# PROPERTIES OF BINARY OXIDES: A DFT STUDY

OLGA MIROSHNICHENKO

*University of Oulu Graduate School*
*University of Oulu, Faculty of Science*
*Nano and Molecular Systems Research Unit*

Academic dissertation to be presented with the assent of the Doctoral Training Committee of Technology and Natural Sciences of the University of Oulu for public defence in the Auditorium Arina-sali (TA 105), Linnanmaa, on 9 August 2019, at 12 o'clock.

REPORT SERIES IN PHYSICAL SCIENCES Report No. 128
OULU 2019 • UNIVERSITY OF OULU

**Opponent**
Professor Adam Kiejna, University of Wrocław, Poland

**Reviewers**
Professor Adam Foster, Aalto University, Finland
Professor Kai Nordlund, University of Helsinki, Finland

**Custos**
Professor Matti Alatalo, University of Oulu, Finland

ISBN 978-952-62-2300-1
ISBN 978-952-62-2301-8 (PDF)
ISSN 1239-4327
PunaMusta OY
Oulu 2019

# Abstract

**Miroshnichenko, Olga, Properties of binary oxides: a DFT study**
University of Oulu Graduate School; University of Oulu, Faculty of Science;
Nano and Molecular Systems Research Unit
Report Series in Physical Sciences No. 128 (2019)
University of Oulu, P.O. Box 8000, FI-90014 University of Oulu, Finland

Titanium dioxide nanoparticles are used in an enormous amount of applications. Their properties are different from bulk $TiO_2$ and are affected by adsorbates that are unavoidably present on the surface. In this thesis, the effect of OH and $SO_4$ groups (the adsorbants present on the surface during manufacturing) on the properties of anatase-structured $TiO_2$ nanoparticles is studied. It was found that the above mentioned groups change both the geometric and electronic structure of nanoparticles, resulting in changes in the photoabsorption spectrum.

Bader charges are calculated using electron density from Density Functional Theory calculations. They can be used for determination of the oxidation state of the atom. The relation between computed partial charges and oxidation states for binary oxides using data from open materials database has been demonstrated in this work using a linear regression. The applicability of the oxidation state determination by Bader charges for mixed valence compounds and surfaces is considered.

*Keywords:* titanium dioxide, nanoparticles, density functional theory, electronic structure, photoabsorption, oxidation state, Bader charge, data analytics, open materials databases

# Tiivistelmä

**Miroshnichenko, Olga, Binääristen oksidien ominaisuudet: DFT-tutkimus**
Oulun yliopiston tutkijakoulu; Oulun yliopisto, Luonnontieteellinen tiedekunta;
Nano- ja molekyylisysteemien ryhmä
Report Series in Physical Sciences No. 128 (2019)
Oulun yliopisto, PL 8000, 90014 Oulun yliopisto

Titaanidioksidinanopartikkeleita käytetään lukuisissa sovelluksissa. Niiden ominaisuudet poikkeavat kiinteän $TiO_2$:n ominaisuuksista, ja niihin vaikuttavat pinnalle väistämättä absorboituvat aineet. Tässä työssä on tutkittu OH- ja $SO_4$-ryhmien vaikutusta anataasirakenteisten $TiO_2$-nanopartikkelien ominaisuuksiin. Tällaisia ryhmiä esiintyy yleisesti nanopartikkelien pinnalla valmistusprosessien aikana. Työssä havaittiin, että nämä ryhmät muuttavat nanopartikkelien rakenteellisia ja sähköisiä ominaisuuksia, ja siten vaikuttavat myös fotoabsorptiospektriin.

Baderin varaukset voidaan laskea käyttäen tiheysfunktionaaliteoriaan perustuvista laskuista saatavaa elektronitiheyttä. Niitä voidaan käyttää atomin hapetustilan laskemiseen. Tässä työssä on osoitettu, että binääristen oksidien tapauksessa laskettujen osittaisvarauksien ja hapetustilan välillä on yhteys. Tämä yhteys voitiin osoittaa käyttämällä lineaarista regressiota. Työssä tarkastellaan myös menetelmän soveltuvuutta hapetustilojen määrittämiseen sekavalenssiyhdisteille ja pinnoille.

*Asiasanat:* titaanidioksidi, nanopartikkelit, tiheysfunktionaaliteoria, elektronirakenne, fotoabsorptio, hapetustila, Baderin varaus, data-analytiikka, avoimet materiaalitietokannat

# Acknowledgements

This work started as a Master thesis in Lappeenranta University of Technology and continued as a Ph.D. project there. After moving to Oulu, this work was continued as Ph.D. studies in the Nano and Molecular Systems Research Unit (NANOMO), University of Oulu. Thanks to my supervisor, Prof. Matti Alatalo, for making it possible and for the great support during my studies. Thank you for becoming our second family and for the amazing rock concerts. Thanks to Prof. Marko Huttula for providing possibilities to finish my doctoral studies and to travel.

I would like to thank UV-TSM project members: Dr. Sami Auvinen, Ph.D. Juho-Pertti Jalava, Prof. Erik Vartiainen, Prof. Heikki Haario, Ph.D. Veli-Matti Taavitsainen, Ph.D. Ralf-Johan Lamminmäki, and M.Sc. Minna Lindholm for their comments and collaboration.

Thanks to our collaborators in sunny Florida: Prof. Talat S. Rahman, Ph.D. Duy Le, and Dr. Volodymyr Turkowski. Special thanks go to Talat Rahman and Lyman Baker for accommodating us in Orlando during the research visit and for all the entertainment.

Thanks to the University of Oulu Graduate School for provided travel grants and CSC-Scientific Computing Ltd, Espoo, Finland for computational resources and support.

I want to express my gratitude to my follow-up group members: Prof. Erkki Thuneberg, Prof. Juha Vaara, Doc. Saana-Maija Huttula for their guidance. Thanks to reviewers Prof. Adam Foster and Prof. Kai Nordlund for their valuable comments and suggestions.

I would like to thank all my friends and colleagues in Lappeenranta and Oulu for a fun and productive environment. Evgenia, Olga, Marina, thank you for hilarious lunch discussions and parties in Lappeenranta and for staying in touch after moving. Thanks, Vera and Simon, for sauna and marshmallows. I would like to thank people we met in Oulu, who are responsible for many fun evenings and barbeques: Poluianov family, Anton, Frolov family, Georgia, Nønne, Jack. Thank you also goes to my friends from Rostov, who are far physically but still close mentally: Katerina, Maria, Elena. Vita, Marina, I am so happy that we met while studying in St. Petersburg, and thank you for visiting us in Oulu.

Special thanks go to my parents for all the support and freedom they have given me. Grandma, thank you for teaching me Mendeleev table in the mornings when I was a

child and for chemistry "swear phrases". Big thanks go to my grandpa for his infinite and contagious optimism. I would like to thank my beloved husband Sergei for great support both at work and at home, I am so lucky to have you in my life. Finally, I want to thank our son, Nikola, for all the joy and sleep deprivation.

Oulu, May 2019 Olga Miroshnichenko

# List of abbreviations

| | |
|---|---|
| CB | *Conduction Band* |
| DOS | *Density of States* |
| DFT | *Density Functional Theory* |
| GGA | *Generalized Gradient Approximation* |
| HOMO | *the Highest Occupied Molecular Orbital* |
| LUMO | *the Lowest Unoccupied Molecular Orbital* |
| MV | *Mixed Valence* |
| OS | *Oxidation State* |
| PAW | *the Projector-Augmented Wave* |
| REST-API | *the REpresentational State Transfer–Application Programming Interface* |
| SE | *Schrödinger Equation* |
| SV | *Single Valence* |
| TDSE | *Time Dependent Schrödinger Equation* |
| TDDFT | *Time Dependent Density Functional Theory* |
| VASP | *Vienna Ab initio Simulation Package* |
| VB | *Valence Band* |
| XAS | *X-ray Absorption Spectroscopy* |
| XC | *Exchange and Correlation* |
| XPS | *X-ray Photoelectron Spectroscopy* |

# List of original publications

This thesis is based on the following articles, which are referred to in the text by their Roman numerals (I–III):


I Miroshnichenko O., Auvinen S., & Alatalo M. (2015). A DFT study of the effect of OH groups on the optical, electronic, and structural properties of $TiO_2$ nanoparticles. *Phys. Chem. Chem. Phys.*, *17*, 5321–5327. doi: 10.1039/c4cp02789b

II Miroshnichenko O., Posysaev S., & Alatalo M. (2016). A DFT study of the effect of $SO_4$ groups on the properties of $TiO_2$ nanoparticles. *Phys. Chem. Chem. Phys., 18*, 33068–33076. doi: 10.1039/c6cp05681d

III Posysaev S., Miroshnichenko O., Alatalo M., Le D., & Rahman T.S. (2019). Oxidation states of binary oxides from data analytics of the electronic structure. *Comput. Mater. Sci., 161*, 403–414. doi: 10.1016/j.commatsci.2019.01.046


**Author's contribution:** In Papers I and II the author has performed the majority of calculations and written most of the text. The author has performed part of the calculations and code development, and participated in writing of the text in Paper III.

In addition, the author has contributed to the following papers that are not included in the thesis:


I Shi X., Posysaev S., Huttula M., Pankratov V., Hoszowska J., Dousse J.-C., Zeeshan F., Niu Y., Zakharov A., Li T., Miroshnichenko O., Zhang M., Wang X., Huang Z., Saukko S., López González D., van Dijken S., Alatalo M., & Cao W. (2018). Metallic Contact between $MoS_2$ and Ni via Au Nanoglue. *Small, 14*, 1704526. doi: 10.1002/smll.201704526

II Shi X., Huttula M., Pankratov V., Hoszowska J., Dousse J.-Cl., Zeeshan F., Niu Y., Zakharov A., Huang Z., Wang G., Posysaev S., Miroshnichenko O., Alatalo M., & Cao W. (2018). Quantification of bonded Ni atoms for Ni-$MoS_2$ metallic contact through X-ray photoemission electron microscopy. *Microsc. Microanal., 24*, 458–459. doi: 10.1017/S1431927618014526

# Contents

# 1 Introduction

In our modern world, computers play an important role and computational science has grown to be the third mode of science along with the traditional theoretical and experimental ones. In the field of physics, computational scientists are able to model and simulate complex physical problems, predict the outcomes of the experiments, and also analyze them. Calculations are often able to substitute the experiments or at least give suggestions for more effective experimental setups, reducing the time and resources spent on the task. With present day supercomputers, high-throughput calculations, which can be used to understand properties of virtually all materials, are possible. One of such ambitious projects is the Materials Genome Initiative (*Materials Genome Initiative*, 2019).

With the desire of mankind to make things as small as possible and to control matter on the atomic and molecular scale, nanotechnology has emerged. It has been a very fast progressing field alongside with Quantum Mechanics that allowed to explain the behavior of very small systems and particles. Now many quantum mechanical computational techniques exist. Scientific computational tools are becoming more and more sophisticated, and computer resources are getting bigger and faster each year, which allows us to use models that describe reality more and more accurately.

$TiO_2$ is a very popular wide-bandgap semiconductor material. Its popularity arises from the abundance of this material and its interesting properties in many applications. Apart from abundance on Earth, $TiO_2$ plays a role in star dust formation (Kamiński et al., 2013) and TiO was recently found in the atmosphere of a hot Jupiter exoplanet (Sedaghati et al., 2017). When it was discovered that $TiO_2$ can be used for photocatalysis (Fujishima & Honda, 1972), it opened up new promising applications for $TiO_2$ (Hashimoto, Irie, & Fujishima, 2005). Since then thousands of researchers dedicated their work to this semiconductor. Due to a large amount of data, $TiO_2$ became also a benchmark material to test theories and methods. Nanosized titania is used in a great amount of applications, but the properties of very small particles are still not fully understood, especially the optical properties. To conduct experiments with such small nanoparticles could be difficult and expensive, so computational physics would come in handy.

Particles of different sizes are used in different applications, for example, particles about 200 nm are used as pigments as they attenuate visible light. Smaller particles

are more reactive and attenuate UV light, which explains their usage, for example as catalysts and in sunscreens (Varner, Rindfusz, Gaglione, & Viveiros, 2010). As the size of particles determines their application, particle size control during manufacturing is important. This control is often done by light scattering methods, where knowledge of optical properties of particles of different sizes is needed. In the past the optical constants of bulk $TiO_2$ were used, which gave wrong results, and the quality of the particles produced was compromised (Jalava et al., 2015). It was shown computationally that the properties of nanoparticles depend largely on their structure, bonding environment of atoms, and size (Auvinen, Alatalo, Haario, Jalava, & Lamminmäki, 2011; Cho et al., 2016).

The focus of my work was the properties of small, 1–2 nanometers in diameter, $TiO_2$ particles. Previous studies show that the structure of $TiO_2$ nanoparticles is affected strongly by surrounding media and the particle size determines the structural configuration (Zhang & Banfield, 2014). The aim of my studies was to understand how different adsorbed species influence the properties of small $TiO_2$ particles, using density functional theory (DFT) and time-dependent DFT (TDDFT) to address optical properties.

Besides computers, another miracle of our time is the internet. The world wide web has allowed scientists to share their findings with incredible speed. Now the openness of science is encouraged: not only articles should be published in open access journals so that anyone could read them without any subscription needed, but also results and scientific data are shared online for everybody to use. This fast accumulating data opens new possibilities for using results of others to build predictive models.

After getting experience in DFT calculations, the second task of my PhD studies was to utilize online available open DFT data to predict oxidation states of atoms in $TiO_2$ and other binary oxides. Oxidation state (OS) plays an important role in many fields of chemistry and physics. In solids, different OSs can mean different properties of the material. There are several experimental techniques that can measure the OS (for example, X-ray photoelectron spectroscopy (XPS), X-ray absorption spectroscopy (XAS)) (Vickerman & Gilmore, 2009; Walsh, Sokol, Buckeridge, Scanlon, & Catlow, 2018). Naturally, scientists have been developing computational techniques mimicking the experimental ones (Akbari, Hashemi, Niskanen, Huotari, & Hakala, 2015; Bare et al., 2011). They can compute the OS pretty well, but usually the calculations are quite demanding. Along with these sophisticated methods some people use Bader charge calculations for estimating the OS (Xu et al., 2014). Bader charge is an approximation

for calculating the charge transfer between atoms based on partitioning of the electron density. The connection between Bader charges and actual OS has been debatable (Jansen & Wedig, 2008; Raebiger, Lany, Resta, & Zunger, 2009; Raebiger, Lany, & Zunger, 2008; Resta, 2008; Walsh et al., 2018). With increasing amount of computational data in open databases we can use the large amount of computed Bader charges for different compounds to find out if there is a correlation between OS and Bader charges.

The structure of this thesis is the following: the background and underlying theory are introduced in Chapter 2, in Chapter 3 the results from Papers I–II are summarized side by side to have an overview of the effect of both $SO_4$ and OH groups on the properties of $TiO_2$ nanoparticles; then the results of Paper III for correlation between oxidation states of binaries and Bader charges are summarized. Chapters 4 and 5 contain conclusions and corrigenda, respectively.

# 2 Theory and background

## 2.1 Titanium dioxide ($TiO_2$)

$TiO_2$ is a popular semiconductor used in many fields. One of the main applications for titanium dioxide particles is as pigments. This usage is due to excellent light scattering properties, inertness, and affordability (Diebold, 2003; Henderson, 2011). One of the main manufacturing methods of $TiO_2$ particles is the sulfate process, during which titanium-containing ore is reacted with sulfuric acid with subsequent hydrolysis and calcination (Jalava, 2000). The quality of pigments depends a lot on the size and structure of the particles (Jalava et al., 2015).

$TiO_2$ has many crystal structures, the most common are being rutile, anatase, and brookite. The most used and studied $TiO_2$ polymorph is rutile, but at the scale of several nanometers anatase is more stable than the rutile phase (Zhang & Banfield, 2014). Calculations show that anatase is more stable than rutile at 0 K (Muscat, Swamy, & Harrison, 2002). To study properties of $TiO_2$ nanoparticles we concentrate on the anatase polymorph only. It has been shown in the literature that at very small sizes anatase and rutile particles are very similar (Blagojevic, Chen, Steigerwald, Brus, & Friesner, 2009). They have the same chemical constitution, but their structures differ by orientation of $TiO_6$ octahedra (see Fig. 1). During the early stages of crystal formation, the above mentioned octahedra join in chains. Depending on the orientation of the third octahedron, the chain will be linear or zigzag (see the top of Fig. 1), which will form later in rutile crystal structure or anatase respectively. $SO_4$ and OH groups play an important role during hydrolysis, governing among other things the orientation of the $TiO_6$ octahedra (Jalava, 2000).

The conventional cell of anatase has a tetragonal space group $I4_1/amd$ with lattice parameters a=3.784 Å c=9.515 Å bonds=1.93 Å and 1.98 Å (Horn, Schwebdtfeger, & Meagher, 1972). Rutile has also a tetragonal space group $P4_2/mnm$ with lattice parameters: a= 4.594 Å and c=2.958 Å whereas brookite has an orthorhombic *Pbca* space group with lattice parameters: a= 9.184 Å b=5.447 Å and c=5.145 Å (Wyckoff, 1963).

Anatase has a band gap of 3.2 eV. The valence band consists mostly of O$p$ states and the conduction band of Ti$d$ states (Mo & Ching, 1995). Optical properties of $TiO_2$ are important for many applications, for example photocatalysis and solar cells (Diebold,

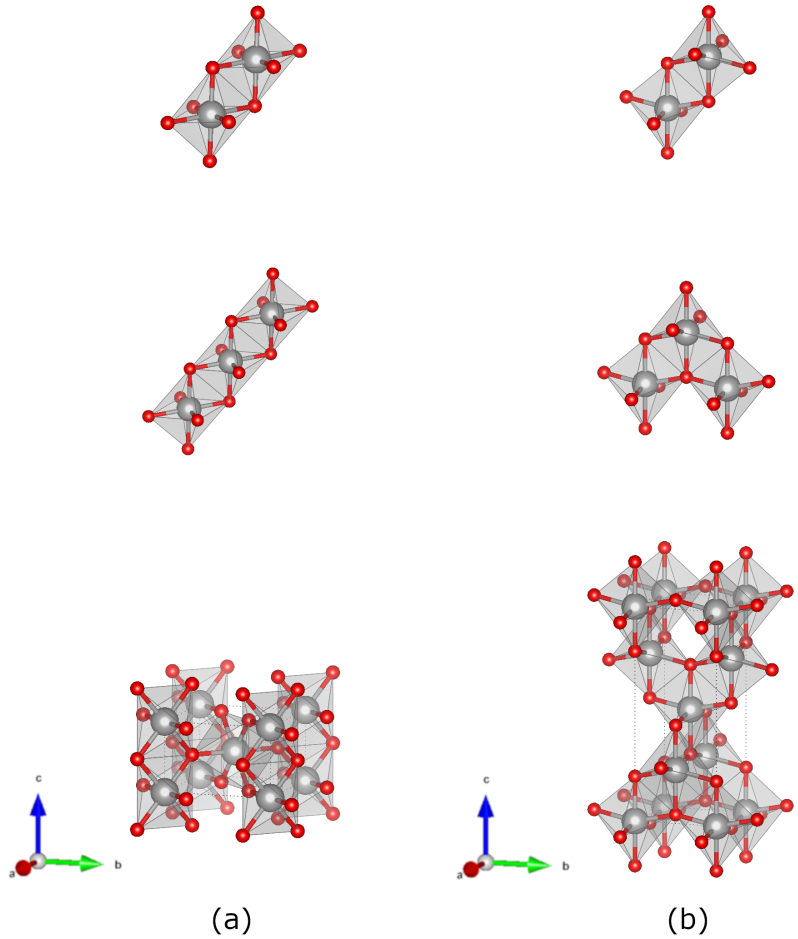



**Fig. 1. Crystal structures and chain formation of (a) rutile, (b) anatase. Oxygen atoms are marked with red color, titanium atoms as gray. The red, green, and blue coordinate axes represent a, b, and c crystal axes, respectively.**

2003). Active research on $TiO_2$ photocatalysis started from the discovery of titanium dioxide's ability to dissociate water under light irradiation (Fujishima & Honda, 1972). As the absorption of visible light is very weak, scientists try to modify $TiO_2$ to change its' bandgap, for example by doping (Asahi, 2001; Peng, Huang, & Huang, 2012). Oxidation states of Ti and O in bulk are usually +4 and -2, respectively. The normal coordination of Ti atoms in the bulk is 6 and of O is 3 (Mo & Ching, 1995).

## 2.2 Density functional theory

The Density Functional Theory is a computational tool for understanding the properties of materials "from first principles" or "ab initio" by applying the fundamental laws of quantum physics to the atoms in the material. However, in the heart of any computational technique there is some simplification of reality, and DFT is not an exception. In principle DFT is an exact theory without any approximations if only we knew all the details of the interactions that occur between electrons, which are hidden inside the exchange-correlation (XC) functional.

In order to understand properties of the matter, or in other words a collection of atoms, we need to know their energy, and how this energy changes if we move some of the atoms. Electrons and nuclei make up atoms. The nucleus is almost 2000 times heavier than electrons, therefore it makes sense to treat electrons and nuclei separately. For every set of nuclei we solve the Schrödinger equation for the set of electrons, moving in the field of these nuclei. This simplification of the problem is called the Born-Oppenheimer approximation (Born & Oppenheimer, 1927).

The famous time-independent Schrödinger equation is

$$H\psi = E\psi, \tag{1}$$

where $H$ is the Hamiltonian of the system, $\psi$ is the many-body electron wavefunction, and $E$ is the energy of the system. The form of the Hamiltonian depends on the problem, but usually in the case of the system with multiple nuclei and many electrons, it can be written as

$$\left[\frac{h^2}{2m}\sum_{i=1}^{N}\bigtriangledown^2 + \sum_{i=1}^{N}V_{ext}(\mathbf{r}_i) + \sum_{i=1}^{N}\sum_{j\neq i}U(\mathbf{r}_i,\mathbf{r}_j)\right]\psi = E\psi. \tag{2}$$

Here, the Hamiltonian consists of (from left to right): the kinetic energy of each electron in the system, the interaction of electrons with nuclei, and the electron-electron interaction. The internuclear interaction is omitted here. If we solve the Schrödinger equation we get all the information about the system, except possible relativistic effects on core electron motion. $\psi$ has a dimension of three times the total number of electrons in the system, which makes it unbearably huge for most systems, therefore approximations are made for the wavefunction, too. One of the most popular is a presentation of the all-electron wavefunction as a product of one-electron wavefunctions ($\psi = \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)...\psi_N(\mathbf{r}_N)$), called the Hartree product. Hartree was the first who proposed the use of independent-electron approximation, which allows to decouple the

Schrödinger equation into a set of one-particle equations (Hartree, 1928). The problem with this approach is that the wavefunction of the system does not obey the principle of antisymmetry. Antisymmetry of the wavefunction assures that the Pauli exclusion principle holds, which does not allow two identical electrons (or other particles with half-integer spin) to occupy the same quantum state simultaneously.

In 1930, Vladimir Fock developed the Hartree approximation further by proposing that the many-electron wavefunction can be represented as a Slater determinant, which satisfies the antisymmetry property (Fock, 1930). This approximation is called the Hartree-Fock approximation. Still, this method requires finding the all-electron wavefunction of the system, which is a demanding computational task.

The all-electron wavefunction is not a physical quantity, only the probability that N electrons are at a particular set of $r_1, r_2..r_N$ coordinates is. Given that we cannot label electrons with $1, 2..N$, the order at which they are at $r_1, r_2..r_N$ coordinates is not important. The probability of finding $N$ electrons in any order at coordinates $r_1, r_2..r_N$ gives us the electron density at a particular position of space

$$n(\mathbf{r}) = 2\sum_i \psi_i^*(\mathbf{r})\psi_i(\mathbf{r}). \tag{3}$$

The multiplier 2 here accounts for electron spin as, according to Pauli principle, the same point of space can be occupied by two electrons if they have different spins.

The origin of DFT was introduced in 1927 by the Thomas-Fermi method (Fermi, 1927; Thomas, 1927), which uses the electron density as the basic variable. The drawback of the approach was that the energy of the system was approximated as a functional of local density at any given point and the exchange and correlation was not taken into account in this method. Lack of exchange means that this method neglected the Pauli exclusion principle. As mentioned earlier, the Hartree-Fock method takes into account the exchange energy. The correlation energy then is called the difference between the energy of the system computed using the Hartree-Fock method and the exact energy of the system. Correlation energy describes interactions between electrons. The exchange was added to the Thomas-Fermi method later by Dirac (Dirac, 1930). Nevertheless, this method used a local approximation of density and therefore it is not accurate enough and is not applicable to the wide range of problems.

The DFT rests upon two Hohenberg-Kohn theorems (Hohenberg & Kohn, 1964; Martin, 2004):

Theorem I. The external potential $V_{ext}(\mathbf{r})$, and hence the total energy $E$, is a unique functional of the ground state electron density $n_0(\mathbf{r})$.

Theorem II. The ground state energy can be obtained variationally: the density that minimizes the total energy is the exact ground state density.

The ground state density determines the total energy and wavefunction of the system. Therefore, we can solve the SE using just the electron density, the function of three coordinates, which allows us in principle to solve the equation regardless of the amount of electrons in the system.

If we knew the form of the functional, we could vary the density, find which density minimizes the functional, and the problem is solved. This approach is used, but with an approximate form of the functional.

The energy functional can be written in terms of individual electron wavefunctions:

$$E[\psi_i] = E_{known}[\psi_i] + E_{XC}[\psi_i], \tag{4}$$

where

$$E_{known}[\{\psi_i\}] = \frac{h^2}{2m}\sum_i \int \psi_i^* \bigtriangledown^2 \psi_i d^3r + \int V_{ext}(\mathbf{r}) n(\mathbf{r}) d^3r + \frac{e^2}{2}\int\int \frac{n(\mathbf{r})n(\mathbf{r}')}{\mathbf{r}-\mathbf{r}'} d^3r d^3r'. \tag{5}$$

The first term in equation (5) is the kinetic energy, second is the electron-nuclei Coulomb interaction, third is the electron-electron electrostatic interaction. In the full description of the functional (4) the unknown term is the exchange-correlation (XC) functional, which includes all unknown quantum mechanical effects.

Even if we knew the form of the XC functional, we would still have a many-electron SE to solve (to find the minimum energy). Kohn and Sham came up with a solution, they divided the problem to a set of one-electron equations (Kohn & Sham, 1965a)

$$\left[\frac{h^2}{2m}\bigtriangledown^2 + V_{ext}(\mathbf{r}) + V_H(\mathbf{r}) + V_{XC}(\mathbf{r})\right]\psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r}). \tag{6}$$

It appears similar to (2), but the solutions of this equation are one-electron wavefunctions. The idea is to use an auxiliary system of noninteracting particles. The first term is the kinetic energy of the electron (quasi-electron or quasi-particle), the second term is a potential that describes interactions of the electron with all nuclei, the third term is a Hartree potential, which is a Coulomb interaction of the electron in the potential generated by all electrons in the system,

$$V_H(\mathbf{r}) = e^2 \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d^3r'. \tag{7}$$

The Hartree potential includes the interaction of the electron with itself, the so-called self-interaction. Correction for it is included in the XC potential, the functional derivative of the XC energy.

The Kohn-Sham equations are solved iteratively: first, we define trial electron density and solve the KS equations to find $\psi_i(\mathbf{r})$. Then, we find electron density from (3) and compare it with our trial electron density. If they are the same within some defined tolerance, we have found the ground state electron density. If they differ, we use the obtained electron density, or mix it with the old trial density and repeat all steps until the difference between densities is very small, or in other words they have reached self-consistency.

As the XC functional is not known, people have come up with approximations. The only situation when XC is known is a hypothetical system called a uniform gas or "jellium" where nuclei and electrons are homogeneously distributed. It underestimates the XC energy, leading to, for example, too high binding energies and short bond lengths in solids, especially in materials, where the electron density varies a lot or fast. This approximation of the XC potential ($V_{XC}(\mathbf{r}) = V_{XC}^{electron\ gas}[n(\mathbf{r})]$) is called Local Density Approximation (LDA) (Kohn & Sham, 1965b), because we use only local density at each point of space. There are many other approximations for the XC functionals, meaning that for now there is no functional suitable for all kinds of systems, but scientists are working to find a universal one. The hierarchy of XC functionals constitute the so-called "Jacob's Ladder" with LDA on the ground step. The next step is an improvement of LDA, which includes also the local gradient of the electron density, called Generalized Gradient Approximation (GGA). There are many ways to include the gradient in the functional, so there are many different GGA XC functionals. The most popular is the Perdew, Burke, and Ernzerhof functional (PBE) (Perdew, Burke, & Ernzerhof, 1996). The next step is Meta-GGA functionals, where also the kinetic energy density is added. Hybrid functionals provide further improvement by including the exact exchange. Climbing up the ladder increases the accuracy but decreases the efficiency of the calculations, one should therefore make a compromise between the two.

DFT describes systems with strongly correlated $d$ and $f$ electrons rather poorly due to the self-interaction error. DFT tends to underestimate bandgaps of these systems. The electrons from partly filled $d$ or $f$ orbitals feel high Coulombic repulsion, which is not described well with the ordinary DFT, as the electrons are assumed to be delocalized. One of the methods to overcome these errors is to use the U correction to the XC functional, for example PBE+U (Dudarev, Botton, Savrasov, Humphreys, & Sutton,

1998; Liechtenstein, Anisimov, & Zaanen, 1995). This correction uses a Hartree-Fock-like potential to describe the interactions of electrons localized on the same atom with the parameter $U_{eff}$ = U-J, where U is the Coulomb interaction and J is the exchange interaction of electrons of a particular angular momentum (Mosey, Liao, & Carter, 2008). The rest of the system is then described with the ordinary DFT.

## 2.3 Time dependent density functional theory

Excitations are of great interest in electronic structure studies. We can investigate two different types of excitations: when the number of electrons changes in the system and when it stays constant. A problem with DFT is that neither Kohn-Sham eigenvalues nor their differences correspond to physical quantities, because these eigenvalues are solutions for non-interacting particles. Ordinary DFT, in principle, can provide a possibility to study excited state properties (Ullrich, 2011), but more practical is to use the extension of it, called the Time-Dependent Density Functional Theory (TDDFT). The idea behind it is the same as in DFT: the time-dependent wavefunction of the system is equivalent to time-dependent electron density.

Experimentally people generally study optical properties of materials by means of spectroscopy. The meaning of this method is to send a probe to a material and investigate the response. The probe in case of properties of atomic materials can be any of the elementary particles, such as, e.g., photons or electrons. The probe triggers a change in the material, for example, electronic transitions or induced dipoles. The change is then measured and analyzed. Computationally, excitations are response functions and the density response function has poles at the excitation frequencies due to changes in external potential. In TDDFT the time-dependent Schrödinger equation is used:

$$H(t)\psi(t) = i\hbar\frac{\partial}{\partial t}\psi(t). \tag{8}$$

The time-dependent theory is based on the Runge-Gross theorem (Runge & Gross, 1984), which is a time-dependent analogue of the Hohenberg theorem. To solve the problem, the auxiliary time-dependent Kohn-Sham system is used with the same electron density as the original system of interacting particles:

$$H(t)\psi_i(t) = i\hbar\frac{\partial}{\partial t}\psi_i(t). \tag{9}$$

The time-dependent effective Hamiltonian reads as

$$H_{eff}(t) = -\frac{1}{2}\bigtriangledown^2 + V_{ext}(\mathbf{r},t) + \int \frac{n(\mathbf{r}',t)}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + V_{xc}[n](\mathbf{r},t). \tag{10}$$

An approximation for $V_{xc}$ is used: it depends on density at time $t$ without taking into account memory effects. The accuracy of TDDFT is limited by the exchange-correlation functional used.

There are two flavours of TDDFT used in practical calculations: linear response in frequency domain and real-time propagation (Yabana & Bertsch, 1996; Yabana, Nakatsukasa, Iwata, & Bertsch, 2006). Linear response is a method of choice if the perturbation of the system is small. With real-time propagation approach we can model also the non-linear phenomena. Real-time perturbation TDDFT allows the use of only occupied states, which lowers the computational load as compared to linear response TDDFT, where a large amount of unoccupied states is needed. These aspects are important to consider in case of a large system.

To perform calculations one can iteratively propagate the time-dependent Schrödinger equation in real time. It is done by expansion of one-particle wave functions $\psi_i(t)$ in a fixed time-independent basis

$$\psi_i(t) = \sum_\alpha c_{i,\alpha}(t)\chi_\alpha. \tag{11}$$

Then the iteration from time $t^n$ to $t^{n+1} = t^n + \delta t$ is

$$c_{i,\alpha}^{n+1} = \sum_{\alpha'} [e^{-iH\delta t}]_{\alpha,\alpha'} c_{i,\alpha'}, \tag{12}$$

where $H$ is a matrix in the basis $\alpha, \alpha'$. $H$ must be considered constant over the interval $\delta t$. Several approaches exist for the expansion of the exponential from (12) to allow the practical calculations to be possible (Ullrich, 2011).

From TDDFT calculations we can get the dipole strength function $S(\omega)$. It is important for calculations of the photoabsorption spectra. The dipole strength function is proportional to the experimental absorption cross-section, which is a probability for a photon to be absorbed by the system. The $S(\omega)$ is calculated from the polarizability:

$$S(\omega) = \frac{2m}{\pi e^2 \hbar} \omega \Im(\alpha(\omega)). \tag{13}$$

The polarizability is

$$\alpha(\omega) = \frac{d(\omega)}{E(\omega)}, \tag{14}$$

where $d(\omega)$ is the dipole moment and $E(\omega)$ is the applied electric field.

To calculate the polarizability, the following approach is used: the solutions $\psi_i^E$ are found for a system with time independent Hamiltonian $H = H_0 - eEx$, where $x$ is a direction of a constant applied electric field $E$. Then, at time t=0, the applied field $E$ is suddenly removed, and the system is allowed to freely evolve for propagated time $t > 0$, with the initial independent particle states $\psi_i^E$ and the Hamiltonian $H_0$. From the Fourier transform of the time-dependent dipole moment, we obtain the dipole strength function, from which we can derive optical constants, such as the dielectric function (Martin, 2004).

## 2.4 Oxidation state

The concept of oxidation state (OS) has been widely used in physics and chemistry. Simply speaking, it shows how many electrons an atom has lost or gained. It is a formal concept and usually it is assumed that bonds between atoms are ionic, which is rarely completely true in nature. Despite the seeming simplicity of the concept it is better not to underestimate it, since it provides useful information about the system and the change of the OS means change in physical and chemical properties of the compound as well as its structure (Walsh et al., 2018).

The definition of OS depends largely on the field and on the application. The International Union of Pure and Applied Chemistry even published a lengthy technical report "toward a comprehensive definition of oxidation state" (Karen, McArdle, & Takats, 2014) with sets of rules for OS determination for different fields. The definition and algorithm for OS calculation have been debated in literature as well as correlation between computed atomic charges and OSs (Jansen & Wedig, 2008; Raebiger et al., 2009, 2008; Resta, 2008; Walsh et al., 2018). Recently, the commonly accepted OS of Ti in $TiO_2$ as +4 was questioned (Koch & Manzhos, 2017; Walsh, Sokol, Buckeridge, Scanlon, & Catlow, 2017).

One chemical element can exist in different OSs in different compounds, for example carbon can be in OSs from -4 to +4. Additionally, one chemical element can be in different OSs in one compound. These compounds are called mixed valence (MV) compounds as opposed to single valence (SV) compounds. An example of MV

compound can be $Co_3O_4$ with Co atoms present both in +2 and +3 OSs. The name "valence" in SV and MV is for historical reasons, since the concept that gave rise to the term OS is valence (Day, Hush, & Clark, 2008).

MV compounds exhibit unique properties and are interesting model systems for the investigation of electron or charge transfer phenomena (Day et al., 2008; Heckmann & Lambert, 2012). Valence electrons could be mixed to a different extent in compounds, depending largely on their crystal structure. Robin-Day came up with three classes for MV compounds: class 1 is for compounds where valence electrons are localized on different atoms; class 3 is for compounds where valence electrons are completely delocalized; class 2 is for intermediate situations where valence electrons are somewhat localized (Robin & Day, 1968). It is impossible to determine integral OS for class 3 MV compounds.

OS can be measured experimentally, for example, with X-ray photoelectron spectroscopy (XPS), X-ray absorption spectroscopy (XAS), K-edge X-ray absorption near edge spectra (XANES) (Vickerman & Gilmore, 2009; Walsh et al., 2018). OSs of atoms are obtained from measured spectra, by using the fact that the binding energy of core electrons in an atom changes with its OS (i.e. core level shift). These spectra can be obtained theoretically (Akbari et al., 2015; Bare et al., 2011) and there have been developments in theoretical calculations of OS from first principles (Postils, Delgado-Alonso, Luis, & Salvador, 2018) and from topology of electronic states (Jiang, Levchenko, & Rappe, 2012). OS can be assigned also based on the atom environment (bonds, angles), this method is called a valence bond analysis (Brown, 2016).

## 2.5 Bader charge analysis

The concept of molecules as atoms bonded together, that is so familiar to us, started to develop only in the 19th century. It is not a trivial task to distinguish atoms between each other in a molecule and define a bond joining them. Studying individual atoms and knowing their contributions can help with understanding the properties of a molecule or material as a whole. A quantum chemist, Richard Bader, proposed to use electron density to divide molecules into atoms (Bader, 1990). By his approach, volumes around each electron density maximum are separated from others by zero-flux surfaces, or surfaces on which the electron density is minimal. These volumes are named Bader volumes. The zero-flux surfaces lie in the bonding regions between atoms.

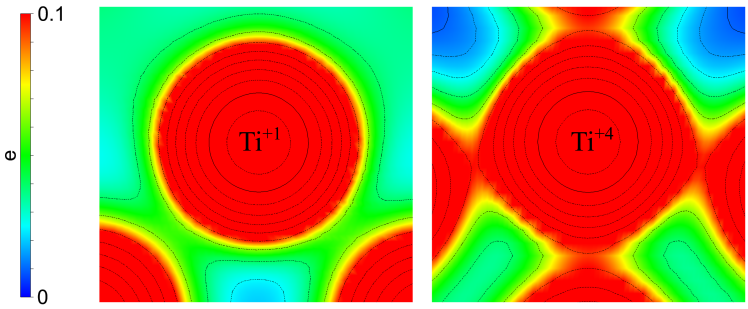



**Fig. 2. Electron density (e/Å$^3$) for Ti in OS +1 (left) and +4 (right). Reproduced from Paper III with permission from Elsevier B.V.**

Apart from the use of the Bader method to divide atoms it can be used to calculate a net charge on atoms. The charge on an atom will be an integral of the electron density over the Bader volume. Graeme Henkelman's group has developed a code for Bader analysis that uses a grid of calculated charge density values (Henkelman, 2017; Henkelman, Arnaldsson, & Jónsson, 2006; Sanville, Kenny, Smith, & Henkelman, 2007; Tang, Sanville, & Henkelman, 2009; Yu & Trinkle, 2011). In a majority of computational codes one can extract the electronic charge density as a 3D grid, for example from VASP (Kresse & Furthmüller, 1996; Kresse & Joubert, 1999) in a CHGCAR file or a CUBE file in the Gaussian package (Frisch et al., 2016). These files serve as an input for the Bader analysis code that will generate the output files with net charges on all atoms in the system.

Let us consider an example. The charge densities of Ti atoms in the OS of +1 and +4 are depicted in Figure 2. We can clearly see the charge density maximum and the bonding regions with minimal charge density (see the color code). Bader charge analysis yields a 0.75e charge on Ti in +1 OS and 2.4e on Ti in +4 OS. These values are not integers as OS usually is, therefore there is no direct correspondence to OS. However, Bader charges and trends in their values across a large amount of compounds can be used for OS determination, as we show in Paper III.

There are other methods to calculate atom charges, such as Mulliken population analysis (Mulliken, 1955). It is an atomic orbitals based approach and therefore its accuracy depends on the basis set used in the calculation. Bader charge analysis is based on the electron density and is insensitive to the basis set if the calculation is well converged. Moreover, electron density is a charge density which can be measured

experimentally. There are other less popular orbital based methods: Löwdin population analysis (Löwdin, 1950), natural population analysis (Reed, Weinstock, & Weinhold, 1985); and electron density based methods: Hirshfeld (Hirshfeld, 1977). All the above mentioned methods have their advantages and disadvantages, the choice of the method depends on the code used for calculations, on the system studied, etc. In Papers I–III only Bader charge analysis was used for partial charges calculations.

## 2.6 Open materials databases and data analytics

Mankind produces lots of data everyday. The amount of scientific data available is tremendous. Why not to use it? One should seize an opportunity to utilize the previously computed data to obtain new results, discover and design materials with desired properties, to find a correlation between different characteristics, and also to avoid duplication of efforts. Computational science has accelerated scientific discovery, providing tools to predict properties of materials without need to perform numerous experiments. Material databases accelerate it even further, providing openly available calculated data without the need to perform lots of calculations. People have made an effort to fill the databases with the results of their work and high-throughput calculations. Machine learning techniques applied to the big amount of data enable scientists to investigate correlations between different material properties.

There are open materials databases available online that accumulate data from millions of first-principles calculations on known and predicted materials. For example, AFLOWLIB, Automatic flow for materials discovery (Curtarolo, Setyawan, Hart, et al., 2012; Curtarolo, Setyawan, Wang, et al., 2012), provides the tools for easy data extraction from the database (REST-API, the REpresentational State Transfer - Application Programming Interface (Taylor et al., 2014)) and has a standard for calculations that requires to use the same calculation details to assure comparability of results (Calderon et al., 2015).

The majority of the time in data analytics is consumed by data processing and data curation. After downloading the data one should make sure that it is all in the same format and has the same features. Once the data is in an appropriate form, we need to clean it, to remove duplicates, outliers, and records with missing or excessive information. It is important to have clean data in order not to end up with wrong conclusions. Visualization of the data usually helps not only to explore the data but also to find errors and outliers.

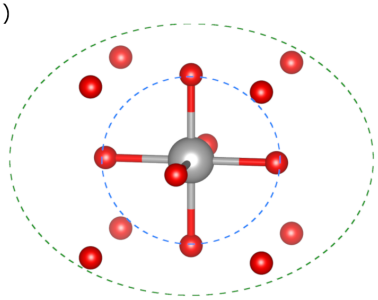
The databases contain data which can be used to calculate new attributes for machine learning algorithms and for data curation such as duplicate removal. For example, in Paper III we have calculated new attributes based on available geometrical data (Fig. 3 (b)): 1) minimal anion-cation bond (*minO_Me*, Fig. 3 (c)), 2) minimal O-O distance (*minO_O*, Fig. 3 (c)), 3) first (*O_nghbrs*, *Me_nghbrs*, Fig. 3 (c)) and second (*O_nghbrs*2, *Me_nghbrs*2, Fig. 3 (c)) coordination numbers of anions and cations around each atom. These attributes were calculated from coordinates and unit cell parameters available from AFLOW. Coordination numbers or the numbers of closest atoms around a particular atom were calculated taking into account the periodicity of the structures. Second coordination numbers mean the second closest atoms to a particular atom. For example, Ti atom in structure with auid aflow:b4b77b98f2bbf457 has 6 O neighbors in first coordination sphere (atoms inside blue dashed circle in Fig. 3 (a)) and 8 in the second (atoms inside green dashed ellipse in Fig. 3 (a)). Coordination numbers were used to find duplicates, compounds with the same constitution, similar unit cells but different coordinates.

## 2.7 Computational details

The full computational details are presented in each publication, so here I briefly describe which codes were used in my work. The main code used in Papers I and II was the GPAW package (Bahn & Jacobsen, 2002; Enkovaara et al., 2010; Mortensen, Hansen, & Jacobsen, 2005). It is a real-space implementation of the projector-augmented wave (PAW) method (Blöchl, 1994; Blöchl, Först, & Schimpl, 2003), that allows the performance of non-periodic calculations, which are more natural for isolated systems such as clusters. This particular package was chosen because of the possibility to perform both DFT and TDDFT calculations for the systems studied. DFT calculations were used for studying structural and electronic properties of the systems, whereas TDDFT calculations were used to obtain photoabsorption spectra.

Another code used in calculations for Paper III is VASP (Vienna Ab initio Simulation Package) (Kresse & Furthmüller, 1996; Kresse & Joubert, 1999), which also uses the PAW formalism. All computed data in AFLOW database was obtained with VASP, therefore we used it for test calculations and verification of results for different XC functionals and for $TiO_2$ surfaces to assure comparability. For Bader charge analysis, the code developed by Henkelman *et al.* was used (Henkelman, 2017; Henkelman et al., 2006; Sanville et al., 2007; Tang et al., 2009; Yu & Trinkle, 2011).

(a)



(b)

| Index | auid | composition | species | natoms | positions_fractional | cell | |
|---|---|---|---|---|---|---|---|
| 0 | aflow:b4b77b98f2bbf457 | 1,1 | O,Ti | 2 | 0.5,0.5,0.5;… | 0.00000000000000 | 2.177825024220… |
| 1 | aflow:e73737a3edc3e07d | 1,1 | O,Ti | 2 | 0.5,0.5,0.5;… | 0.00000000000000 | 2.234579776864… |
| 2 | aflow:b9a416287fe9f994 | 1,1 | O,Ti | 2 | 0.5,0.5,0.5;… | 0.00000000000000 | 2.234888498864… |
| 3 | aflow:b462f391fe6bae64 | 10,6 | O,Ti | 16 | 0.67777366,0… | 1.94140035200259 | 4.951046647029… |
| 4 | aflow:fbdb5ace1cc74df7 | 10,6 | O,Ti | 16 | 0.67807087,0… | 1.94139915271792 | 4.950463021612… |
| 5 | aflow:c1e02612351ed165 | 8,4 | O,Ti | 12 | 0.63140303,0… | 1.91671316611413 | 6.225267552695… |

(c)

| Index | auid | composition | species |
|---|---|---|---|
| 0 | aflow:b4b77b98f2bbf457 | 1,1 | O,Ti |
| 1 | aflow:e73737a3edc3e07d | 1,1 | O,Ti |
| 2 | aflow:b9a416287fe9f994 | 1,1 | O,Ti |
| 3 | aflow:b462f391fe6bae64 | 10,6 | O,Ti |
| 4 | aflow:fbdb5ace1cc74df7 | 10,6 | O,Ti |
| 5 | aflow:c1e02612351ed165 | 8,4 | O,Ti |

…

| O_nghbrs | minO_O | Me_nghbrs | minO_Me | O_nghbrs2 | Me_nghbrs2 |
|---|---|---|---|---|---|
| [0.,6.] | 0.0 | [ 6.,12.] | 2.177… | [0.,8.] | [8.,6.] |
| [0.,6.] | 0.0 | [ 6.,12.] | 2.234… | [0.,8.] | [8.,6.] |
| [0.,6.] | 0.0 | [ 6.,12.] | 2.234… | [0.,8.] | [8.,6.] |
| [0.,0.… | 2.630… | [2.,2.,2… | 1.885… | [0.,0.… | [2.,2.… |
| [0.,0.… | 2.630… | [2.,2.,2… | 1.886… | [0.,0.… | [2.,2.… |
| [0.,0.… | 2.533… | [2.,2.,2… | 1.809… | [0.,0.… | [6.,6.… |

**Fig. 3. (a) First and second coordination spheres for a Ti atom in the structure with auid aflow:b4b77b98f2bbf457, (b) part of AFLOW information for $Ti_xO_y$ structures collected in a database, (c) part of AFLOW information for $Ti_xO_y$ structures plus new computed attributes. Row with data for aflow:b4b77b98f2bbf457 is marked with red.**

# 3 Summary of results

This chapter presents a review of articles included in the thesis. The first part is a summary of Papers I and II combined together. The second part is a review of results of Paper III.

## 3.1 The effect of OH and $SO_4$ groups on the properties of $TiO_2$ nanoparticles

### *3.1.1 Structural properties*

Previously it was shown that the structural properties of $TiO_2$ nanoparticles differ strongly from bulk and these differences have a profound effect on the electronic structure of particles (Auvinen et al., 2011). Thorough analysis of the structure was performed for particles of different sizes. In our work we performed the same analysis for the particles with adsorbed species on their surface, to see how bondlengths and sizes of the cluster are affected by adsorption.

The model structure for a $TiO_2$ nanoparticle in this work was chosen to be a $TiO_2$ anatase particle consisting of 16 $TiO_2$ units, carved from the experimental bulk as described in the work by Persson *et al.* (Persson, Gebhardt, & Lunell, 2003): $TiO_2$ units were stripped off until the desired size was reached, which guaranteed stoichiometry and sufficiently high coordination of all constituent atoms. This size of the particle was chosen as it is small enough for calculations to be feasible but large enough to have a distinguishable anatase structure. An anatase $TiO_2$ polymorph was chosen as it is more stable than rutile in the size of a few nanometers (Zhang & Banfield, 2014) and at 0 K, as predicted by first-principles calculations by Muscat et al. (2002). In a previous work it was found that the symmetrical and spherical particles are more stable than needle-like shapes (Auvinen et al., 2011). Therefore the shape of the studied particle is as symmetrical as possible. After the model structure was constructed it was relaxed in GPAW. Further computational details can be found in Paper I.

To study the effect of the adsorbant coverage on a $(TiO_2)_{16}$ cluster, we considered particles covered with 1, 3, 6, and 16 OH groups and 1–4 $SO_4$ groups. OH groups were attached to on-surface titanium atoms as suggested in Bredow and Jug (1995). Possible OH adsorption sites were tested by the Universal Force Field optimizations (Rappe,

Casewit, Colwell, Goddard III, & Skiff, 1992) (using built-in molecular dynamics auto-optimizer tool in Avogadro molecular editor (*Avogadro: an open-source molecular builder and visualization tool. Version 1.2.0.*, 2019; Hanwell et al., 2012)) and/or additionally by geometry relaxations in GPAW. According to the literature (Liu, Li, Peng, Lu, & Li, 2013; Raj, Shanmugam, Mahalakshmi, & Viswanathan, 2010; Yang, Lu, & Wang, 2003), $SO_4$ groups tend to attach to two on-surface titanium atoms via their single-bonded oxygens. The number of test structures for the favorable adsorption site was tested by relaxation in GPAW. The detailed procedures for testing the adsorption sites for OH and $SO_4$ groups can be found correspondingly in Papers I and II. In both cases the lowest energy configurations were chosen for further study. The structures obtained are not necessarily the most stable geometries, as no global optimization was performed.

All relaxed structures including the bare cluster are depicted in Figs. 4–5. The structural information is collected in Tables 1–2. The bonds between the OH and $SO_4$ groups and the cluster and inside the groups were not taken into account to see the changes in the geometry of the cluster itself. In the bond analysis we have assumed that the Ti-O bond exists if the distance between the Ti and O atoms is smaller than 2.2 Å, the distance between Ti atoms smaller than 2.94 Å for Ti-Ti bond, the distance between O atoms smaller than 1.5 Å for O-O. In Paper I we have used the O-O bond cut off value 1.46 Å, as a sum of covalent radii, but in Paper II due to clear existence of the bond between two oxygens, this cut off radius was increased to 1.5 Å. In this review we are

**Table 1. Number of bonds and the average bond lengths in the relaxed structures. The OH and $SO_4$ groups are not taken into account. Reproduced from Papers I–II with permission from the PCCP Owner Societies.**

| Structure | $n_{Ti-O}$ | $d_{Ti-O}$/Å | $n_{Ti-Ti}$ | $d_{Ti-Ti}$/Å | $n_{O-O}$ | $d_{O-O}$/Å | $n_{All}$ | $d_{All}$/Å |
|---|---|---|---|---|---|---|---|---|
| $Ti_{16}O_{32}$ bare | 68 | 1.872 | 12 | 2.789 | 0 | — | 80 | 2.009 |
| $Ti_{16}O_{32}$+1OH | 66 | 1.867 | 11 | 2.832 | 0 | — | 77 | 2.005 |
| $Ti_{16}O_{32}$+3OH | 65 | 1.878 | 9 | 2.822 | 1 | 1.443 | 75 | 1.985 |
| $Ti_{16}O_{32}$+6OH | 62 | 1.868 | 2 | 2.810 | 2 | 1.459 | 66 | 1.884 |
| $Ti_{16}O_{32}$+16OH | 58 | 1.882 | 1 | 2.848 | 3 | 1.431 | 62 | 1.876 |
| $Ti_{16}O_{32}$+1$SO_4$ | 70 | 1.900 | 11 | 2.814 | 1 | 1.492 | 82 | 2.018 |
| $Ti_{16}O_{32}$+2$SO_4$ | 66 | 1.883 | 10 | 2.803 | 2 | 1.473 | 78 | 1.990 |
| $Ti_{16}O_{32}$+3$SO_4$ | 64 | 1.880 | 9 | 2.815 | 2 | 1.393 | 75 | 1.979 |
| $Ti_{16}O_{32}$+4$SO_4$ | 62 | 1.878 | 7 | 2.850 | 3 | 1.418 | 72 | 1.953 |

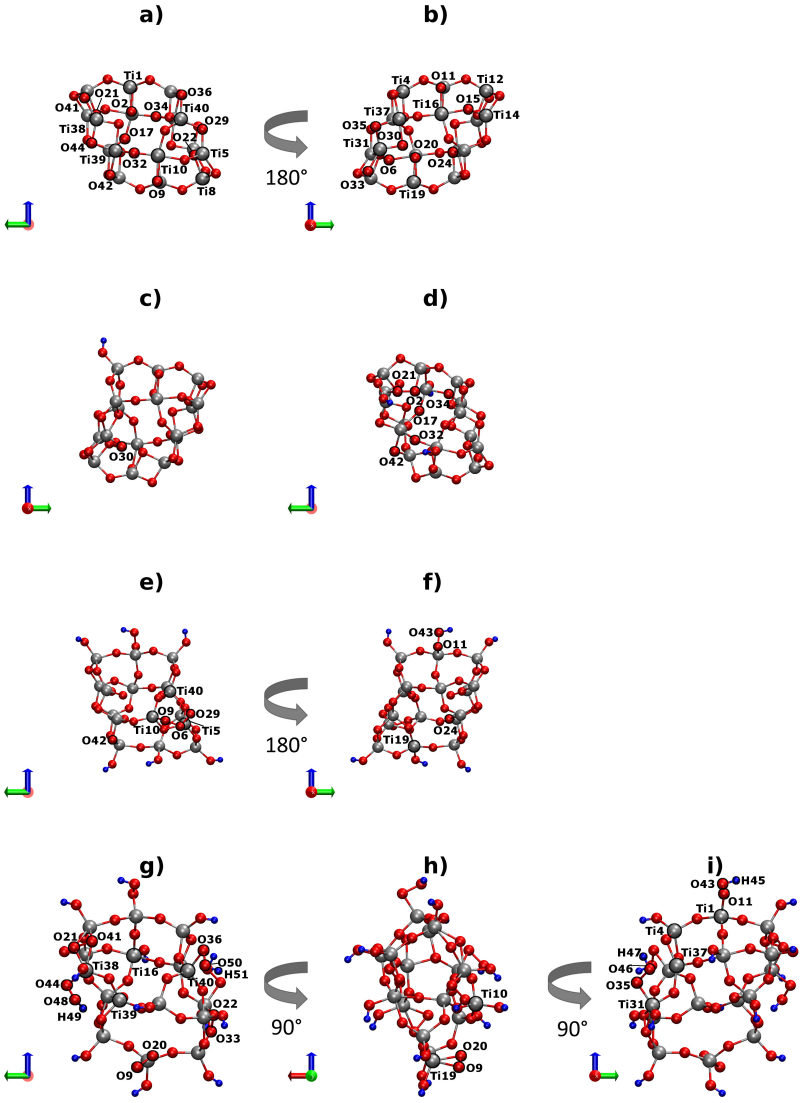



**Fig. 4. The relaxed cluster structures: a) bare $Ti_{16}O_{32}$, b) bare $Ti_{16}O_{32}$ turned by 180° around the z axis relative to a, c) $Ti_{16}O_{32}$ with 1 OH group, d) $Ti_{16}O_{32}$ with 3 OH groups, e) $Ti_{16}O_{32}$ with 6 OH groups, f) $Ti_{16}O_{32}$ with 6 OH groups turned by 180° around the z axis relative to e, g) $Ti_{16}O_{32}$ with 16 OH groups, h) $Ti_{16}O_{32}$ with 16 OH groups turned by 90° around the z axis relative to g, i) $Ti_{16}O_{32}$ with 16 OH groups turned by 90° around the z axis relative to h. Oxygen atoms are marked with red color, titanium atoms as gray color, and hydrogen atoms are blue. The red, green, and blue coordinate axes represent x, y, and z axes, respectively. Reproduced from Paper I with permission from the PCCP Owner Societies.**

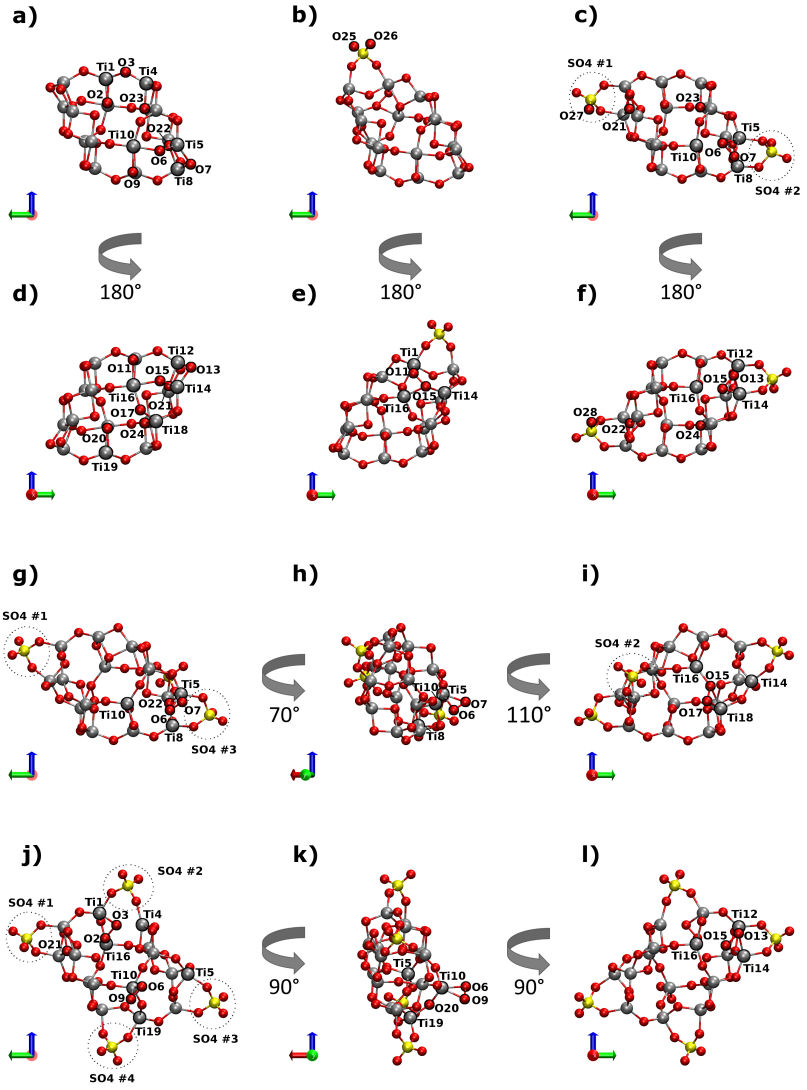



**Fig. 5. The relaxed cluster structures: a) bare $Ti_{16}O_{32}$, b) $Ti_{16}O_{32}$ with 1 $SO_4$ group, c) $Ti_{16}O_{32}$ with 2 $SO_4$ groups, d) bare $Ti_{16}O_{32}$ turned by 180° around the z axis relative to a, e) $Ti_{16}O_{32}$ with 1 $SO_4$ group turned by 180° around the z axis relative to b, f) $Ti_{16}O_{32}$ with 2 $SO_4$ groups turned by 180° around the z axis relative to c, g) $Ti_{16}O_{32}$ with 3 $SO_4$ groups, h) $Ti_{16}O_{32}$ with 3 $SO_4$ groups turned by 70° around the z axis relative to g, i) $Ti_{16}O_{32}$ with 3 $SO_4$ groups turned by 110° around the z axis relative to h, j) $Ti_{16}O_{32}$ with 4 $SO_4$ groups, k)$Ti_{16}O_{32}$ with 4 $SO_4$ groups turned by 90° around the z axis relative to j, and l) $Ti_{16}O_{32}$ with 4 $SO_4$ groups turned by 90° around the z axis relative to k. Oxygen atoms are marked with red color, titanium atoms as gray color, and sulfur atoms are yellow. Reproduced from Paper II with permission from the PCCP Owner Societies.**

using the 1.5 Å value which leads to slightly different structural results for clusters with OH groups on the surface than in Paper I, since more O-O bonds were taken into account.

As can be seen from Tables 1–2, the adsorbed OH and $SO_4$ groups cause the elongation or breakage of Ti-O and Ti-Ti bonds and also the appearance of O-O bonds. OH groups have more pronounced effect on the number of Ti-O and Ti-Ti bonds than $SO_4$ groups. The average bond lengths are decreasing upon adsorption with the exception of the structure with 1 $SO_4$ group, where the appearing O-O bond is relatively long. The average bond lengths are decreasing and at the same time the Ti-O and Ti-Ti bonds are getting longer. This can be explained by the appearance of relatively short O-O bonds. The number of Ti-O bonds is decreasing upon adsorption, again with the exception of the structure with 1 $SO_4$. The average dimensions of the clusters increase upon adsorption, with the exception of 2 $SO_4$ groups which tend to compress the cluster in the *y* direction while only slightly changing the other two dimensions. The $SO_4$ groups in this case are oriented in the *y* direction and cause the outermost corner oxygen atoms to move away from $SO_4$ inside the cluster (Fig. 5 c,f). In other configurations where 2 $SO_4$ groups are oriented also in the *y* direction, this dimension does not change that dramatically, probably due to other distortions occurring in the structure. The relaxed cluster with 16 OH groups attached to every Ti atom on the surface undergoes the biggest change in size by expanding in all three directions with average increase of 20%. After adsorption of 3 and 4 $SO_4$ groups and 16 OH groups side-on coordinated O-O species appear ($\eta^2$-$O_2$, where $\eta$ stands for the hapticity of O-O, $\eta^2$ means that two contiguous O atoms are connected to one Ti atom)(Cramer, Tolman, Theopold, & Rheingold, 2003; Martirez et al., 2015) pointing in the *x* direction, causing the enlarged size in this direction as compared to the bare cluster (Fig. 5 g–l, Fig. 4 g–h). In the case of 6 OH the *x* dimension is increased significantly (31% larger than bare cluster) due to the appearance of triple O-O-O bond on the surface, pointing in the *x* direction (O6-O9-O29, Fig. 4 e), among many other reconstructions.

In Paper I we concentrated mostly on the electronic and optical properties of the clusters, and therefore did not study the structural reconstruction in great detail. In Paper II we took a more thorough look at the formation of O-O bonds in relaxed structures upon $SO_4$ adsorption. In this review it is interesting to perform the same analysis for the structures with adsorbed OH groups. The O-O bonds and Bader charges per O-O bonds are tabulated in Table 3. In the structure with 6 OH groups there is a triple O-O-O bond, which has a sum of Bader charges roughly equal to the average charge of other O-O

**Table 2. Percentage changes in the dimensions of the relaxed structures in three cartesian directions relative to the bare cluster. Negative values stand for the decrease in the dimensions. The OH and $SO_4$ groups are not taken into account. Reproduced from Papers I–II with permission from the PCCP Owner Societies.**

| Structure | x | y | z | Avg. |
|---|---|---|---|---|
| $Ti_{16}O_{32}$+1OH | 7 | -5 | 14 | 5 |
| $Ti_{16}O_{32}$+3OH | 13 | -1 | 6 | 5 |
| $Ti_{16}O_{32}$+6OH | 31 | -10 | 12 | 7 |
| $Ti_{16}O_{32}$+16OH | 15 | 12 | 32 | 20 |
| $Ti_{16}O_{32}$+1$SO_4$ | 4 | -1 | 5 | 2 |
| $Ti_{16}O_{32}$+2$SO_4$ | 1 | -13 | 2 | -5 |
| $Ti_{16}O_{32}$+3$SO_4$ | 6 | 3 | 9 | 6 |
| $Ti_{16}O_{32}$+4$SO_4$ | 36 | -2 | 5 | 10 |

species. Adsorption of 16 OH, 3 $SO_4$, and 4 $SO_4$ groups causes the creation of one $\eta^2$-$O_2$ O-O moiety (O9-O20, O6-O7, and O6-O9 correspondingly, depicted in Fig. 4 g–h, Fig. 5 g–l). These O-O bonds have shorter bond lengths than other O-O bonds and the charge transfer from the cluster to these bonds is smaller than for others. Smaller charge transfer between $\eta^2$-$O_2$ and the cluster is supported by longer than average Ti-O bonds between these oxygens and titaniums to which they are attached. These facts point out the superoxo character of $\eta^2$-$O_2$ (Cramer et al., 2003; Martirez et al., 2015). To verify that the appearance of O-O is not merely a shortcoming of our model, calculations with 3–7 water molecules surrounding the cluster with 1 $SO_4$ group were performed to create a more realistic configuration. During geometrical relaxations the same O11-O15 bond appeared, giving us more confidence in our model.

In the structure with 16 OH groups, four O-O bonds are forming between O atoms from the cluster and O from OH groups due to major structure reconstructions during geometry optimization. These bonds were not mentioned in Table 1, because OH groups were not taken into account there. It is interesting to note that in some cases OH groups got disconnected from Ti atoms to which they were connected before the relaxation, leaving some Ti atoms (Ti37, Ti39, and Ti1) with no OH groups, whereas others have 2 OH groups near them simultaneously (Ti38, Ti40), see Fig. 4 g, i. Ti38 has now 2 OH groups, making it the only Ti atom with 6-fold coordination in this structure, and with two O-O bonds connected to it (Fig. 4 g).

**Table 3. Calculated bond lengths and Bader charges of the O-O species. Reproduced from Paper II with permission from the PCCP Owner Societies.**

| Structure | Atoms | Bond length/Å | Charge/*e* |
|---|---|---|---|
| $Ti_{16}O_{32}$+3OH | O2-O21 | 1.443 | -1.139 |
| $Ti_{16}O_{32}$+6OH | O6-O9 | 1.452 | -0.614 |
| | O6-O29 | 1.466 | -0.575 |
| | O11-O43 | 1.479 | -1.087 |
| $Ti_{16}O_{32}$+16OH | O22-O33 | 1.499 | -1.257 |
| | O9-O20 ($\eta^2$-$O_2$) | 1.334 | -0.522 |
| | O41-O21 | 1.461 | -1.042 |
| | O36-O50 | 1.484 | -1.114 |
| | O35-O46 | 1.481 | -1.175 |
| | O11-O43 | 1.476 | -1.068 |
| | O44-O48 | 1.464 | -1.078 |
| $Ti_{16}O_{32}$+1$SO_4$ | O11-O15 | 1.492 | -1.215 |
| $Ti_{16}O_{32}$+2$SO_4$ | O13-O15 | 1.473 | -1.099 |
| | O6-O7 | 1.474 | -1.100 |
| $Ti_{16}O_{32}$+3$SO_4$ | O15-O17 | 1.442 | -1.050 |
| | O6-O7 ($\eta^2$-$O_2$) | 1.345 | -0.556 |
| $Ti_{16}O_{32}$+4$SO_4$ | O13-O15 | 1.455 | -1.051 |
| | O2-O3 | 1.471 | -1.148 |
| | O6-O9 ($\eta^2$-$O_2$) | 1.327 | -0.452 |

The performed analysis shows that adsorbed OH and $SO_4$ groups introduce a lot of changes into the structure of nanoparticles. Previous work (Auvinen et al., 2011) showed that the Ti-O bondlengths in particles are longer than bulk values, and with increasing particle size these bondlengths are getting longer and therefore closer to bulk. According to our study, Ti-O bondlengths are mostly getting longer upon adsorption of different groups. Sizes of corresponding particles are also getting larger upon the adsorption of OH and $SO_4$ groups, therefore our results are in agreement with earlier findings by Auvinen at al. Adsorbates cause strong rearrangements in the structure, resulting in the appearance of O-O bonds, triple O bonds, and $\eta^2$-$O_2$ species that should introduce substantial changes in the electronic structure as we shall see in the next section. The appearance of $\eta^2$-$O_2$ species is interesting because they are intermediate species in the production of molecular oxygen from water on transition metal based oxides (Martirez

et al., 2015). The adsorbed groups increased the number of undercoordinated atoms (see Papers I–II), which should increase the reactivity of the clusters.

### 3.1.2 Electronic properties

The Densities of States (DOS) for all structures are depicted in Fig. 6. The DOS of the bare cluster is plotted in blue color for reference. In Paper I, the DOS for each structure was plotted relative to their own Fermi levels. In this review all DOS were shifted relative to the Fermi level of the bare cluster to see changes in the band gap more clearly. The Highest Occupied Molecular Orbital (HOMO) and Lowest Unoccupied Molecular Orbital (LUMO) levels of each structure are marked with black dashed lines, the HOMO and LUMO of the bare cluster are plotted with green dots. The energies of the HOMO and LUMO, orbitals of the atoms, on which these levels are located, and differences between HOMO and LUMO are collected in Table 4. In the case of adsorption of 1 OH, 3 OH, 16 OH, 3 $SO_4$, and 4 $SO_4$, the HOMO and LUMO levels almost coincide due to new empty states introduced to the electronic structure (Fig. 6 a, b, d, g, h). The actual band gap then was estimated from the DOS as the difference between the highest valence band (VB) peak and the lowest conduction band (CB) peak. These values are in parentheses in Table 4.

In the bare cluster the HOMO level is on the p-orbitals of O15 and O6 (Fig. 4 b) and the LUMO is on the d-orbitals of Ti14, Ti5, Ti12, Ti8, which are at the corners of the cluster (Fig. 4 a, b). The adsorption of OH groups causes the appearance of new states at the bottom of the VB due to the p-orbitals of O atoms in the OH groups and O-O bonds. The intensity of DOS at the top of the CB is increasing due to the increase of the number of atoms in the system and has no particular predominant contribution. This study shows that OH groups introduce the empty states in the band gap or at the top of the VB, lowering the HOMO-LUMO difference dramatically.

1 OH adsorption causes O30 (Fig. 4 c) to separate from a Ti atom during relaxation and to be 1-fold coordinated, which is energetically unfavorable, causing both HOMO and LUMO to be on the p-orbital of this oxygen. The structure with 3 OH groups has HOMO and LUMO levels mostly on the p-orbitals of O atoms close to the adsorbed OH groups.

The only structure that does not undergo the strong lowering of the HOMO-LUMO gap is a structure with 6 adsorbed OH groups, probably due to the symmetry of the structure and no strong reconstructions during the relaxation as compared to the also

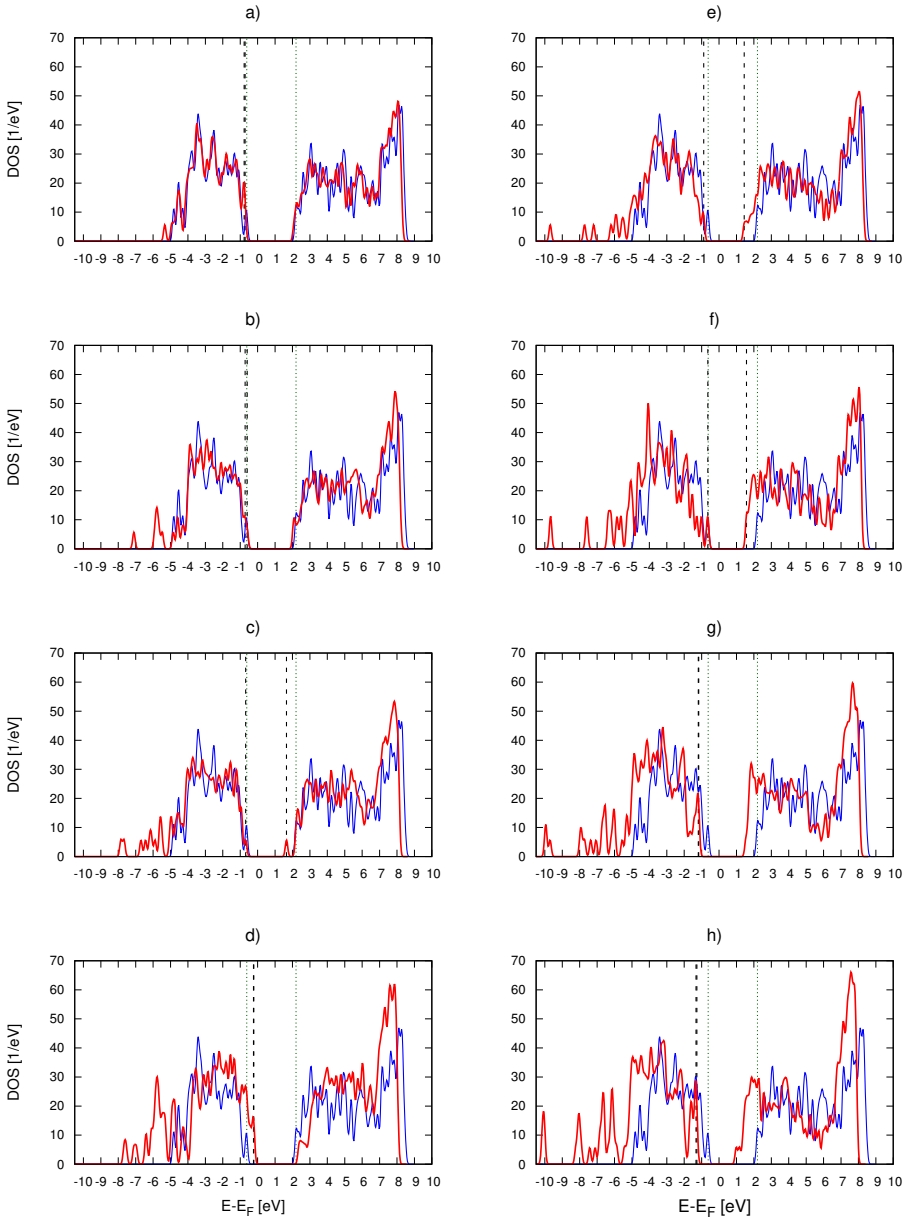



**Fig. 6. DOS for a) $Ti_{16}O_{32}$ with 1 OH group, b) $Ti_{16}O_{32}$ with 3 OH groups, c) with 6 OH groups, d) 16 OH groups, e) 1 $SO_4$ group, f) 2 $SO_4$ groups, g) 3 $SO_4$ groups, and h) $Ti_{16}O_{32}$ with 4 $SO_4$ groups relative to the Fermi level of bare $Ti_{16}O_{32}$. HOMO and LUMO levels are plotted with black vertical dashed lines. In a), b), d), g), and h) HOMO and LUMO levels almost coincide. DOS for the bare cluster is depicted in blue color on each plot for comparison and its HOMO and LUMO levels are plotted with green dots. Reproduced from Papers I–II with permission from the PCCP Owner Societies.**

symmetrical 16 OH structure. However, the empty state in the band gap is introduced at around 1.5 eV due to the d-orbitals of Ti10 and Ti5 (Fig. 4 e). The coordination of these atoms has changed during the relaxation and they have an exceptional triple O coordination environment next to it (O9-O6-O29). The LDOS of the O atoms in the triple O bond and neighboring Ti atoms are depicted in Fig. 7, which shows strong hybridization of the p-orbitals of these oxygens, supporting the formation of the triple O bond. For fully OH saturated cluster both HOMO and LUMO are on the p-orbitals of O9-O20 from $\eta^2$-$O_2$ (Fig. 4 g, h). These oxygens have very strong hybridized, almost molecule-like p-orbitals (Fig. 8). The actual band gap estimated from the DOS peaks in the top of the VB and the bottom of the CB in the case of adsorption of 1 and 3 OH is 0.207 eV and 0.042 eV bigger than in the bare cluster. Adsorption of 6 and 16 OH groups lowers the band gap on average by 0.482 eV.

The adsorption of 1, 2, and 4 $SO_4$ groups lowers the band gap on average by 0.541 eV, whereas 3 $SO_4$ groups cause only a 0.08 eV lower band gap relative to the bare cluster. The adsorption of $SO_4$ groups introduces new states in the bottom of the VB which consist of orbitals of atoms in $SO_4$ groups, atoms to which $SO_4$ groups are attached, and O atoms forming O-O bonds. The increased DOS intensity of the top of the CB is due to the p-orbitals of S atoms.

In the case of the adsorption of $SO_4$ groups the small HOMO-LUMO differences appear in the structures where the $\eta^2$-$O_2$ bonds appear and the bottom state of the CB consists of the d-orbital of a Ti atom, neighboring the O-O bond. After the adsorption of 3 and 4 $SO_4$ groups both HOMO and LUMO levels lie on p-orbitals of oxygens from $\eta^2$-$O_2$. Martirez *et al.* reported the appearance of the states near the Fermi level in $SrTiO_3$ surface due to adsorbate $\eta^2$-$O_2$ orbitals (Martirez et al., 2015). The p-orbitals of the oxygens from $\eta^2$-$O_2$ have a strong localized character and very weak hybridization with d-orbitals of neighboring Ti atoms (Fig. 9–10). Detailed description of new states in the electronic structure created upon the adsorption of $SO_4$ groups as well as LDOS figures for O atoms forming O-O bonds and their Ti neighbors is collected in the Electronic Supplementary Material (ESI) to Paper II.

The average band gap for all structures is 2.564 eV, which is 0.444 eV larger than our earlier computed value for bulk anatase (2.12 eV) (Auvinen et al., 2011). This band gap broadening corresponds to previously reported values of 0.1–0.6 eV (Satoh, Nakashima, Kamikura, & Yamamoto, 2008). The electronic structure of small particles is very sensitive to their structure and bonding environment. During the relaxation of the particles they have undergone significant structural changes. The O-O bonds, triple O

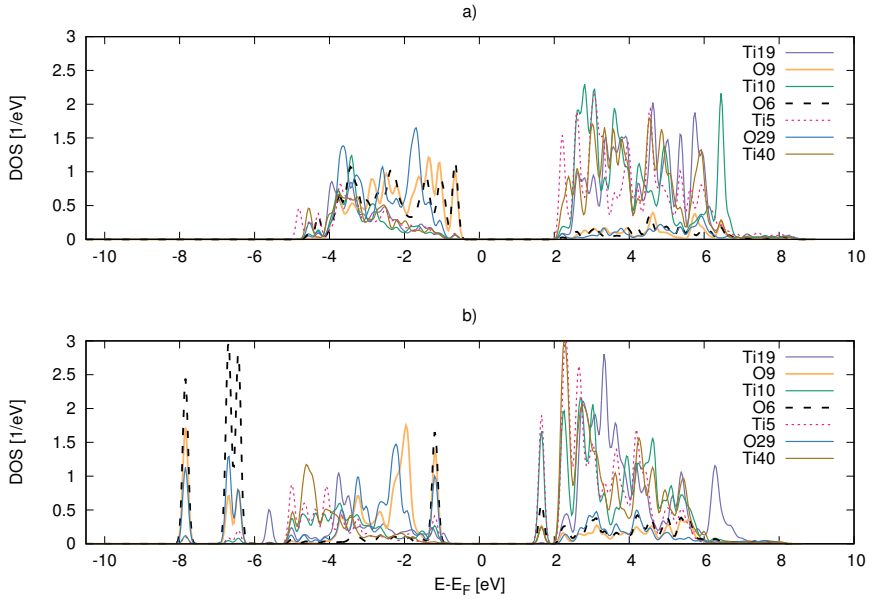



**Fig. 7. LDOS before (a) and after (b) the adsorption of 6 OH groups for the p-orbitals of O atoms, forming an O-O-O bond and the d-orbitals of neighboring Ti atoms. The atoms are numbered as in Fig. 4 e–f.**

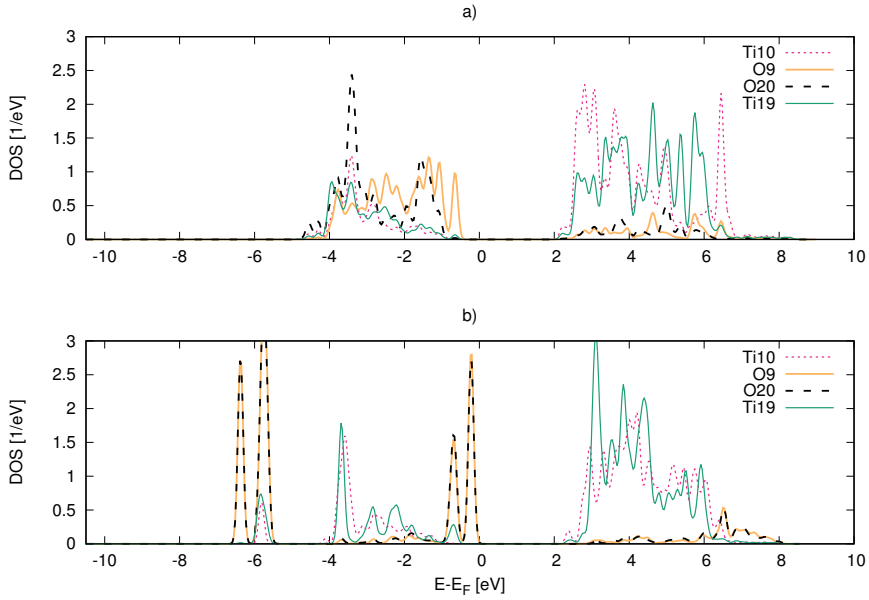



**Fig. 8. LDOS before (a) and after (b) the adsorption of 16 OH groups for the p-orbitals of O atoms, forming an O-O bond and the d-orbitals of neighboring Ti atoms. The atoms are numbered as in Fig. 4 h.**

**Table 4. HOMO-LUMO energies, corresponding atoms and orbitals, and band gaps in eV for relaxed clusters. The zero energy level is shifted to the Fermi level of the bare $Ti_{16}O_{32}$ cluster.*The HOMO-LUMO difference was calculated as the difference between the edges of valence and conduction bands. Reproduced from Papers I–II with permission from the PCCP Owner Societies.**

| Structure | HOMO | LUMO | $\Delta_{HOMO-LUMO}$ |
|---|---|---|---|
| $Ti_{16}O_{32}$ bare | -0.626 | 2.207 | 2.833 |
| | O15$p$, O6$p$ | Ti14$d$, Ti5$d$, Ti12$d$, Ti8$d$ | |
| $Ti_{16}O_{32}$+1OH | -0.776 | -0.715 | 0.061 (3.040*) |
| | O30$p$ | O30$p$ | |
| $Ti_{16}O_{32}$+3OH | -0.718 | -0.598 | 0.120 (2.875*) |
| | O34$p$, O42$p$, O17$p$ | O32$p$, O42$p$ | |
| $Ti_{16}O_{32}$+6OH | -0.686 | 1.657 | 2.343 |
| | O42$p$, O24$p$ | Ti10$d$, Ti5$d$ | |
| $Ti_{16}O_{32}$+16OH | -0.240 | -0.219 | 0.021 (2.360*) |
| | O9$p$, O20$p$ | O9$p$, O20$p$ | |
| $Ti_{16}O_{32}$+1$SO_4$ | -0.880 | 1.448 | 2.328 |
| | O25$p$, O26$p$ | Ti1$d$ | |
| $Ti_{16}O_{32}$+2$SO_4$ | -0.643 | 1.577 | 2.220 |
| | O13$p$, O7$p$, O15$p$, O6$p$ | Ti8$d$, Ti12$d$ | |
| $Ti_{16}O_{32}$+3$SO_4$ | -1.190 | -1.167 | 0.023 (2.753*) |
| | O7$p$, O6$p$, O17$p$ | O6$p$, O7$p$ | |
| $Ti_{16}O_{32}$+4$SO_4$ | -1.320 | -1.274 | 0.046 (2.327*) |
| | O9$p$, O6$p$ | O9$p$, O6$p$ | |

bonds, and $\eta^2$-$O_2$ formed have introduced defect states into the electronic structure of particles, resulting in very small HOMO-LUMO gaps, which should be reflected in the optical spectra of the particles, as we shall see in the next section.

### *3.1.3 Optical properties*

The photoabsorption spectra were calculated with the time-propagation TDDFT, the computational details can be found in Papers I–II. The spectra for all relaxed structures

in $x, y, z$ coordinate directions and the total averaged spectra are depicted in Fig. 11. The absorption shoulder starts at about 3 eV for all structures, which is consistent with the literature values of 3–3.55 eV for bigger nanoparticles (Karthick et al., 2011; Li, White, & Lim, 2004; Pugachevskii, 2013; Reddy, Reddy, & Manorama, 2001). The adsorbed species do not affect the position of the absorption peak. However, they change the intensity of the absorption, especially in the lower and higher energy regions (3–9 eV and 12–30 eV). The features of the electronic structure are reflected also in the optical spectrum: for the structures with a small HOMO-LUMO gap, absorption starts at very low energies (see insets of Fig. 11).

In both cases the most photoactive direction is the $z$-coordinate direction (intensity is about 14 $eV^{-1}$ and 12 $eV^{-1}$ for the clusters with OH and $SO_4$ groups, respectively). The total averaged refractive index functions (RIFs) derived from the photoabsorption spectra for the structures with OH groups on the surface are presented in Paper I.

The changes in the intensity of absorption and absorption peaks in the low energy region should be taken into account in light scattering methods that are often used for

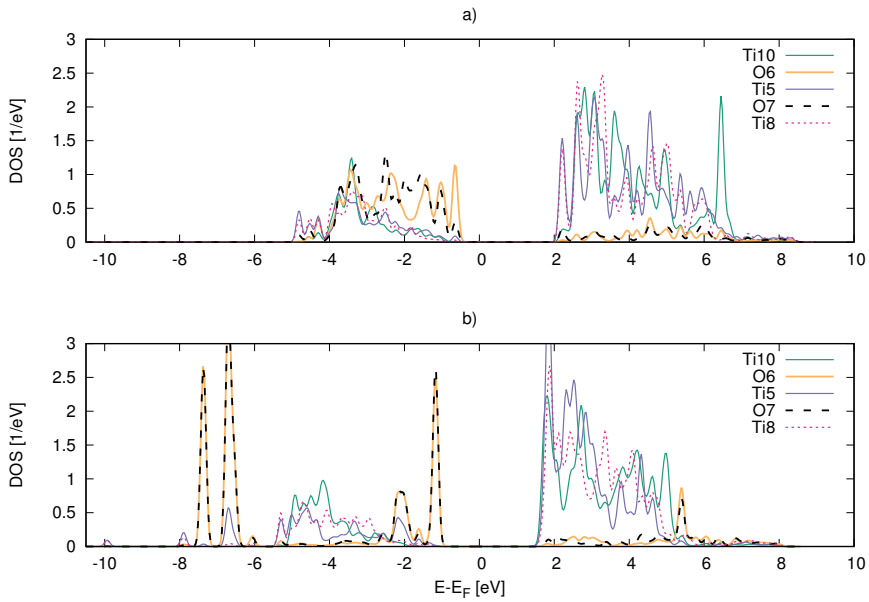



**Fig. 9. LDOS before (a) and after (b) the adsorption of $3SO_4$ groups for the p-orbitals of O atoms, forming an O-O bond and the d-orbitals of neighboring Ti atoms. The atoms are numbered as in Fig. 5 g–h. Reproduced from Paper II with permission from the PCCP Owner Societies.**

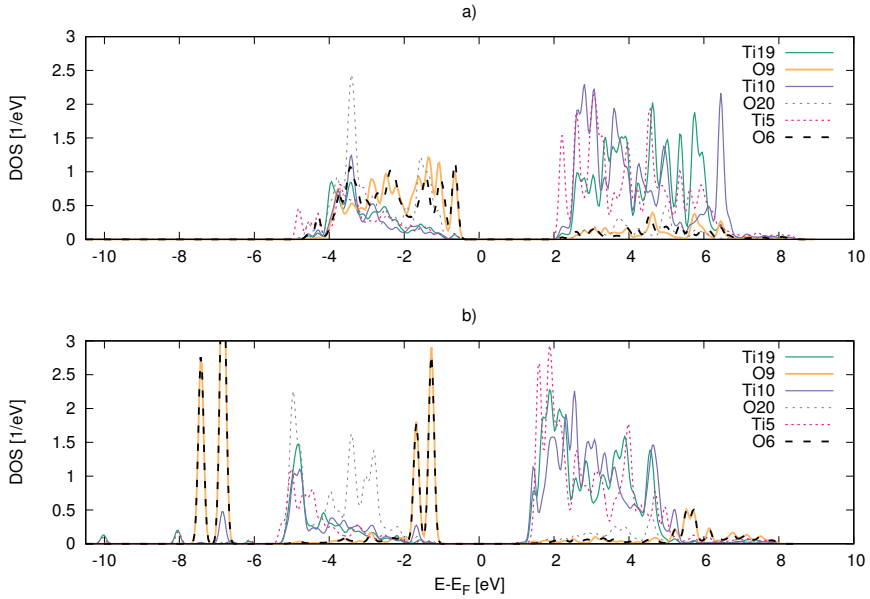



**Fig. 10. LDOS before (a) and after (b) the adsorption of $4SO_4$ groups for the p-orbitals of O atoms, forming an O-O bond, and the d-orbitals of neighboring Ti atoms. The atoms are numbered as in Fig. 5 j–k. Reproduced from Paper II with permission from the PCCP Owner Societies.**

quality and size control during manufacturing. It is worth noting that this method of calculation of optical properties does not take into account possible excitonic effects. Recently it was shown that bound excitons do exist in $TiO_2$ nanoparticles (Baldini et al., 2017). Their binding energy can be calculated by using many-body perturbation theory or by using TDDFT with more advanced XC kernels (Turkowski, Din, & Rahman, 2017).

## 3.2 Oxidation states of binary oxides from data analytics of the electronic structure

Paper III aims to find a correlation between the OSs of cations in binary oxides and Bader charges. In the basis of our analysis is an assumption that the OS of the anion, or in our case oxygen's OS, in binary oxides equals -2. From this we can evaluate the OS per compound and OS of cation in the following way

$$OS\ per\ compound = 2 * number\ of\ anions \tag{15}$$

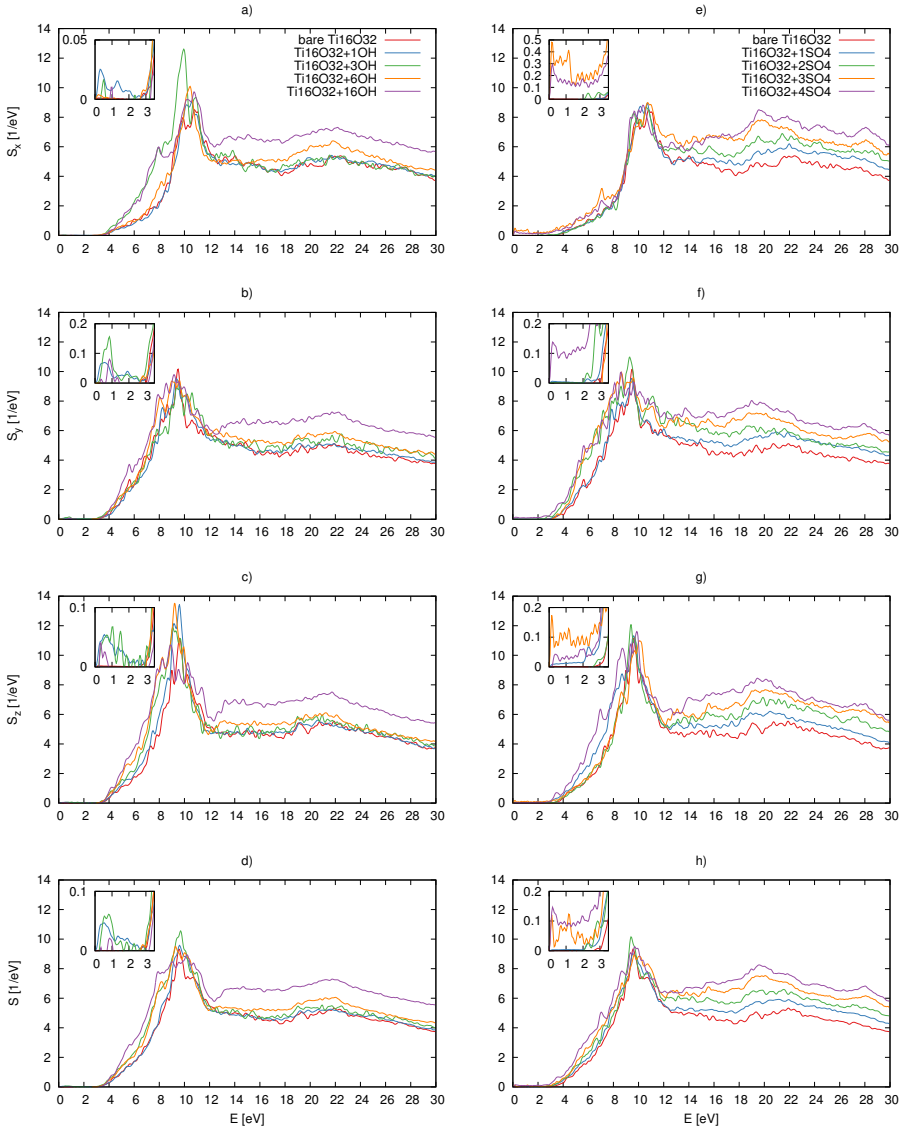



**Fig. 11. a) Photoabsorption spectrum for the *x* coordinate direction, b) for the *y* coordinate direction, c) for the *z*, and d) total averaged photoabsorption spectra for bare $Ti_{16}O_{32}$ structure, $Ti_{16}O_{32}$ with 1 OH group, with 3 OH groups, 6 OH groups, and with 16 OH groups; e) Photoabsorption spectrum for the *x* coordinate direction, f) for the *y* coordinate direction, g) for the *z*, and h) total averaged photoabsorption spectra for bare $Ti_{16}O_{32}$ structure, $Ti_{16}O_{32}$ with 1 $SO_4$ group, with 2 $SO_4$ groups, 3 $SO_4$ groups, and with 4 $SO_4$ groups. Reproduced from Papers I–II with permission from the PCCP Owner Societies.**

$$OS\ of\ cation = \frac{OS\ per\ compound}{number\ of\ cations}. \tag{16}$$

Data for this study was downloaded from the AFLOW database. It was curated and split in two sets of SV and MV compounds using procedures described in detail in Appendix to Paper III. The resulting data sets consist of total 697 compounds or 10743 atoms.

We found that for each binary, the Bader charges form non intersecting intervals corresponding to different OSs (Figure 2 in Paper III). Cations in one OS have Bader charges that differ by less than 0.03e in most cases and always less than 0.1e (Figure 17 from Appendix in Paper III). On the other hand, for cations to be in consecutive OSs difference of 0.025e Bader charge is enough.

Investigation of SV compounds showed a linear dependence between Bader charges and OSs. Linear regression plots for binaries allow to predict the OS, as can be seen from Figure 4 in Paper III and from Figure S1 in Supporting information to Paper III.

The applicability of the results to calculations with different exchange-correlation functionals was studied by test Bader charge calculations for several $Ti_xO_y$ binaries as usual benchmarking material, where Ti atom is in +1, +2, +3, and +4 OSs. Discrepancy between different functionals was found to be 2–9% from PBE+U values. We assume that similar discrepancy should be in the case of other binaries and results obtained with different functionals can be compared.

Characterization of cations in a MV compound is a nontrivial task. A workflow for the process of assigning OS in MV compounds is depicted in Figure 7 of the Paper III. MV compounds were investigated by means of Bader charges bisected intervals. These intervals are based on intervals from investigations of the SV compounds and are being adjusted by Bader charges of cations in MV compounds that fall into the earlier defined intervals. The procedure for obtaining and adjusting the bisected intervals is thoroughly explained in the Supporting information for Paper III. The OS of cations are assumed to be defined if all cation's Bader charges of a compound lie in the bisected intervals and the sum of OSs of each atom in the compound is equal to the OS per compound (see eq. (15)).

The presented workflow for characterization of the MV compounds works well for many binaries. We have been able to define the OS of cations in 49 MV compounds out of 87 records in our dataset. For example, for the $Ti_xO_y$ compound in our curated data set there are 18 MV compounds. OSs of cations from 12 $Ti_xO_y$ MV compounds

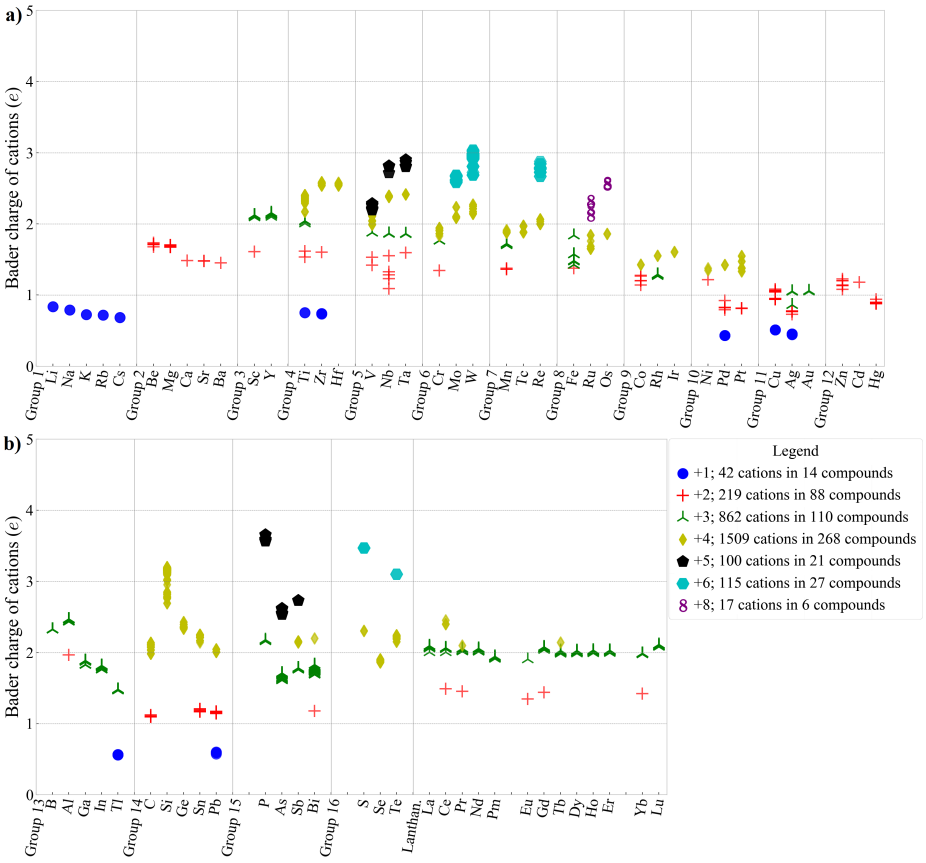



**Fig. 12. Bader charges for cations. Each point on the graph is a Bader charge for one atom. The numbers of atoms are in the legend. Atoms are grouped according to the Mendeleev table. a. Groups 1–12; b. Groups 13–16 and lanthanoids. Reproduced from Paper III with permission from Elsevier B.V.**

were successfully determined. However, for $Fe_xO_y$ this procedure was not appropriate, since none of the 10 MV compounds has been defined. Graphs for these results for MV compounds are presented in Figures 8–10 in Paper III and in Figures S3–S21 in Supporting Information for Paper III. A summary of MV compounds can be found in Table S1 in Supporting Information for Paper III.

MV compounds with OS of cation lower than +1 were disregarded from investigation due to difficulty in assigning 0 or +1 OS to cations. Charge in such compounds is spread among the cations, we assumed their OS is a fractional number.

We have analyzed Bader charges of O atoms in SV and MV compounds together and also outliers that were disregarded due to short O-O distance (less than 1.64 Å). Distribution of Bader charges of O atoms is depicted in Figure 14 in Paper III. We can conclude that O in -1 or 0 OS can have Bader charges in the interval [-0.96, 0.05]e, oxygens with lower Bader charges are in OS -2.

Analysis of maximal Bader charge difference between cations in the same OS for both SV and MV datasets (Figure 15 in Paper III) resulted in a "rule of thumb": with high probability cations of the same element in the binary oxides are in different OSs if maximal Bader charge difference is more than 0.1e.

With the proposed method we can determine the OS of cation in most binaries fast by using computed charges. It is worth noting that the proposed method is valid only for binary oxides. If we compare other materials, for example nitrides (Piskorska-Hommel, Winiarski, Kunert, & Hommel, 2017), we can see that Bader charges of 1.12e and 1.47e are assigned to Mn cations in OS +2 and +3 respectively. These Bader charge values for Mn cations in our work would fall in the intervals for OS +1 and +2, which is a disagreement. This disagreement arises from our assumption, that anion should be in -2 OS, which is not the case in nitrides.

### 3.2.1 Application to $TiO_2$ surface

In AFLOW most of the data is for periodic bulk compounds. It was interesting to test our method for other types of systems. For this purpose we chose the $TiO_2$ (110) rutile surface, where few types of Ti atoms are present in different OS (Fig. 13). This particular surface was chosen because of the availability of experimental and theoretical results for the Ti atom's OS. The Bader charges do not necessary have the same correlation with OS for the uppermost surface atoms, due to the environment of the surface atoms, different from the bulk.

Pure rutile (110) surface, surface with Ti adatom, and surface with one O vacancy were relaxed, and consequently the Bader charge analysis was performed. The OSs of Ti atoms were determined from bisected intervals, defined after analysis of SV and MV compounds (Figure 12 in Paper III). Bader charges for pure $TiO_2$ surface lie in an interval between 2.329e and 2.374e, which correspond to the bisected interval for OS +4.

The combined theoretical and experimental study by Nolan *et al.* (Nolan et al., 2008) showed that Ti adatom on rutile (110) surface is reduced to +2 OS and the neighboring

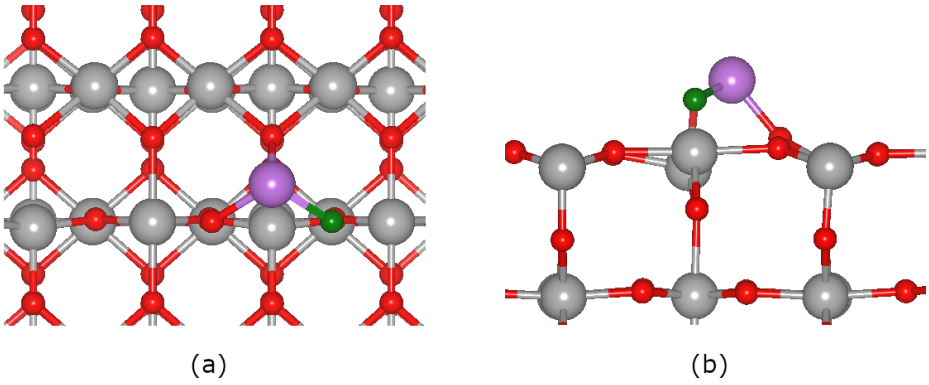



**Fig. 13. Rutile (110) surface (a) top view, (b) side view. Oxygen atoms are marked with red color, titanium atoms as gray. The oxygen vacancy is in green, titanium adatom in purple. Reproduced from Paper III with permission from Elsevier B.V.**

surface 5-fold coordinated Ti atoms is reduced to +3. Our computed Bader charges for these Ti atoms (1.686e for adatom, 2.132e for surface Ti) fall into intervals for +2 and +3 OSs respectively (Figure 12 in Paper III). The rest of the surface titanium atoms have computed Bader charges that fall into interval for OS +4.

An experimental study by Lu, Linsebigler, and Yates Jr (1994) reported that Ti is reduced to OS +3 near an O vacancy. According to our calculation, the Bader charges of Ti atoms neighboring the O vacancy are not less than 2.303e, which falls into interval for OS +4. In this case our method is not in agreement with experimental data. Apparently, for materials with vacancies more advanced methods for defining OS should be used.

# 4 Conclusions

A study of the effect of varying number of OH and $SO_4$ groups on anatase-structured $TiO_2$ nanoparticles by means of DFT and TDDFT has been performed. It showed that adsorbates introduce changes into the geometrical and electronic structure of $TiO_2$ clusters, resulting in altered photoabsorption characteristics. Since OH and $SO_4$ groups are typically present on the surface of nanoparticles during manufacture and in applications, the obtained results are important when comparing experimental data with computed values.

Changes in the properties of particles are important to be taken into account during the size and quality control procedures when manufacturing $TiO_2$ particles for different applications. As quality control is often performed by means of light scattering methods, detailed knowledge of optical properties of particles is needed. It was shown earlier that optical constants for nano-$TiO_2$ are different from bulk and are sensitive to the geometrical structure of particles (Auvinen, Alatalo, Haario, Vartiainen, & Jalava, 2013). Computed optical spectra of nanoparticles exhibit large blueshifts as compared to experimental bulk values (Auvinen et al., 2013). The blueshifts exceed the ones observed in experimental spectra of $TiO_2$ particles (J.-P. Jalava, personal communication). Our work showed that this blueshift is reduced by adsorbates. The adsorption of OH and $SO_4$ groups caused on average the enlargement of particle sizes, which also should be taken into account during size control. The adsorption of 3 and $4SO_4$ groups causes side-on coordinated $\eta^2$-$O_2$ species to appear on the surface, which is interesting, because they are intermediate species in the production of molecular oxygen from water on transition metal based oxides (Martirez et al., 2015).

The information obtained will be useful both in the manufacturing process of $TiO_2$nanoparticles and designing new applications. For example, in hydrogen energy applications, a $TiO_2$ surface covered with OH groups is needed (Du et al., 2012). In a similar manner, $TiO_2$ nanoparticles can be functionalized, given that the effect of functional groups on their properties is well understood. Deeper understanding of the electronic and optical properties could improve the quality of the $TiO_2$ products and efficiency in the photocatalytic applications of $TiO_2$ clusters.

This study was an attempt to address the properties of $TiO_2$ nanoparticles in solutions, providing insights into how the environment affects the particles. In this work only

stoichiometric clusters were studied to reduce the effect of defects on the electronic structure, to concentrate on the effect of adsorbates. Including them in the model will help to achieve more realistic conditions. Further work direction could be to consider excitonic effects, that were shown recently to be present in $TiO_2$ particles (Baldini et al., 2017). For achieving this, methods beyond ordinary TDDFT should be used.

A correlation between oxidation states and Bader charges was demonstrated for a large set of binary oxides as compounds that comprise the majority in online material databases. The concepts of oxidation states and Bader charge are widely used in the analysis of experimental data and theoretical calculations for gaining physical and chemical insights. However, a direct connection between them has been missing so far. Our results establish a correlation and provide an easy and fast approach to theoretical OS determination from DFT calculations that are routinely carried out. Additionally it can be used for the inverse task - to find compounds where an atom is in a particular OS. The proposed method can be incorporated to existing tools and applications in databases widening their functionality. Fast OS determination can be useful in developing the Material Genome.

The definition of OS varies from field to field and raised debates in the literature. Our work has shown the connection of the OS to the electron density of the material. However, our proposed correlation does not work in all cases, which could be attributed to the fact that there is no unambiguous way to partition the electron density between atoms. Another issue is the validity of DFT calculations for systems with strong electron correlations. Our results have been shown to depend weakly on the choice of the exchange-correlation functionals. Future work can be done to widen the results to other compounds available in open databases, to surfaces and clusters, and taking into account other properties of materials available, for example XAS spectra (Zheng et al., 2018).

# 5 Corrigenda

During the preparation of this thesis, errors in the published articles were discovered. This chapter addresses these errors.

In Paper I Table 3 the percentage change of dimension $x$ of cluster with 16 OH groups must be 15, not 0.15. This typo does not affect our conclusions as the average values are correct.

In Paper I Table 4 the number of 4-fold coordinated Ti atoms in the cluster with 1 OH group should be two and the number of 3-fold coordinated should be zero. This miscalculation does not affect our conclusions as the total number of undercoordinated Ti atoms in this structure remains the same.

# References


Akbari, A., Hashemi, J., Niskanen, J., Huotari, S., & Hakala, M. (2015). Identification of the dye adsorption modes in dye-sensitised solar cells with X-ray spectroscopy techniques: a computational study. *Phys. Chem. Chem. Phys.*, *17*(16), 10849–10855. doi: 10.1039/C4CP05980H

Asahi, R. (2001). Visible-Light Photocatalysis in Nitrogen-Doped Titanium Oxides. *Science*, *293*(5528), 269–271. doi: 10.1126/science.1061051

Auvinen, S., Alatalo, M., Haario, H., Jalava, J.-P., & Lamminmäki, R.-J. (2011). Size and shape dependence of the electronic and spectral properties in $TiO_2$ nanoparticles. *J. Phys. Chem. C*, *115*, 8484–8493. doi: 10.1021/jp112114p

Auvinen, S., Alatalo, M., Haario, H., Vartiainen, E., & Jalava, J.-P. (2013). Refractive Index Functions of $TiO_2$ Nanoparticles. *J. Phys. Chem. C*, *117*(7), 3503–3512. doi: 10.1021/jp305303q

*Avogadro: an open-source molecular builder and visualization tool. version 1.2.0.* (2019, May 03). Retrieved from `http://avogadro.cc/`

Bader, R. F. W. (1990). *Atoms in molecules: A quantum theory*.

Bahn, S. R., & Jacobsen, K. W. (2002). An object-oriented scripting interface to a legacy electronic structure code. *Comput. Sci. Eng.*, *4*(3), 56–66. doi: 10.1109/5992.998641

Baldini, E., Chiodo, L., Dominguez, A., Palummo, M., Moser, S., Yazdi-Rizi, M., … Chergui, M. (2017). Strongly bound excitons in anatase $TiO_2$ single crystals and nanoparticles. *Nat. Commun.*, *8*(1), 1–40. doi: 10.1038/s41467-017-00016-6

Bare, S. R., Kelly, S. D., Vila, F. D., Boldingh, E., Karapetrova, E., Kas, J., … Rehr, J. J. (2011). Experimental (XAS, STEM, TPR, and XPS) and Theoretical (DFT) Characterization of Supported Rhenium Catalysts. *J. Phys. Chem. C*, *115*, 5740–5755. doi: 10.1021/jp1105218

Blagojevic, V., Chen, Y.-R., Steigerwald, M., Brus, L., & Friesner, R. A. (2009). Quantum chemical investigation of cluster models for $TiO_2$ nanoparticles with water-derived ligand passivation: Studies of excess electron states and implications for charge transport in the gratzel cell. *The Journal of Physical Chemistry C*, *113*(46), 19806–19811. doi: 10.1021/jp905332z

Blöchl, P. E. (1994). Projector augmented-wave method. *Phys. Rev. B*, *50*, 17953–17979. doi: 10.1103/PhysRevB.50.17953

Blöchl, P. E., Först, C. J., & Schimpl, J. (2003). Projector augmented wave method: ab initio molecular dynamics with full wave functions. *Bull. Mater. Sci.*, *26*, 33–41. doi: 10.1007/BF02712785

Born, M., & Oppenheimer, R. (1927). Zur quantentheorie der molekeln. *Annalen der Physik*, *389*(20), 457–484. doi: 10.1002/andp.19273892002

Bredow, T., & Jug, K. (1995). Theoretical investigation of water-adsorption at rutile and anatase surfaces. *Surf. Sci.*, *327*, 398–408. doi: 10.1016/0039-6028(94)00851-5

Brown, I. D. (2016). *The chemical bond in inorganic chemistry: the bond valence model* (2nd ed.). Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780198742951.001.0001

Calderon, C. E., Plata, J. J., Toher, C., Oses, C., Levy, O., Fornari, M., … Curtarolo, S. (2015). The AFLOW standard for high-throughput materials science calculations. *Computational Materials Science*, *108*, 233–238. doi: https://doi.org/10.1016/j.commatsci.2015.07.019

Cho, D., Ko, K. C., Lamiel-García, O., Bromley, S. T., Lee, J. Y., & Illas, F. (2016). Effect of Size and Structure on the Ground-State and Excited-State Electronic Structure of $TiO_2$ Nanoparticles. *J. Chem. Theory Comput.*, *12*(8), 3751–3763. doi: 10.1021/acs.jctc.6b00519

Cramer, C. J., Tolman, W. B., Theopold, K. H., & Rheingold, A. L. (2003). Variable character of O-O and M-O bonding in side-on ($\eta 2$) 1:1 metal complexes of $O_2$. *Proc. Natl. Acad. Sci. U.S.A.*, *100*(7), 3635–3640. doi: 10.1073/pnas.0535926100

Curtarolo, S., Setyawan, W., Hart, G. L., Jahnatek, M., Chepulskii, R. V., Taylor, R. H., … Morgan, D. (2012). AFLOW: An automatic framework for high-throughput materials discovery. *Comput. Mater. Sci.*, *58*, 218–226. doi: 10.1016/j.commatsci.2012.02.005

Curtarolo, S., Setyawan, W., Wang, S., Xue, J., Yang, K., Taylor, R. H., … Levy, O. (2012). AFLOWLIB.ORG: A distributed materials properties repository from high-throughput ab initio calculations. *Comput. Mater. Sci.*, *58*, 227–235. doi: 10.1016/J.COMMATSCI.2012.02.002

Day, P., Hush, N. S., & Clark, R. J. (2008). Mixed valence: origins and developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *366*(1862), 5–14. doi: 10.1098/rsta.2007.2135

Diebold, U. (2003). The surface science of titanium dioxide. *Surface Science Reports*, *48*(5), 53–229. doi: 10.1016/S0167-5729(02)00100-0

Dirac, P. A. M. (1930). Note on Exchange Phenomena in the Thomas Atom. *Mathematical Proceedings of the Cambridge Philosophical Society*, *26*(3), 376–385. doi: 10.1017/S0305004100016108

Du, Y., Petrik, N. G., Deskins, N. A., Wang, Z., Henderson, M. A., Kimmel, G. A., & Lyubinetsky, I. (2012). Hydrogen reactivity on highly-hydroxylated $TiO_2$(110) surfaces prepared via carboxylic acid adsorption and photolysis. *Phys. Chem. Chem. Phys.*, *14*(9), 3066–3074. doi: 10.1039/c1cp22515d

Dudarev, S. L., Botton, G. A., Savrasov, S. Y., Humphreys, C. J., & Sutton, A. P. (1998). Electron-energy-loss spectra and the structural stability of nickel oxide: An LSDA+U study. *Phys. Rev. B*, *57*, 1505–1509. doi: 10.1103/PhysRevB.57.1505

Enkovaara, J., Rostgaard, C., Mortensen, J. J., Chen, J., Dułak, M., Ferrighi, L., … Jacobsen, K. W. (2010). Electronic structure calculations with GPAW: a real-space implementation of the projector augmented-wave method. *J. Phys.: Condens. Matter*, *22*(25), 253202. doi: 10.1088/0953-8984/22/25/253202

Fermi, E. (1927). Statistical method to determine some properties of atoms. *Rend. Accad. Naz. Lincei*, *6*, 602–607.

Fock, V. (1930). Näherungsmethode zur lösung des quantenmechanischen mehrkörperproblems. *Zeitschrift für Physik*, *61*(1), 126–148. doi: 10.1007/BF01340294

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., … Fox, D. J. (2016). *Gaussian 16 Revision B.01.* (Gaussian Inc. Wallingford CT)

Fujishima, A., & Honda, K. (1972). Electrochemical photolysis of water at a semiconductor electrode. *Nature*, *238*(5358), 37–38. doi: 10.1038/238037a0

Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., & Hutchison, G. R. (2012). Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics*, *4*(1), 17. doi: 10.1186/1758-2946-4-17

Hartree, D. (1928). The wave mechanics of an atom with a non-Coulombic central field. Part I. In *Math. Proc. Cambridge Phil. Soc* (Vol. 24, pp. 89–110). doi: 10.1017/S0305004100011919

Hashimoto, K., Irie, H., & Fujishima, A. (2005). $TiO_2$ Photocatalysis: A Historical Overview and Future Prospects. *Jpn. J. Appl. Phys.*, *44*(12), 8269–8285.

Heckmann, A., & Lambert, C. (2012). Organic mixed-valence compounds: A playground for electrons and holes. *Angewandte Chemie International Edition*, *51*(2), 326–392. doi: 10.1002/anie.201100944

Henderson, M. A. (2011). A surface science perspective on $TiO_2$ photocatalysis. *Surf. Sci. Rep.*, *66*(6-7), 185–297. doi: 10.1016/j.surfrep.2011.01.001

Henkelman, G. (2017, September 25). *Bader charge analysis.* Retrieved from `http://theory.cm.utexas.edu/henkelman/code/bader/`

Henkelman, G., Arnaldsson, A., & Jónsson, H. (2006). A fast and robust algorithm for bader decomposition of charge density. *Comp. Mater. Sci.*, *36*(3), 354–360. doi: 10.1016/j.commatsci.2005.04.010

Hirshfeld, F. L. (1977). Bonded-atom fragments for describing molecular charge densities. *Theoretica chimica acta*, *44*(2), 129–138. doi: 10.1007/BF00549096

Hohenberg, P., & Kohn, W. (1964). Inhomogeneous electron gas. *Phys. Rev.*, *136*, B864–B871. doi: 10.1103/PhysRev.136.B864

Horn, M., Schwebdtfeger, C., & Meagher, E. (1972). Refinement of the structure of anatase at several temperatures. *Zeitschrift für Kristallographie-Crystalline Materials*, *136*(1–6), 273–281.

Jalava, J.-P. (2000). *Formation of $TiO_2$ pigment particles in the sulphate process – a methodological study* (PhD thesis). University of Turku.

Jalava, J.-P., Taavitsainen, V.-M., Lamminmäki, R.-J., Lindholm, M., Auvinen, S., Alatalo, M., ... Haario, H. (2015). Modeling $TiO_2$s refractive index function from bulk to nanoparticles. *J. Quant. Spectrosc. Radiat. Transf.*, *167*, 105–118. doi: 10.1016/J.JQSRT.2015.08.007

Jansen, M., & Wedig, U. (2008). A Piece of the Picture-Misunderstanding of Chemical Concepts. *Angew. Chemie Int. Ed.*, *47*(52), 10026–10029. doi: 10.1002/anie.200803605

Jiang, L., Levchenko, S. V., & Rappe, A. M. (2012). Rigorous Definition of Oxidation States of Ions in Solids. *Phys. Rev. Lett.*, *108*(16), 166403. doi: 10.1103/PhysRevLett.108.166403

Kamiński, T., Gottlieb, C. A., Menten, K. M., Patel, N. A., Young, K. H., Brünken, S., ... Decin, L. (2013). Pure rotational spectra of TiO and $TiO_2$ in VY Canis Majoris. *A&A*, *551*(A113), 1–13. doi: 10.1051/0004-6361/201220290

Karen, P., McArdle, P., & Takats, J. (2014). Toward a comprehensive definition of oxidation state (IUPAC technical report). *Pure and Applied Chemistry*, *86*(6), 1017–1081.

Karthick, S., Prabakar, K., Subramania, A., Hong, J.-T., Jang, J.-J., & Kim, H.-J. (2011). Formation of anatase $TiO_2$ nanoparticles by simple polymer gel technique and their properties. *Powder Technol.*, *205*(1–3), 36–41. doi: 10.1016/j.powtec.2010.08.061

Koch, D., & Manzhos, S. (2017). On the Charge State of Titanium in Titanium Dioxide. *J. Phys. Chem. Lett.*, *8*(7), 1593–1598. doi: 10.1021/acs.jpclett.7b00313

Kohn, W., & Sham, L. J. (1965a). Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, *140*, A1133–A1138. doi: 10.1103/PhysRev.140.A1133

Kohn, W., & Sham, L. J. (1965b). Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.*, *140*(4A), 1133–1138. doi: 10.1103/PhysRev.140.A1133

Kresse, G., & Furthmüller, J. (1996). Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B*, *54*, 11169–11186. doi: 10.1103/PhysRevB.54.11169

Kresse, G., & Joubert, D. (1999). From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B*, *59*, 1758–1775. doi: 10.1103/PhysRevB.59.1758

Li, Y., White, T., & Lim, S. (2004). Low-temperature synthesis and microstructural control of titania nano-particles. *J. Solid State Chem.*, *177*(4), 1372–1381. doi: 10.1016/j.jssc.2003.11.016

Liechtenstein, A. I., Anisimov, V. I., & Zaanen, J. (1995). Density-functional theory and strong interactions: Orbital ordering in Mott-Hubbard insulators. *Phys. Rev. B*, *52*, R5467–R5470. doi: 10.1103/PhysRevB.52.R5467

Liu, X., Li, Y., Peng, S., Lu, G., & Li, S. (2013). Modification of $TiO_2$ with sulfate and phosphate for enhanced eosin Y-sensitized hydrogen evolution under visible light illumination. *Photochem. Photobiol. Sci.*, *12*(10), 1903–1910. doi: 10.1039/c3pp50167a

Löwdin, P. (1950). On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals. *The Journal of Chemical Physics*, *18*(3), 365–375. doi: 10.1063/1.1747632

Lu, G., Linsebigler, A., & Yates Jr, J. T. (1994). Ti3+ defect sites on $TiO_2$ (110): production and chemical detection of active sites. *The Journal of Physical Chemistry*, *98*(45), 11733–11738. doi: 10.1021/j100096a017

Martin, R. M. (2004). *Electronic structure: Basic theory and practical methods*. Cambridge University Press. doi: 10.1017/CBO9780511805769

Martirez, J. M. P., Kim, S., Morales, E. H., Diroll, B. T., Cargnello, M., Gordon, T. R., ... Rappe, A. M. (2015). Synergistic oxygen evolving activity of a $TiO_2$-rich reconstructed $SrTiO_3$(001) surface. *J. Am. Chem. Soc.*, *137*(8), 2939–2947. doi: 10.1021/ja511332y

*Materials genome initiative.* (2019, April 3). Retrieved from `https://www.mgi.gov/`

Mo, S. D., & Ching, W. Y. (1995). Electronic and optical properties of three phases of titanium dioxide: Rutile, anatase, and brookite. *Phys. Rev. B*, *51*(19), 13023–13032. doi: 10.1103/PhysRevB.51.13023

Mortensen, J. J., Hansen, L. B., & Jacobsen, K. W. (2005). Real-space grid implementation of the projector augmented wave method. *Phys. Rev. B*, *71*(3), 035109. doi: 10.1103/PhysRevB.71.035109

Mosey, N. J., Liao, P., & Carter, E. A. (2008). Rotationally invariant ab initio evaluation of coulomb and exchange parameters for DFT+U calculations. *The Journal of Chemical Physics*, *129*(1), 014103. doi: 10.1063/1.2943142

Mulliken, R. S. (1955). Electronic population analysis on LCAO–MO molecular wave functions. I. *The Journal of Chemical Physics*, *23*(10), 1833–1840. doi: 10.1063/1.1740588

Muscat, J., Swamy, V., & Harrison, N. M. (2002). First-principles calculations of the phase stability of $TiO_2$. *Phys. Rev. B*, *65*, 224112. doi: 10.1103/PhysRevB.65.224112

Nolan, M., Elliott, S. D., Mulley, J. S., Bennett, R. A., Basham, M., & Mulheran, P. (2008). Electronic structure of point defects in controlled self-doping of the $TiO_2$ (110) surface: Combined photoemission spectroscopy and density functional theory study. *Phys. Rev. B*, *77*, 235424. doi: 10.1103/PhysRevB.77.235424

Peng, Y.-H., Huang, G.-F., & Huang, W.-Q. (2012). Visible-light absorption and photocatalytic activity of Cr-doped $TiO_2$ nanocrystal films. *Adv. Powder Technol.*, *23*(1), 8–12. doi: 10.1016/J.APT.2010.11.006

Perdew, J. P., Burke, K., & Ernzerhof, M. (1996). Generalized gradient approximation made simple. *Phys. Rev. Lett.*, *77*(18), 3865–3868. doi: 10.1103/PhysRevLett.77.3865

Persson, P., Gebhardt, J. C. M., & Lunell, S. (2003). The smallest possible nanocrystals of semiionic oxides. *J. Phys. Chem. B*, *107*(15), 3336–3339. doi: 10.1021/jp022036e

Piskorska-Hommel, E., Winiarski, M., Kunert, G., & Hommel, D. (2017). Polarization-dependent XAFS and density functional theory investigations of the quality of the epitaxial GaMnN structure. *Journal of Alloys and Compounds*, *725*, 632–638. doi: 10.1016/j.jallcom.2017.07.174

Postils, V., Delgado-Alonso, C., Luis, J. M., & Salvador, P. (2018). An Objective Alternative to IUPAC's Approach To Assign Oxidation States. *Angew. Chemie*, *130*(33), 10685–10689. doi: 10.1002/anie.201802745

Pugachevskii, M. (2013). Ultraviolet absorption spectrum of laser-ablated titanium dioxide nanoparticles. *Tech. Phys. Lett.*, *39*(1), 36–38. doi: 10.1134/S1063785013010239

Raebiger, H., Lany, S., Resta, R., & Zunger, A. (2009). Oxidation numbers as Social Security Numbers: Are they predictive or postdictive? *Nat. Preced.*. Retrieved from `http://hdl.handle.net/10101/npre.2009.4012.1`

Raebiger, H., Lany, S., & Zunger, A. (2008). Charge self-regulation upon changing the oxidation state of transition metals in insulators. *Nature*, *453*(7196), 763–766. doi: 10.1038/nature07009

Raj, K. J. A., Shanmugam, R., Mahalakshmi, R., & Viswanathan, B. (2010). XPS and IR spectral studies on the structure of phosphate and sulphate modified titania - A combined DFT and experimental study. *Indian J Chem A*, *49*, 9–17.

Rappe, A. K., Casewit, C. J., Colwell, K. S., Goddard III, W. A., & Skiff, W. M. (1992). UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.*, *114*, 10024–10035. doi: 10.1021/ja00051a040

Reddy, K. M., Reddy, C. G., & Manorama, S. (2001). Preparation, characterization, and spectral studies on nanocrystalline anatase $TiO_2$. *J. Solid State Chem.*, *158*(2), 180–186. doi: 10.1006/jssc.2001.9090

Reed, A. E., Weinstock, R. B., & Weinhold, F. (1985). Natural population analysis. *The Journal of Chemical Physics*, *83*(2), 735–746. doi: 10.1063/1.449486

Resta, R. (2008). Charge states in transition. *Nature*, *453*(7196), 735. doi: 10.1038/453735a

Robin, M. B., & Day, P. (1968). Mixed Valence Chemistry – A Survey and Classification. *Adv. Inorg. Chem. Radiochem.*, *10*, 247–422. doi: 10.1016/S0065-2792(08)60179-X

Runge, E., & Gross, E. K. U. (1984). Density-Functional Theory for Time-Dependent Systems. *Phys. Rev. Lett.*, *52*(12), 997–1000.

Sanville, E., Kenny, S. D., Smith, R., & Henkelman, G. (2007). Improved grid-based algorithm for bader charge allocation. *J. Comput. Chem.*, *28*(5), 899–908. doi: 10.1002/jcc.20575

Satoh, N., Nakashima, T., Kamikura, K., & Yamamoto, K. (2008). Quantum size effect in $TiO_2$ nanoparticles prepared by finely controlled metal assembly on dendrimer templates. *Nat. Nanotechnol.*, *3*(2), 106–111. doi: 10.1038/nnano.2008.2

Sedaghati, E., Boffin, H. M. J., MacDonald, R. J., Gandhi, S., Madhusudhan, N., Gibson, N. P., ... Rauer, H. (2017). Detection of titanium oxide in the atmosphere of a hot Jupiter. *Nature*, *549*(7671), 238–241. doi: 10.1038/nature23651

Tang, W., Sanville, E., & Henkelman, G. (2009). A grid-based bader analysis algorithm without lattice bias. *J. Phys.: Condens. Matter*, *21*(8), 084204. doi: 10.1088/0953-8984/21/8/084204

Taylor, R. H., Rose, F., Toher, C., Levy, O., Yang, K., Buongiorno Nardelli, M., & Curtarolo, S. (2014). A RESTful API for exchanging materials data in the AFLOWLIB.org consortium. *Comput. Mater. Sci.*, *93*, 178–192. doi: 10.1016/j.commatsci.2014.05.014

Thomas, L. H. (1927). The calculation of atomic fields. *Mathematical Proceedings of the Cambridge Philosophical Society*, *23*(5), 542–548. doi: 10.1017/S0305004100011683

Turkowski, V., Din, N., & Rahman, T. (2017). Time-Dependent Density-Functional Theory and Excitons in Bulk and Two-Dimensional Semiconductors. *Computation*, *5*(4), 39. doi: 10.3390/computation5030039

Ullrich, C. A. (2011). *Time-dependent density-functional theory: Concepts and applications.* OUP Oxford. doi: 10.1093/acprof:oso/9780199563029.001.0001

Varner, K., Rindfusz, K., Gaglione, A., & Viveiros, E. (2010). *State of the Science Literature Review: Nano Titanium Dioxide Environmental Matters* (Tech. Rep.). Washington, DC: U.S. Environ. Prot. Agency.

Vickerman, J. C., & Gilmore, I. S. (2009). *Surface analysis: the principal techniques* (2nd ed.). John Wiley & Sons. doi: 10.1002/9780470721582

Walsh, A., Sokol, A. A., Buckeridge, J., Scanlon, D. O., & Catlow, C. R. A. (2017). Electron Counting in Solids: Oxidation States, Partial Charges, and Ionicity. *J. Phys. Chem. Lett.*, *8*(9), 2074–2075. doi: 10.1021/acs.jpclett.7b00809

Walsh, A., Sokol, A. A., Buckeridge, J., Scanlon, D. O., & Catlow, C. R. A. (2018). Oxidation states and ionicity. *Nat. Mater.*, *17*(11), 958–964. doi: 10.1038/s41563-018-0165-7

Wyckoff, R. (1963). *Crystal structures* (2nd ed., Vol. 1). New York: Interscience Publishers.

Xu, M., Xiao, P., Stauffer, S., Song, J., Henkelman, G., & Goodenough, J. B. (2014). Theoretical and Experimental Study of Vanadium-Based Fluorophosphate Cathodes for Rechargeable Batteries. *Chem. Mater.*, *26*, 3089–3097. doi: 10.1021/cm500106w

Yabana, K., & Bertsch, G. F. (1996). Time-dependent local-density approximation in real time. *Phys. Rev. B*, *54*(7), 4484–4487. doi: 10.1103/PhysRevB.54.4484

Yabana, K., Nakatsukasa, T., Iwata, J.-I., & Bertsch, G. F. (2006). Real-time, real-space implementation of the linear response time-dependent density-functional theory. *phys. stat. sol. (b)*, *243*(5), 1121–1138. doi: 10.1002/pssb.200642005

Yang, H., Lu, R., & Wang, L. (2003). Study of preparation and properties on solid superacid sulfated titania–silica nanomaterials. *Mater. Lett.*, *57*(5–6), 1190–1196. doi: 10.1016/S0167-577X(02)00954-0

Yu, M., & Trinkle, D. R. (2011). Accurate and efficient algorithm for Bader charge integration. *J. Chem. Phys.*, *134*(6), 064111. doi: 10.1063/1.3553716

Zhang, H., & Banfield, J. F. (2014). Structural characteristics and mechanical and thermodynamic properties of nanocrystalline $TiO_2$. *Chem. Rev.*, *114*(19), 9613–9644. doi: 10.1021/cr500072j

Zheng, C., Mathew, K., Chen, C., Chen, Y., Tang, H., Dozier, A., ... Ong, S. P. (2018). Automated generation and ensemble-learned matching of X-ray absorption spectra. *npj Comput. Mater.*, *4*(1), 12. doi: 10.1038/s41524-018-0067-x

# Original publications

I Miroshnichenko O., Auvinen S., & Alatalo M. (2015). A DFT study of the effect of OH groups on the optical, electronic, and structural properties of $TiO_2$ nanoparticles. *Phys. Chem. Chem. Phys.*, *17*, 5321–5327. doi: 10.1039/c4cp02789b

II Miroshnichenko O., Posysaev S., & Alatalo M. (2016). A DFT study of the effect of $SO_4$ groups on the properties of $TiO_2$ nanoparticles. *Phys. Chem. Chem. Phys., 18*, 33068–33076. doi: 10.1039/c6cp05681d

III Posysaev S., Miroshnichenko O., Alatalo M., Le D., & Rahman T.S. (2019). Oxidation states of binary oxides from data analytics of the electronic structure. *Comput. Mater. Sci., 161*, 403–414. doi: 10.1016/j.commatsci.2019.01.046