys. Rev. Lett.*, 48:1425, 1981.

[155] F. P. Koffyberg, K. Dwight, and A. Wold. *Solid State Commun.*, 30:433, 1979.

[156] W. Kohn and J. M. Luttinger. *Phys. Rev.*, 98:915, 1955.

[157] W. Kohn and L. J. Sham. *Phys. Rev.*, 140:A1133, 1965.

[158] D. Koller, P. Blaha, and F. Tran. *J. Phys.: Condens. Matter*, 25:435503, 2013.

[159] H.-P. Komsa, T. T. Rantala, and A. Pasquarello. *Phys. Rev. B*, 86:045112, 2012.

[160] G. Korotcenkov. *Mat. Sci. Eng. B*, 139:1, 2007.

[161] M. Krack. *Theor. Chem. Acc.*, 114:145, 2005.

[162] B. Králik, E. K. Chang, and S. G. Louie. *Phys. Rev. B*, 57:7027, 1998.

[163] M. Krauss and W. J. Stevens. *Ann. Rev. Phys. Chem.*, 35:357, 1984.

[164] G. Kresse and D. Joubert. *Phys. Rev. B*, 59:1758, 1999.

[165] M. Kröger, S. Hamwi, J. Meyer, T. Riedl, W. Kowalsky, and A. Kahn. *Appl. Phys. Lett.*, 95:123301, 2009.

[166] H. M. Krutter. *Phys. Rev.*, 48:664, 1935.

[167] W. Ku and A. G. Eguiluz. *Phys. Rev. Lett.*, 89:126401, 2002.

[168] A. Kudo and Y. Miseki. *Chem. Soc. Rev.*, 38:253, 2009.

[169] S. Kümmel and L. Kronik. *Rev. Mod. Phys.*, 80:3, 2008.

[170] F. Labat, P. Baranek, C. Domain, C. Minot, and C. Adamo. *J. Chem. Phys.*, 126:154703, 2007.

[171] J. Lægsgaard and K. Stokbro. *Phys. Rev. B*, 61:12590, 2000.

[172] J. Lægsgaard and K. Stokbro. *Phys. Rev. Lett.*, 86:2834, 2001.

[173] C. Lambert-Mauriat and V. Oison. *J. Phys.: Condens. Matter*, 18:7361, 2006.

[174] L. D. Landau. *Sov. Phys. JETP*, 3:920, 1957.

[175] L. D. Landau. *Sov. Phys. JETP*, 5:101, 1957.

[176] M. Landmann, E. Rauls, and W. G. Schmidt. *J. Phys.: Condens. Matter*, 24:195503, 2012.

[177] D. C. Langreth and J. P. Perdew. *Phys. Rev. B*, 21:5469, 1980.

[178] S. Lany and A. Zunger. *Phys. Rev. B*, 78:235104, 2008.

[179] S. Lany and A. Zunger. *Modelling Simul. Mater. Sci. Eng.*, 17:084002, 2009.

[180] C. Lee, W. Yang, and R. G. Parr. *Phys. Rev. B*, 37:785, 1988.

[181] G. H. Lee, T. Kawazoe, and M. Ohtsu. *Solid State Commun.*, 124:163, 2002.

[182] M. Leslie and M. J. Gillan. *J. Phys. C*, 18:973, 1985.

[183] M. Levy. *Proc. Natl. Acad. Sci.*, 76:6062, 1979.

[184] M. Levy. *Phys. Rev. A*, 26:1200, 1982.

[185] W. Y. Liang and A. D. Yoffe. *Phys. Rev. Lett.*, 20:59, 1968.

[186] D. R. Lide. CRC Press/Taylor and Francis, 79th edition, 1998/1999.

[187] D. R. Lide. CRC Press/Taylor and Francis, 90th edition, 2010.

[188] E. H. Lieb. *Int. J. Quantum Chem.*, 24:243, 1983.

[189] A. I. Liechtenstein, V. I. Anisimov, and J. Zaanen. *Phys. Rev. B*, 52:R5467, 1995.

[190] B. O. Loopstra and H. M. Rietveld. *Acta Crystallogr., Sect. B*, 25:1420, 1969.

[191] S. Lutfalla, V. Shapovalov, and A. T. Bell. *J. Chem. Theory Comput.*, 7:2218, 2011.

[192] Grüning M., R. Shaltaf, and G. M. Rignanese. *Phys. Rev. B*, 81:035330, 2010.

[193] M. Magagnini, P. Gianozzi, and A. Dal Corso. *Phys. Rev. B*, 61:2621, 2000.

[194] G. Makov and M. C. Payne. *Phys. Rev. B*, 51:4014, 1995.

[195] A. Malashevich, M. Jain, and S.G. Louie. *Phys. Rev. B*, 89:075205, 2014.

[196] B. D. Malone and M. L. Cohen. *J. Phys.: Condens. Matter*, 25:105503, 2013.

[197] A. Marini, G. Onida, and R. Del Sole. *Phys. Rev. Lett.*, 88:016403, 2001.

[198] M. A. L. Marques, J. Vidal, M. J. T. Oliveira, L. Reining, and S. Botti. *Phys. Rev. B*, 83:035119, 2011.

[199] P. Mars and D. W. van Kravelen. *Chem. Eng. Sci.*, 3:41, 1954.

[200] M. Marsman, J. Paier, A. Stroppa, and G. Kresse. *J. Phys.: Condens. Matter*, 20:064201, 2008.

[201] R. M. Martin. *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press, 2004.

[202] L. F. Mattheiss. *J. Phys.: Condens. Matter*, 8:5987, 1996.

[203] G. Mattioli, P. Alippi, F. Filippone, R. Caminiti, and A. Amore Bonapasta. *J. Phys. Chem. C*, 114:21694, 2010.

[204] G. Mattioli, F. Filippone, P. Alippi, and A. Amore Bonapasta. *Phys. Rev. B*, 78:241201(R), 2008.

[205] E. W. McFarland and H. Metiu. *Chem. Rev.*, 113:4391, 2013.

[206] D. Mearns. *Phys. Rev. B*, 38:5906, 1988.

[207] B. K. Meyer, F. Lohse, J. M. Spaeth, and J. A. Weil. *J. Phys. C: Solid State Phys.*, 17:L31, 1984.

[208] J. Meyer, Kröger, S. Hamwi, F. Gnam, T. Riedl, W. Kowalsky, and A. Kahn. *Appl. Phys. Lett.*, 96:193302, 2010.

[209] A. Migani, D. J. Mowbray, A. Iacomino, J. Zhao, H. Petek, and A. Rubio. *J. Am. Chem. Soc.*, 135:11429, 2013.

[210] A. Migani, D. J. Mowbray, J. Zhao, and H. Petek. *J. Chem. Theory Comput.*, 11:239, 2014.

[211] D. B. Migas, V. L. Shaposhnikov, V. N. Rodin, and V. E. Borisenko. *J. Appl. Phys.*, 108:093713, 2010.

[212] J. Moellmann, S. Ehrlich, R. Tonner, and S. Grimme. *J. Phys.: Condens. Matter*, 24:424206, 2012.

[213] H. J. Monkhorst and J. D. Pack. *Phys. Rev. B*, 13:5188, 1976.

[214] B. J. Morgan and G. W. Watson. *J. Phys. Chem. C*, 114:2321, 2010.

[215] J. Muscat, V. Swamy, and N. M. Harrison. *Phys. Rev. B*, 65:224112, 2002.

[216] J. Muscat, A. Wander, and N. M. Harrison. *Chem. Phys. Lett.*, 342:397, 2001.

[217] R. Nada, C. R. A. Catlow, R. Dovesi, and P. Pisani. *Phys. Chem. Miner.*, 17:353, 1990.

[218] K. Nassau and B. E. Prescott. *Phys. Status Solidi A*, 29:659, 1975.

[219] G. A. Niklasson and C. G Granqvist. *J. Mater. Chem.*, 17:127, 2007.

[220] Y. Noel, C. M. Zicovich-Wilson, B. Civalleri, P. D'Arco, and R. Dovesi. *Phys. Rev. B*, 65:014111, 2001.

[221] M. Nolan, S. D. Elliot, J. S. Mulley, R. A. Bennet, M. Basham, and P. Mulheran. *Phys. Rev. B*, 77:235424, 2008.

[222] M. Nolan and G. W. Watson. *J. Chem. Phys.*, 125:144701, 2006.

[223] R. H. D. Nuttal and J. A. Weil. *Solid State Commun.*, 35:789, 1980.

[224] R. H. D. Nuttal and J. A. Weil. *Can. J. Phys.*, 59:1696, 1981.

[225] M. C. M. O'Brien. *Proc. R. Soc. London, Ser. A*, 231:404, 1955.

[226] G. Onida, L. Reining, and A. Rubio. *Rev. Mod. Phys.*, 74:601, 2002.

[227] R. Orlando, V. Lacivita, R. Bast, and K. Ruud. *J. Chem. Phys.*, 132:244106, 2010.

[228] H. Over, Y. D. Kim, A. P. Seitsonen, S. Wendt, E. Lundgren, M. Schmid, P. Varga, A. Morgante, and G. Ertl. *Science*, 287(5457):1474, 2000.

[229] Ü. Özgür, Ya. I. Alivov, C. Liu, A. Teke, M. A. Reshchikov, S. Doğan, V. Avrutin, S.-J. Cho, and H. Morkoç. *J. Appl. Phys.*, 98:041301, 2005.

[230] G. Pacchioni. *Solid State Sci.*, 2:161, 2000.

[231] G. Pacchioni. *Chem. Phys. Chem.*, 4:1041, 2003.

[232] G. Pacchioni. *Catal. Lett.*, 145:80, 2015.

[233] G. Pacchioni. In J. Jupille and G. Thornton, editors, *Defects at Oxide Surfaces*, pages 1–28. Springer, 2015.

[234] G. Pacchioni, F. Frigoli, D. Ricci, and J. A. Weil. *Phys. Rev. B*, 63:054102, 2000.

[235] J. Paier, M. Marsman, K. Hummer, G. Kresse, I. C. Gerber, and J. G. Ángyán. *J. Chem. Phys.*, 124:154709, 2006.

[236] J. Paier, M. Marsman, and G. Kresse. *Phys. Rev. B*, 78:121201(R), 2008.

[237] C. Paracchini and G. Schianchi. *Phys. Status Solidi A*, 72:K129, 1982.

[238] J. Pascual, J. Camassel, and H. Mathieu. *Phys. Rev. Lett.*, 39:1490, 1977.

[239] C. E. Patrick and F. Giustino. *J. Phys.: Condens. Matter*, 24:202201, 2012.

[240] M.F. Peintinger, D. Vilela Oliveira, and T. Bredow. *J. Comput. Chem.*, 34:451, 2013.

[241] J. P. Perdew and K. Burke. *Int. J. Quantum Chem.*, 57:309, 1996.

[242] J. P. Perdew, K. Burke, and M. Ernzerhof. *Phys. Rev. Lett.*, 77:3865, 1996.

[243] J. P. Perdew, K. Burke, and M. Ernzerhof. *J. Chem. Phys.*, 105:9982, 1996.

[244] J. P. Perdew, M. Ernzerhof, and K. Burke. *J. Chem. Phys.*, 105:9982, 1996.

[245] J. P. Perdew and M. Levy. *Phys. Rev. Lett.*, 51:1884, 1983.

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

[247] J. P. Perdew and Y. Wang. *Phys. Rev. B*, 45:13244, 1992.

[248] J. P. Perdew and A. Zunger. *Phys. Rev. B*, 23:5048, 1981.

[249] C. Persson and A. Ferreira da Silva. *Appl. Phys. Lett.*, 86:231912, 2005.

[250] J. C. Phillips and L. Kleinman. *Phys. Rev.*, 116:287, 1959.

[251] Y. Ping and G. Galli. *J. Phys. Chem. C*, 118:6019, 2014.

[252] Y. Ping, W. A. Goddard III, and G. Galli. *J. Am. Chem. Soc.*, 137:5264, 2015.

[253] Y. Ping, D. Rocca, and G. Galli. *Chem. Soc. Rev.*, 42:2437, 2013.

[254] Y. Ping, D. Rocca, and G. Galli. *Phys. Rev. B*, 87:165203, 2013.

[255] J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley. *Int. J. Quantum Chem.*, 16:225, 1979.

[256] S. Rangan, S. Katalinic, R. Thorpe, R. A. Bartynski, J. Rochford, and E. Galoppini. *J. Phys. Chem. C*, 114:1139, 2010.

[257] S. Refaely-Abramson, S. Sharifzadeh, M. Jain, R. Baer, J. B. Neaton, and L. Kronik. *Phys. Rev. B*, 88:081204, 2013.

[258] D. C. Reynolds, D. C. Look, B. Jogai, C. W. Litton, G. Cantwell, and W. C. Harsch. *Phys. Rev. B*, 60:2340, 1999.

[259] G. Ribes, S. Bruyere, D. Roy, C. Parthasarthy, M. Muller, M. Denais, V. Huard, T. Skotnicki, and G. Ghibaudo. *IEEE Trans. Device Mater. Reliab.*, 6:132, 2006.

[260] C. E. Rice and W. R. Robinson. *Acta Crystallogr., Sect. B*, B33:1342, 1977.

[261] J. Robertson. *Eur. Phys. J. Appl. Phys.*, 28:265, 2004.

[262] M. Rohlfing, P. Krüger, and J. Pollmann. *Phys. Rev. Lett.*, 75:3489, 1995.

[263] E. Ruiz, M. Llunel, and P. Alemany. *J. Solid State Chem.*, 176:400, 2003.

[264] V. R. Saunders. *Faraday Symp. Chem. Soc.*, 19:79, 1984.

[265] A. Sawa. *Mater. Today*, 11:28, 2008.

[266] S. Sayan, R. A. Bartynski, X. Zhao, E. P. Gusev, D. Vanderbilt, M. Croft, M. Banaszak Holl, and E. Garfunkel. *Phys. Status Solidi B*, 241:2246, 2004.

[267] U. Scherz and M. Scheffler. *Semicond. Semimetals*, 38:1, 1993.

[268] O. F. Schirmer. *Solid State Commun.*, 18:1349, 1976.

[269] R. Schnadt and A. Räuber. *Solid State Commun.*, 9:159, 1971.

[270] A. K. See and R. A. Bartynski. *Phys. Rev. B*, 50:12064, 1994.

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

[272] M. Setvin, C. Franchini, X. Hao, M. Schmid, A. Janotti, M. Kaltak, C. G. Van de Walle, G. Kresse, and U. Diebold. *Phys. Rev. Lett.*, 113:086402, 2014.

[273] H. Sezen, M. Buchholz, A. Nefedov, C. Natzeck, S. Heissler, C. Di Valentin, and C. Wöll. *Sci. Rep.*, 4:3808, 2014.

[274] L. J. Sham and M Schlüter. *Phys. Rev. Lett.*, 51:1888, 1983.

[275] R. T. Sharp and G. K. Horton. *Phys. Rev.*, 90:317, 1953.

[276] B. Shih, P. Zhang, M. L. Cohen, and S. G. Louie. *Phys. Rev. Lett.*, 105:146401, 2010.

[277] T. Shimazaki and Y. Asai. *J. Chem. Phys.*, 130:164702, 2009.

[278] T. Shimazaki and Y. Asai. *J. Chem. Phys.*, 132:224105, 2010.

[279] T. Shimazaki and T. Nakajima. *J. Chem. Phys.*, 141:114109, 2014.

[280] S. H. Shin, G. V. Chandrashekhar, R. E. Loehman, and J. M. Honig. *Phys. Rev. B*, 8:1364, 1973.

[281] M. Shishkin and G. Kresse. *Phys. Rev. B*, 75:235102, 2007.

[282] H. B. Shore, J. H. Rose, and E. Zaremba. *Phys. Rev. B*, 15:2858, 1977.

[283] J. H. Skone, M. Govoni, and G. Galli. *Phys. Rev. B*, 89:195112, 2014.

[284] J. C. Slater. *Phys. Rev.*, 45:794, 1934.

[285] J. C. Slater. *Adv. Quantum Chem.*, 6:1, 1972.

[286] K. Smits. *Luminiscence of zirconia nanocrystals*. Ph.D. thesis, University of Latvia, 2010.

[287] K. Smits, L. Grigorjeva, W. Łojkowski, and J. D. Fidelus. *Phys. Status Solidi C*, 4:770, 2007.

[288] X. Solans-Monfort, V. Branchadell, M. Sodupe, M. Sierka, and J. Sauer. *J. Chem. Phys.*, 121:6034, 2004.

[289] M. Städele, J. A. Majewski, P. Vogl, and A. Görling. *Phys. Rev. Lett.*, 79:2089, 1997.

[290] M. Stankovski, G. Antonius, D. Waroquiers, A. Miglio, H. Dixit, K. Sankaran, M. Giantomassi, X. Gonze, M. Côté, and G.-M. Rignanese. *Phys. Rev. B*, 84:241201, 2011.

[291] E. V. Stefanovich, A. L. Shluger, and C. R. A. Catlow. *Phys. Rev. B*, 49:11560, 1994.

[292] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch. *J. Phys. Chem.*, 98:11623, 1994.

[293] G. Strinati. *Riv. Nuovo Cimento*, 11:1, 1988.

[294] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications, 1996.

[295] H. Tang, F. Levy, H. Berger, and P. E. Schmid. *Phys. Rev. B*, 52:7771, 1995.

[296] S. E. Taylor and F. Bruneval. *Phys. Rev. B*, 84:075155, 2011.

[297] Y. Tezuka, S. Shis, T. Ishii, T. Ejima, S. Suzuki, and S. Sato. *J. Phys. Soc. Japan*, 63:347, 1994.

[298] A. G. Thomas, W. R. Flavell, A. K. Mallick, A. R. Kumarasinghe, D. Tsoutsou, N. Khan, C. Chatwin, S. Rayner, G. C. Smith, R. L. Stockbauer, S. Warren, T. K. Johal, S. Patel, D. Holland, A. Taleb, and F. Wiame. *Phys. Rev. B*, 75:035105, 2007.

[299] L. H. Thomas. *Proc. Cambridge Philos. Soc.*, 23:542, 1927.

[300] L. Thulis and J. Guerra. *Phys. Rev. B*, 77:195112, 2008.

[301] J. To, A. A. Sokol, S. A. French, N. Kaltsoyannis, and C. R. A. Catlow. *J. Chem. Phys*, 122:144704, 2005.

[302] S. Tosoni, C. Di Valentin, and G. Pacchioni. *J. Phys. Chem. C*, 118:3000, 2014.

[303] S. Tosoni, D. Fernandez Hevia, J. Perez Peña, and F. Illas. *Phys. Rev. B*, 85:115114, 2012.

[304] N. Troullier and J. L. Martins. *Phys. Rev. B*, 43:1993, 1991.

[305] J. Uddin and G. E. Scuseria. *Phys. Rev. B*, 74:245115, 2006.

[306] P. Umari, G. Stenuit, and S. Baroni. *Phys. Rev. B*, 81:115104, 2010.

[307] L. Valenzano, Y. Noel, R. Orlando, C.M. Zicovich-Wilson, M. Ferrero, and R. Dovesi. *Theor. Chem. Acc.*, 117:991, 2007.

[308] R. van de Krol, Y. Liang, and J. Schoonman. *J. Mater. Chem.*, 18:2311, 2008.

[309] C. G. Van de Walle and J. Neugebauer. *J. Appl. Phys.*, 95:3851, 2004.

[310] D. Vanderbilt. *Phys. Rev. B*, 32:8412, 1985.

[311] D. Vanderbilt. *Phys. Rev. B*, 41:7892, 1990.

[312] C. Verdi, E. Mosconi, F. De Angelis, M. Marsili, and P. Umari. *Phys. Rev. B*, 90:155410, 2014.

[313] S. H. Vosko, L. Wilk, and M. Nusair. *Can. J. Phys.*, 58:1200, 1980.

[314] R. Wan, H. Zhang, and J. Peng. *Chem. Phys. Lett.*, 513:17, 2011.

[315] F. Wang, C. Di Valentin, and G. Pacchioni. *J. Chem. Phys. C*, 115:8345, 2011.

[316] F. Wang, C. Di Valentin, and G. Pacchioni. *Phys. Rev. B*, 84:073103, 2011.

[317] F. Wang, C. Di Valentin, and G. Pacchioni. *J. Phys. Chem. C*, 116:10672, 2012.

[318] J. A. Weil, J. R. Bolton, and J. E. Wertz. *Electron Paramagnetic Resonance*. John Wiley & Sons, 1994.

[319] Z. A. Weinberg, G. W. Rubloff, and E. Bassous. *Phys. Rev. B*, 19:3107, 1979.

[320] L. Weinhardt, M. Blum, M. Bär, C. Heske, B. Cole, B. Marsen, and E. L. Miller. *J. Phys. Chem. C*, 112:3078, 2008.

[321] E. Weintraub, T. M. Henderson, and G. E. Scuseria. *J. Chem. Theory Comput.*, 5:754, 2009.

[322] S. H. Wemple. *J. Chem. Phys.*, 67:2151, 1977.

[323] R. C. Whited, C. J. Flaten, and W. C. Walker. *Solid State Commun.*, 13:1903, 1973.

[324] E. P. Wigner and F. Seitz. *Phys. Rev.*, 46:509, 1934.

[325] N. Wiser. *Phys. Rev.*, 129:62, 1963.

[326] J. Wróbel, K. Kurzydłowski, K. Hummer, G. Kresse, and J. Piechota. *Phys. Rev. B*, 80:155124, 2009.

[327] T. Yanai, D. P. Tew, and N. C. Handy. *Chem. Phys. Lett.*, 393:51, 2004.

[328] S. Yang, L. E. Halliburton, A. Manivannan, P. H. Bunton, D. B. Baker, M. Klemm, S. Horn, and A. Fujishima. *Appl. Phys. Lett.*, 94:162114, 2009.

[329] S. B. Zhang and J. E. Northrup. *Phys. Rev. Lett.*, 67:2339, 1991.

[330] Y. Zhang, W. Lin, Y. Li, K. Ding, and J. Li. *J. Phys. Chem. B*, 109:19270, 2005.

[331] J. X. Zheng, G. Ceder, T. Maxisch, W. K. Chim, and W. K. Choi. *Phys. Rev. B*, 75:104112, 2007.

[332] T. Zhu and S.-P. Gao. *J. Phys. Chem. C*, 118:11385, 2014.

[333] C. M. Zicovich-Wilson and R. Dovesi. *J. Phys. Chem. B*, 102:1411, 1998.

[334] A. S. Zyubin, A. M. Mebel, and S. H. Lin. *J. Chem. Phys.*, 119:11408, 2003.