UCLouvain

École polytechnique de Louvain

Design and investigation of new materials for optoelectronics

Author: **Victor TRINQUET**
Supervisor: **Gian-Marco RIGNANESE**
Readers: **Xavier GONZE, Guillaume BRUNIN, Francesco NACCARATO**
Academic year 2021–2022
Master [120] in Physical Engineering

Abstract

Nonlinear optical materials have become vital components in modern optoelectronics. However, various regions of the electromagnetic spectrum are still lacking appropriate compounds. In an attempt to guide experimental research, this work investigates the second-harmonic generation phenomenon from both *ab initio* and machine learning approach. The first-principles computations are performed in the framework of characterizing six promising crystals for use in the mid-infrared region: MgSi(-Ge,-Sn)P_2(As_2). Both $MgGeAs_2$ and $MgSnAs_2$ are shown to display an important second-harmonic generation with the possibility of realizing angular phase matching. Moreover, the six materials constitute an ideal system to inspect the relations between the atomic arrangement, the electronic properties, and the second-harmonic generation response thanks to modifications of their crystalline structure and to the existence of binary analogs. This richness of possibilities allows to explore the interesting concept of quantifying the non-centrosymmetry and its influence on the second-harmonic generation and on Miller's coefficient. It is found that the latter seems to be more impacted by this "degree of centrosymmetry" than the second-harmonic response is. This coefficient relates linear to nonlinear optics by a simple relation which is termed Miller's rule. Its study is thus primordial since it holds the potential of greatly simplifying the determination of nonlinear optical properties. This last ambition is shared by a data-driven approach, made possible thanks to the recent emergence of a database with the relevant features. The present essay thus explores the possibility of predicting the second-harmonic generation response by a materials-oriented machine learning model, MODNet. Although the final results are mitigated, it paves the way for future improvements.

Acknowledgments

First and foremost I would like to warmly thank my promotor, Professor Gian-Marco Rignanese. This project would not have been possible without his help and support. His enthusiasm was a great source of motivation and I am grateful to be able to continue exploring this captivating subject with him and to benefit from his expertise.

I would also like to sincerely thank Guillaume Brunin. He greatly contributed to many first-principles results and never ceased to answer my countless questions. I am also very grateful to Pierre-Paul De Breuck for his help with MODNet both when setting it up and when first using it. My gratitude extends to Francesco Naccarato for his availability and to Professor Xavier Gonze for reading this dissertation.

Finally, many thanks to my family, friends (special mention to the Fyki gang for these last few years at university), and girlfriend for their emotional support and their sympathetic ear.

Contents

Introduction

Nowadays, developing and improving technologies has become a crucial aspect in facing the challenges that await the modern world. Engineers and scientists have their role to play to solve the current crisis, whether societal or ecological. It is thus essential to consider one of the main vectors of progress: materials knowledge. Indeed, the contribution of materials innovation to the overall technological progress has become increasingly important in the last few decades. It can be speculated that it is responsible for at least 20% of the progress in all major areas and for 80% in energy storage which is one of the most important thematics of our time [1]. These statistics are not surprising seeing how many technologies (such as photovoltaic panels, lithium batteries, or hydrogen fuel cells) rely on functional materials, i.e., materials designed for a targeted application. Most of them, no matter how mature they are, could clearly benefit from the development of new compounds (e.g., defect-tolerant absorbers, solid-state electrolytes, or better catalysts). Indeed, improving the properties of the materials at their core can boost their efficiency, increase their safety towards users or even lower their environmental impact.

Unfortunately, the path to developing new functional materials is a long one and is paved with various difficulties as it is often the case when experimenting. This time scale does not incite investments which further slows progress. As a solution, theoretical and computational approaches can efficiently guide the experiments and compensate the lack of information that experimentalists often face [2]. Computational materials design based on quantum simulations such as in the framework of density functional theory (DFT) has enabled many current applications and still hold a vast potential [3]. The success of the computational field has also led to the emergence of large materials databases (DBs). They contain an incredible amount of knowledge that is yet to be extracted. As a result, statistical learning and machine learning (ML) techniques have now become attractive tools to discover, at a fraction of the computational time of quantum simulations, the hidden relations between elements and structures in compounds. Thus, the search for new functional materials can now also be data-driven and promises to speed up materials discovery [4]. Figure 0.1 summarizes the global design process of functional materials.

Optoelectronics is one of the many domains that could benefit from new materials. This highly active research field has a wide variety of applications such as communications (e.g., optical fibers), energy production (e.g., photovoltaics), lighting (e.g., LEDs), and it is foremost used in the laser industry. By itself, the latter covers a considerable range of applications such as spectroscopy, chemical sensing, surgeries, cancer detection, etc. One of its important challenges is to cover the whole frequency spectrum of electromagnetic (EM) fields in order to efficiently address the different utilization. Since the spectral coverage of laser typically depends on the bandgap E_g of the amplifying medium, only parts of it can be reached using conventional methods such as III-V semiconductor lasers. To overcome this

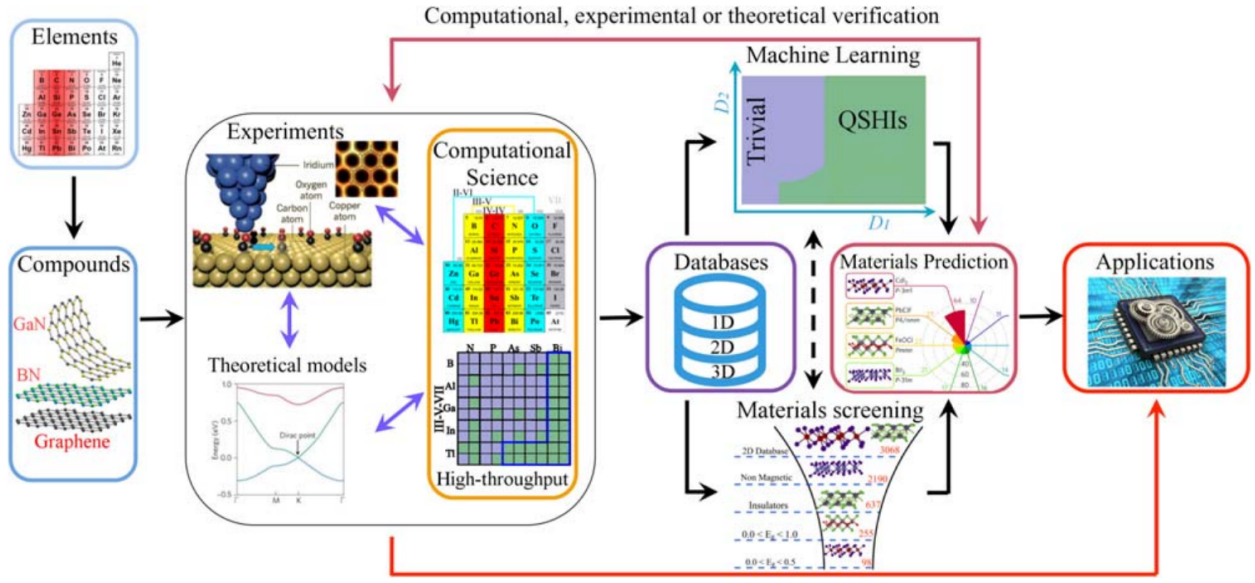


Figure 0.1: Schematic of the design process of functional materials. The possible atomic combinations form a great number of compounds, which can be studied by means of experimental, theoretical or computational approaches, especially with high-throughput calculations. Large quantities of data generated are then stored in databases, which can be used by means of materials screening or machine learning, both of which leads to promising materials candidates. Data-driven or traditional routes select materials suited for specific applications. Taken from Ref. [4].

limitation, the trick is to use particular optical materials, that is, materials that interact with EM fields in devices. These can convert the frequency of the incident radiation, ω_{in}, to a higher or lower one through specific nonlinear optical (NLO) processes such as sum frequency generation. The latter enables the functioning of the common Nd:YAG lasers by doubling ω_{in} for example [5]. Thanks to these NLO crystals, a wider range of frequencies can be explored and exploited. As of today, appropriate materials are, however, still lacking for various parts of the spectrum such as the mid- (2-25 μm) and far- (>25 μm) infrared (IR). This drastically limits the development of new technologies using these wavelengths. Furthermore, the coverage of the frequency spectrum is not the only requirement for optical materials. As a result, only a very limited set of crystals have currently been identified to fulfill most requirements. The goal of the present work is thus to initiate a theoretical investigation of new NLO materials to guide experimentalists in their work.

Firstly, in Chapter 1, a theoretical background of nonlinear optics is given in order to grasp the main concepts. A classical explanation of the NLO phenomena is developed along with a link to linear optics via Miller's rule. Technical characteristics of NLO crystal are then considered and guide the discussion towards crystals birefringence and symmetry which are necessary to perform angular phase matching. The end of Chapter 1 makes a parallel with the experimental world via the Kurtz-Perry powder technique and its effective coefficient.

Secondly, in Chapter 2, promising candidates of recent literature are investigated using first-principles computations. Their structural, electronic, and linear optic properties are analyzed with respect to their second harmonic generation response. Moreover, the variation of the latter is inspected when modifying their atomic arrangement and when considering closely related compounds.

Thirdly, in Chapter 3, common ML techniques as well as one of the materials-oriented ML models, the "material optimal descriptor network" (MODNet), are presented. One dataset with the relevant NLO properties has indeed started to be developed, allowing the use of ML to accelerate the search for new materials in the NLO field [6]. In that regard, within the present project, this already existing DB is analyzed and fed to the MODNet model in order to predict an appropriate property. It is chosen so that it reflects the potential of a material to be a NLO crystal. This materials informatics approach can also provide valuable insights into the physics and is hoped to enable further development of the theory. This can in turn help the systematic discovery or even synthetization of new NLO crystals.

Finally, conclusions are drawn with respect to the presented results and perspectives for future works are suggested at the end of the document.

Chapter 1

Nonlinear optics

This chapter aims at briefly explaining the origin of the nonlinear optical phenomena of interest as well as laying the ground for understanding the results of this work. For the sake of simplicity, the emergence of NLO phenomena is developed in Section 1.1 in a classical framework instead of a rigorous quantum theoretical one. Although not proven here, the results of both methods agree as shown in Ref. [7]. The case of second-harmonic generation is then specified in Section 1.2. The link between nonlinear and linear optics, called Miller's rule, is elaborated in Section 1.3. This leads to the enumeration in Section 1.4 of the various requirements that NLO crystals must fulfill in practice. In order to meet some of them, the property of birefringence must be exploited and is therefore explained in Section 1.5. Moreover, Section 1.6 develops the simplifications that can be applied based on the symmetry of the crystal at hand. Section 1.7 then defines a scalar equation to simplify the mathematics and take the birefringence into account. One of the most important practical criteria to be an efficient NLO crystal, i.e., phase matching, is further developed in Section 1.8 and the typical method to achieve it is presented in Section 1.9. To close this first part of the present work, Section 1.10 describes the Kurtz-Perry powder technique which is a common experimental method to evaluate the potential of NLO crystals.

1.1 Lorentz model

A well-known classical description of matter is the Lorentz model. The solid is assumed to be homogeneously made up of charged particles, each assigned to an equilibrium position. To simplify even more the discussion, let us restrict it to the case of atoms consisting of a fixed nucleus and its single non-fixed electron. The distance between these two can be changed by the application of an electric field. If the latter is fluctuating, this translates to the image of a dipole oscillating at the frequency of the field and with a phase determined by the restoring force towards the equilibrium. Similarly to an antenna, this situation leads to the radiation of an electromagnetic wave. Globally, the solid thus becomes equivalent to a phased array of antennae.

Mathematically, an oscillating dipole can be modeled as a harmonic oscillator (limited to one dimension here)

$$\frac{d^2x}{dt^2} + 2\gamma\frac{dx}{dt} + \omega_0^2 x \;=\; -\frac{e}{m}E, \tag{1.1}$$

where x is the displacement of the electron from its equilibrium position, m its mass, $-e$

its charge, E the applied electric field, and ω_0 the natural frequency of the system. The damping constant, γ, is responsible for the dissipation of energy due to the electromagnetic radiation of the accelerated electron. In a more complex model, other phenomena, such as the oscillation of the nucleus and the interaction between all particles, come into play. However, the present model assumes a single natural frequency and that all the energy driving the variation of the charge separation distance is re-radiated.

The substitution of an alternating field such as

$$E(t) \;=\; \mathcal{E}_0 \cos(\omega t - \phi) \;=\; \frac{1}{2}\mathcal{E}_0 \exp\{i\phi\} \exp\{-i\omega t\} + (\text{c.c.}) \;=\; E_0 \exp\{-i\omega t\} + (\text{c.c.}) \quad (1.2)$$

into Eq. (1.1) leads to the following solution

$$x(\omega, t) \;=\; -\frac{e}{m} E_0 \frac{\exp\{-i\omega t\}}{\omega_0^2 - 2i\gamma\omega - \omega^2}, \quad (1.3)$$

where one recognizes the natural resonance induced by the denominator. Finally the macroscopic polarization density can be obtained as

$$P \;=\; Np \;=\; -Nex \;=\; \epsilon_0 \chi(\omega) E_0 \exp\{-i\omega t\} + (\text{c.c.}), \quad (1.4)$$

where p is the microscopic electric dipole moment and ϵ_0 is the vacuum permittivity. $\chi(\omega)$ is called the linear optical susceptibility and reads

$$\chi(\omega) \;=\; \frac{Ne^2}{\epsilon_0 m} \frac{1}{\omega_0^2 - 2i\gamma\omega - \omega^2}. \quad (1.5)$$

Its contribution to P is called "resonant" in contrast to other sources of polarization contained in a "background" term which is discarded in the present work. The above result can be inserted in the expression of the total displacement field

$$D(\omega) \;=\; \epsilon_0 E(\omega) + P(\omega) \;=\; \epsilon_0 [1 + \chi(\omega)] E(\omega) \;=\; \epsilon_0 \epsilon_r(\omega) E(\omega), \quad (1.6)$$

which, in consideration of Maxwell's equations and in the case of a traveling wave, leads to n^2, the refractive index, being equal to ϵ_r, the relative permittivity [8, 9].

To go one step beyond this simple model, the anharmonicity of the oscillator can be taken into account. The equations developed up to now were indeed only valid for small deviations from the equilibrium [10]. This is accomplished by adding the term $-\xi x^2$ to the left hand side of Eq. (1.1). It effectively modifies the strength of the restoring force such that the parameter ξ now characterizes its nonlinear nature. Compared to the other terms, this new one is small enough that the solution to Eq. (1.1) can be expressed as a power series centered at zero:

$$x \;=\; \sum_l x^{(l)} \;=\; \sum_l a_l E^l, \quad (1.7)$$

where E still denotes the electric field. However, the latter must now be comprised of multiple components with different frequencies, ω_n, and phases since their mutual interaction is at the core of the NLO phenomena. At the first order, one finds the same solution as with the harmonic oscillator but extended to multiple frequencies

$$x^{(1)} = \sum_n x(\omega_n) = -\frac{e}{m}E_0 \sum_n \frac{\exp\{-i\omega_n t\}}{\omega_0^2 - 2i\gamma\omega_n - \omega_n^2} \tag{1.8}$$

with n taking both positive and negative values and such that ω_{-n} corresponds to $-\omega_n$. This result is then used to find the terms at higher orders. They are omitted here for the sake of brevity. As before, the macroscopic polarization density is expressed in terms of the displacement of the electrons. This implies that P also takes the form of a power series with the first two orders being

$$P^{(1)} = \epsilon_0 \sum_n \chi^{(1)}(\omega_n) E_0(\omega_n) \exp\{-i\omega_n t\} \tag{1.9}$$

$$P^{(2)} = \epsilon_0 \sum_{nm} \chi^{(2)}(\omega_n, \omega_m) E_0(\omega_n) E_0(\omega_m) \exp\{-i(\omega_n + \omega_m)t\}, \tag{1.10}$$

where $P^{(1)}$ corresponds to the usual linear polarization while $\chi^{(1)}$ and $\chi^{(2)}$ read

$$\chi^{(1)}(\omega_n) = \frac{Ne^2}{\epsilon_0 m} \frac{1}{\omega_0^2 - 2i\gamma\omega_n - \omega_n^2} \tag{1.11}$$

$$\begin{aligned} \chi^{(2)}(\omega_n, \omega_m) &= \chi^{(2)}(\omega_n + \omega_m, \omega_n, \omega_m) \\ &= -\frac{\epsilon_0^2 m \xi}{N^2 e^3} \chi^{(1)}(\omega_n)\chi^{(1)}(\omega_m)\chi^{(1)}(\omega_n + \omega_m). \end{aligned} \tag{1.12}$$

In order to emphasize the dependence on the three frequencies, $\omega_n + \omega_m$ is added as variable of $\chi^{(2)}$. Similar expressions can be found for higher order terms [11].

1.2 Second-Harmonic Generation

It can be seen from Eq. (1.10) that the second-order polarization is comprised of multiple components with different frequencies, $\omega_n + \omega_m$. Each one of them is responsible for the generation of electromagnetic waves. Indeed, a wave-equation in nonlinear optical media can be found starting from Maxwell's equations and by expressing the polarization in its power series form [12]

$$\nabla^2 \mathbf{E} - \frac{1}{\epsilon_0 c^2} \frac{\partial^2 \mathbf{D}^{(1)}}{\partial t^2} = \frac{1}{\epsilon_0 c^2} \frac{\partial^2 \mathbf{P}^{\mathrm{NL}}}{\partial t^2} \tag{1.13}$$

where the superscript "(1)" and "NL" indicate the linear and the nonlinear part of the expansion, respectively. It must also be noted that the 3D vectorial formalism was adopted here. Eq. (1.13) shows that the nonlinear orders of the polarization can drive an electric field.

This leads to the identification of two main phenomena due to Eq. (1.10):

1. Sum-frequency generation (SFG): $P^{(2)}(\omega_1 + \omega_2)$
2. Difference-frequency generation: $P^{(2)}(\omega_1 - \omega_2)$

Several others stem from these two, e.g., the electrooptic effect, optical rectification, and second-harmonic generation (SHG). The latter is probably the most popular consequence of nonlinear optics and is historically the starting point of this field of research. It consists in generating an output wave at twice the frequency of an incident one. It is thus a specific case of SFG. It can be modeled as a single quantum-mechanical process in which two photons are absorbed via a virtual level and a photon of twice their energy is emitted as depicted in Figure 1.1b. Throughout this work, the main quantity of interest will be the nonlinear susceptibility relative to the SHG as schematically illustrated in Figure 1.1a.

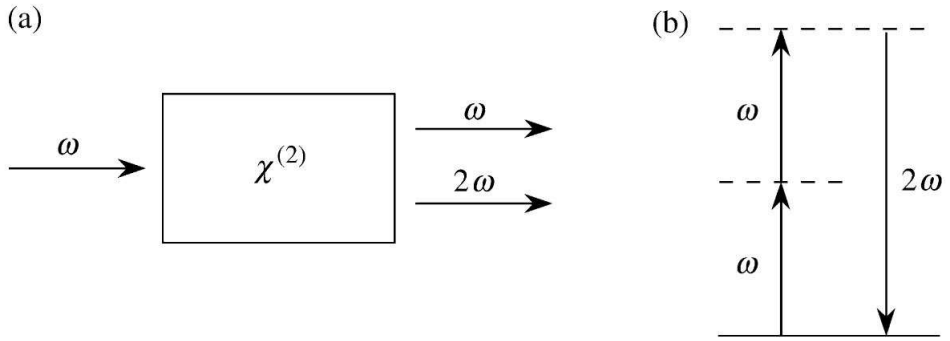


Figure 1.1: (a) Schematic of second-harmonic generation. (b) Energy-level diagram describing second-harmonic generation. The dashed and solid lines indicate virtual levels and the atomic ground state, respectively. Taken from Ref. [12].

1.3 Miller's rule

Up to now, the tensor notation has mostly been avoided by considering only one-dimensional systems. In reality, each frequency component of Eq. (1.10) is rigorously written

$$P_i^{(2)}(\omega_n + \omega_m) = \epsilon_0 \sum_{jk} \sum_{(nm)} \chi_{ijk}^{(2)}(\omega_n + \omega_m, \omega_n, \omega_m) E_j(\omega_n) E_k(\omega_m) e^{-i(\omega_n + \omega_m)t}, \quad (1.14)$$

where the subscripts i, j, and k refer to the three cartesian dimensions and the notation (nm) indicates that, in performing the summation over n and m, the sum $\omega_n + \omega_m$ is fixed. However, ω_n and ω_m are each allowed to vary [12]. The polarization and electric field have now become vectors and thus, the second-order susceptibility corresponds to a third-rank tensor. In the same way, the linear susceptibility is actually a second-rank tensor. Eq. (1.12) is also modified accordingly and even leads to a particularly useful relation, known as Miller's rule

$$\chi_{ijk}^{(2)}(\omega_1, \omega_2, \omega_3) = \left[\chi_{\mathrm{ii}}^{(1)}(\omega_1)\right] \left[\chi_{\mathrm{jj}}^{(1)}(\omega_2)\right] \left[\chi_{\mathrm{kk}}^{(1)}(\omega_3)\right] \Delta_{ijk}, \quad (1.15)$$

where $\boldsymbol{\Delta}$ is referred to as Miller's coefficient, though it is actually a third-rank tensor. It provides a direct link between linear and nonlinear optics. This relation was empirically found by Miller but it is directly visible in the result of the anharmonic oscillator [13]. At first, Δ_{ijk} was thought of as almost being constant and independent of the material. However, recent works have shown the contrary [6]. Investigating this coefficient is of considerable importance as it could hide major parameters favoring high nonlinear optical properties as well as greatly simplify the determination of the NLO response. Indeed, the latter is more complicated to obtain than its linear counterpart, both experimentally and computationally. In any case, the main idea that emanates from this relation is that efficient NLO phenomena can likely be achieved in materials exhibiting a high linear susceptibility and thus a high refractive index. As explained hereafter, this leads to contradicting requirements regarding the usefulness of a compound as NLO crystal.

1.4 NLO crystals: requirements

In order to be useful as a NLO crystal, a material must meet several requirements. Unfortunately, some of them are contradictory. This explains the scarcity of compounds that can be used in practice. The list of requirements reads as follows:

1. transparency for all frequencies of interaction
2. good stability and growability as single crystal
3. noncentrosymmetricity
4. high refractive index
5. high laser damage threshold (LDT)
6. phase matchability

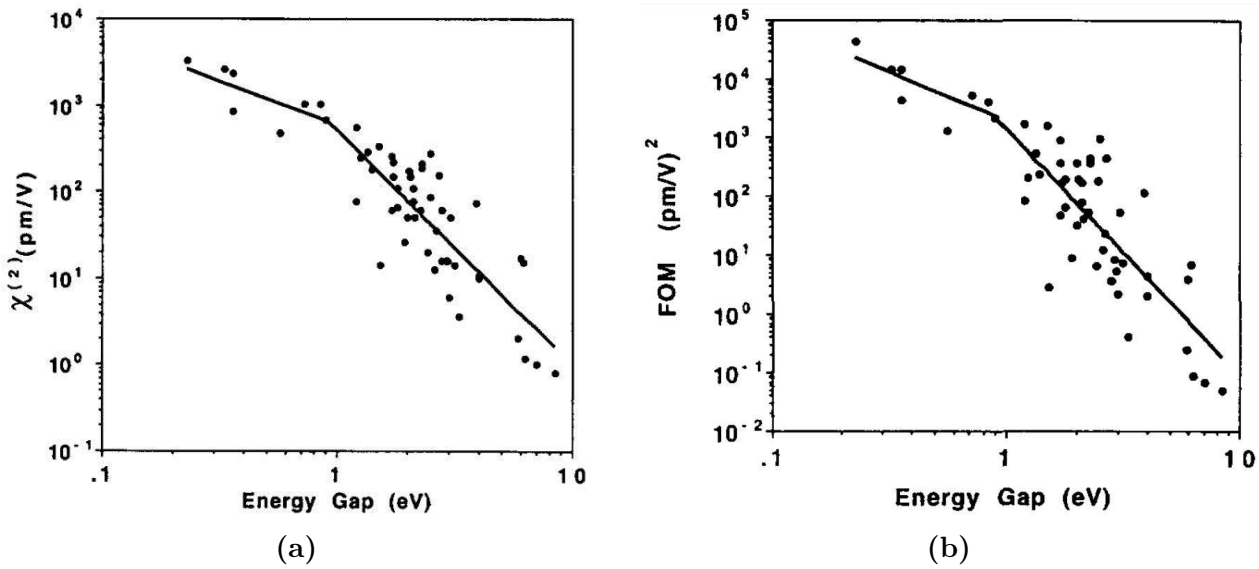
The first two conditions are practical considerations and the second one can be particularly bothersome.

The third one is a direct consequence of the third-rank nature of the nonlinear susceptibility tensor. Neumann's principle states that a tensor representing a physical property of a material must remain invariant under the application of any of the symmetry elements characteristic of the point group at hand. If one of its components changes sign under the transformation, then this component must be equal to zero by symmetry consideration. As a consequence, the 11 centrosymmetric point groups cannot present any non-null third-rank tensor. It also means that each point group has its own form of nonlinear susceptibility tensor with specific components set to 0 [14]. It can even be shown that it is homologous to the piezoelectric matrix [11].

The fourth requirement is a direct consequence of Miller's rule. Although very useful and promising as explained before, this relation also brings disadvantages. As will be developed later, the output of NLO interactions is proportional to χ^2/n^3 where χ is an effective susceptibility. According to Miller's rule, there is a direct proportionality between χ and n which, in turn, limits the output. That is why the ratio χ^2/n^3 can be considered as a figure of merit (FOM) to quantify the potential of a material to be an efficient NLO crystal.

The fifth condition is required to ensure an excellent optical quality of the crystal since NLO phenomena are present in the case of intense electric fields. This last statement is a direct consequence of the expansion in power series of the field as developed in Section 1.1. It may seem trivial but it actually introduces an important problem since it is in contradiction with the need for a high refractive index. Indeed, a large LDT is synonym of a high bandgap while a large refractive index usually requires a low one. An additional challenge is thus to find outliers in this high bandgap/low index trend [15]. Several power law models were developed and tend to show an inverse relationship between an average value of $\chi^{(2)}$ and the bandgap E_g [16] as is represented in Figure 1.2. This means that a compromise is inevitable when considering NLO crystal candidates.

The sixth requirement is also an extremely important point in practice. However, its understanding requires some prerequisites which will be developed in the next sections.



(a) (b)

Figure 1.2: Experimental data of (a) effective values of the second-order nonlinear susceptibility tensor and (b) FOM as defined in the text with respect to the bandgap. The solid lines are power law models fitted on the data. Taken from Ref. [16].

1.5 Birefringence

Taking into account anisotropies by switching to tensor notation does not only affect the susceptibility. Indeed, it also impacts the dielectric constant and thus the refractive index since these three quantities are related to one another. Indeed, the electric field displacement does not have to be aligned with the electric field anymore as shown by the following equation

$$D_i = \sum_j \epsilon_{ij} E_j. \tag{1.16}$$

This section introduces the consequences of this 3D-formalism. The corresponding demonstrations can be found in the literature such as Refs. [11, 17]. After some development related to the electric energy density, one finds that the dielectric tensor has to be symmetric, thus reducing its number of independent components to 6. It also comes that, given a set of cartesian axes x, y, and z,

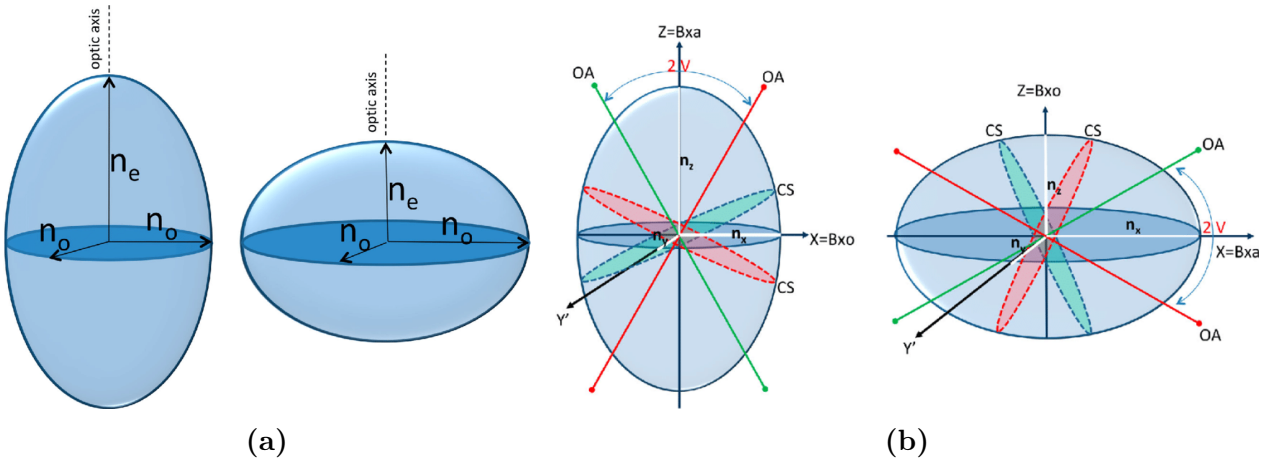
$$\epsilon_{xx}x^2 + \epsilon_{yy}y^2 + \epsilon_{zz}z^2 + 2\epsilon_{yz}yz + 2\epsilon_{xx}xz + 2\epsilon_{xy}xy = \text{constant} \tag{1.17}$$

$$\Rightarrow \epsilon_x x^2 + \epsilon_y y^2 + \epsilon_z z^2 = \text{constant}, \tag{1.18}$$

where the first equation represents an ellipsoid while the second one is its transformation to its principal axes in a particular coordinate system of the crystal. ϵ_x, ϵ_y, ϵ_z are called the principal dielectric constants and they define the principal refractive indices n_i as their respective square root. This implies a new ellipsoid, called the "optical indicatrix" or "ellipsoid of wave normals", whose semiaxes are the n_i. Anisotropic crystals can now be classified in two categories as illustrated in Figure 1.3. Uniaxial crystals are characterized by an ellipsoid of revolution as their indicatrix, i.e., two of their three principal refractive indices are equal, whereas biaxial ones possess three different semiaxes. Their denomination refers to their number of optic axes which are defined as the directions perpendicular to a circular cross section of the indicatrix, i.e., the wave-normal directions in which the direction of polarization does not influence the value of the refractive index. As can be directly understood, in the case of uniaxial crystals, the circular cross section is the plane defined by the two equal semiaxes and it is always along the highest symmetry axis. Biaxials crystals present two optic axes positioned symmetrically around the longest semiaxis. Although not mentioned, the indicatrix of isotropic crystals is a sphere. Thanks to Neumann's principles, the symmetries of each crystal system determine their optical category:

- Isotropic: cubic
- Uniaxial: trigonal, tetragonal, hexagonal
- Biaxial: orthorhombic, monoclinic,triclinic

Although not demonstrated here, an important consequence of the anisotropy is that only two orthogonal directions of linear polarization are allowed for rays propagating inside



(a) **(b)**

Figure 1.3: Representation of the optical indicatrix for positive (on the left) and negative (on the right) for (a) uniaxial and (b) biaxial crystals. In (a), the optic axis and its perpendicular circular cross-section of the ellipsoid are respectively indicated by the dotted line and the dark blue area while it is indicated by "OA" and the green or red areas denoted by "CS" in (b). Both figures were taken from Ref. [18].

the crystals. This implies that the polarization of any incident wave is decomposed into these two directions when entering the crystal. The indicatrix allows to find both of them geometrically. The case of biaxial crystals is complex and will not be developed here. The focus is thus put on the case of uniaxial crystals. The first polarization allowed, called "ordinary", is orthogonal to the plane defined by the incident wave-normal and by the optic axis. The second polarization allowed, called "extraordinary", is the direction which lies in the plane normal to the incident wave-normal and which is perpendicular to the ordinary direction. The situation is schematically represented in Figure 1.4a where each allowed polarization has its refractive index defined by the surface of the indicatrix. This effectively results in two waves inside the crystal, namely the ordinary (o-wave) and the extraordinary (e-wave) one, with their respective refractive index.

By definition, the value of the ordinary index, n_o, does not vary with the direction of propagation. On the contrary, the extraordinary index depends on this direction and it changes "elliptically" when the angle of propagation is varied as can be seen from the xz-plane projection of Figure 1.4b in the case of KH_2PO_4 (KDP), a well-known NLO crystal. Therefore, n_e usually refers to the extraordinary index when the direction of propagation is perpendicular to the optic axis (see Figure 1.3a), called the "principal value of the extraordinary index", while the effective one is given by

$$n_e(\theta) = \frac{n_o n_e}{\sqrt{n_o^2 sin^2(\theta) + n_e^2 cos^2(\theta)}}, \tag{1.19}$$

where θ is the angle between the optic axis and the wave-normal. If the angular dependence is not explicitly written, then the quantity refers to the principal value. In a 3D-space, one can construct volumes whose surfaces correspond to the ordinary and extraordinary

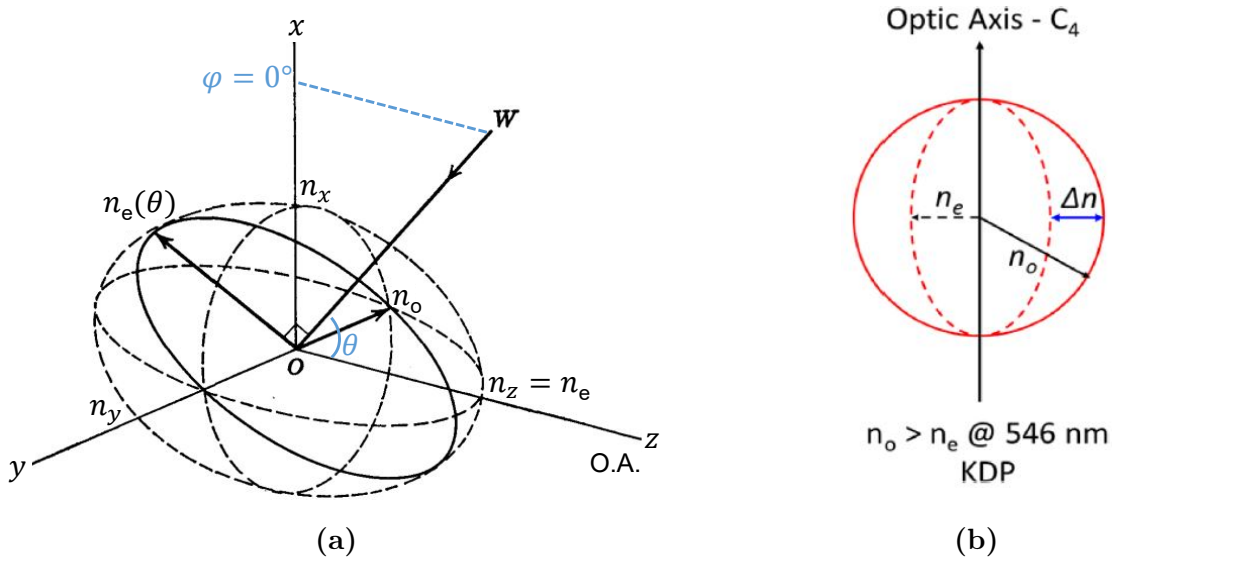


Figure 1.4: (a) Representation of the optical indicatrix for a positive uniaxial crystal and the determination of the ordinary and extraordinary polarizations for their respective refractive index, n_o and $n_\mathrm{e}(\theta)$, in the case of a wave-normal W with incidence angles (θ,$\varphi = 0°$). Adapted from Ref. [11]. (b) Two-dimensional projection of the refractive index of KDP (negative uniaxial) with $n_\mathrm{o} > n_\mathrm{e}$. The birefringence Δn is also indicated. Taken from Ref. [19].

index for every angle possible. The absolute difference between n_o and n_e is known as the "birefringence." As shown in Figure 1.3, a further classification can be realized: an uniaxial or biaxial crystal is called "positive" ("negative") if $n_\mathrm{e} > n_\mathrm{o}$ ($n_\mathrm{o} > n_\mathrm{e}$) or $n_z - n_y > n_y - n_x$ ($n_z - n_y < n_y - n_x$), respectively [19, 20]. As stressed upon in Figure 1.4b, it is important to specify the frequency at which the refractive indices are given since they vary with it. Far from any absorption band, their dispersion is called "normal", i.e., they increase with decreasing wavelength [8].

1.6 Symmetry and contraction

It is now necessary to introduce the usual conventions regarding the form of the nonlinear susceptibility tensor. In practice, a modified tensor is used instead of the usual one. This new quantity is defined as follows

$$d_{ijk} = \frac{1}{2}\chi^{(2)}_{ijk} \tag{1.20}$$

The first step in simplifying the tensor is possible thanks to the first requirement mentioned in the previous section which implies that the linear susceptibilities are real. Thanks to Eq. (1.12), this effectively reduces the number of independent components to 81 since $d_{ijk}(\omega_n + \omega_m, \omega_n, \omega_m)$ now equals $d_{ijk}(-\omega_n - \omega_m, -\omega_n, -\omega_m)$. Moreover, this same equation thus implies that the indices can be freely permuted if the frequencies are permuted

simultaneously [21]:

$$\begin{aligned} d_{ijk}(\omega_3, \omega_1, \omega_2) &= d_{jki}(\omega_1, -\omega_2, \omega_3) = d_{kij}(\omega_2, \omega_3, -\omega_1) \\ &= d_{ikj}(\omega_3, \omega_2, \omega_1) = d_{kji}(\omega_2, -\omega_1, \omega_3) \\ &= d_{jik}(\omega_1, \omega_3, -\omega_2). \end{aligned} \tag{1.21}$$

The number of independent components now drops to 27 following this assumption. A further reduction can be obtained as follows. Intuitively, the sequence of the application of two fields should not influence the result of their interaction. This translates to setting d_{ijk} equal to d_{ikj} and thus to reducing the number of independent components to 18 [11]. This allows to adopt the Voigt notation which is illustrated in Figure 1.5.

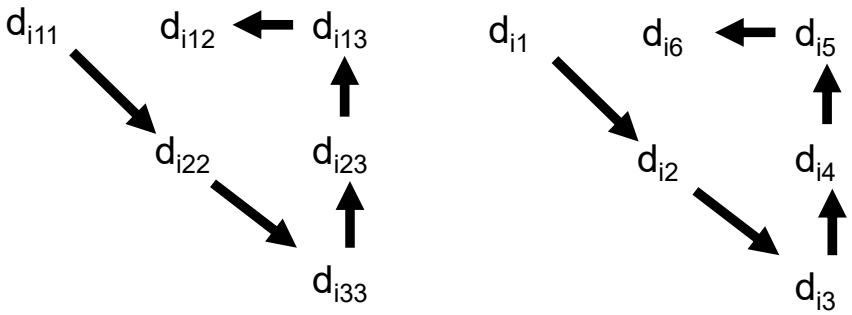


Figure 1.5: Graphical illustration of switching to Voigt notation for the last two indices. Adapted from Ref. [11].

Going even one step further, one can disregard the frequency-dependence in Eq. (1.21), i.e., to assume a dispersionless medium. This assumption is in fact called Kleinman's symmetry condition. It is applicable when all interacting frequencies are much smaller than the resonance [22]. This is a generally valid simplification in the case of SHG and this reduces the number of independent components down to 10 such that the final reduced tensor has the form [12]

$$d_{il} = \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{16} & d_{22} & d_{23} & d_{24} & d_{14} & d_{12} \\ d_{15} & d_{24} & d_{33} & d_{23} & d_{13} & d_{14} \end{bmatrix} \tag{1.22}$$

Moreover, Neumann's principle should not be left aside and can bring further simplifications depending on the type of crystal.

1.7 Effective value of the nonlinear susceptibility

Even after having simplified the main tensor, the sum in Eq. (1.10) is still mandatory to obtain the polarization of the generated wave. However, the decomposition into ordinary and extraordinary rays must also be taken into account in this equation. Moreover, since scalar relations are more convenient to manipulate in general, Eq. (1.10) should be modified to explicitly contain the amplitude of the fields. The following equation is obtained as a

result in the case of SHG, i.e., of fields at frequency ω and resulting polarization at 2ω.

$$P^{\mathrm{o}} \;=\; |\mathbf{P}^{\mathrm{o}}| \;=\; a_i P_i^{\mathrm{o}} \;=\; 2a_i d_{ijk} E_j^{\mathrm{e}} E_k^{\mathrm{e}} \;=\; 2a_i d_{ijk} b_j b_k |\mathbf{E}^{\mathrm{e}}|^2 \;=\; d_{\mathrm{eff}} |\mathbf{E}^{\mathrm{e}}|^2 \tag{1.23}$$

with

$$E_j^{\mathrm{e}} \;=\; |\mathbf{E}^{\mathrm{e}}| \begin{vmatrix} -\cos\varphi\cos\theta \\ -\sin\varphi\cos\theta \\ \sin\theta \end{vmatrix} \;=\; |\mathbf{E}^{\mathrm{e}}|\, b_j \tag{1.24}$$

$$E_j^{\mathrm{o}} \;=\; |\mathbf{E}^{\mathrm{o}}| \begin{vmatrix} \sin\varphi \\ -\cos\varphi \\ 0 \end{vmatrix} \;=\; |\mathbf{E}^{\mathrm{o}}|\, a_j, \tag{1.25}$$

where Einstein summation convention was used, θ is the angle between the propagation direction and the optic axis which coincides with the z-axis while φ is the angle between the xz-plane and the projection of the propagation direction onto the xy-plane as illustrated in Figure 1.4a. It is important to note that the collinear case is always considered in this work. The superscripts "o" and "e" refer respectively to the ordinary and extraordinary components of the fields and polarization. The quantity $d_{\mathrm{eff}}(\theta, \varphi)$ is thus an effective average of d_{ijk} for a specific geometry of incident wave, crystal system, and interaction of e- and o-waves [23, 24].

1.8 Phase matching

Thanks to this scalar relationship, it is easier to work with Maxwell's equations to derive the coupled-wave equations. From the latter, one can obtain the intensity of the generated field at ω_3 from incident ones at ω_1 and ω_2:

$$I_3 \;=\; \frac{8\, d_{\mathrm{eff}}^2\, \omega_3^2\, I_1\, I_2}{n_1\, n_2\, n_3\, \epsilon_0\, c^2} L^2\, \mathrm{sinc}^2\left(\frac{\Delta k L}{2}\right), \tag{1.26}$$

where the numerical indices refer to the respective frequencies, L to the length of the crystal and Δk to the momentum mismatch, i.e., $k_1 + k_2 - k_3$. The interested reader will find the corresponding demonstration in Ref. [12]. If it were not for the last factor, called the "phase mismatch factor", this function would depend on the square of the crystal length. However, the squared sinc function strongly modulates the intensity and it is the primary limiting factor of the conversion efficiency as shown in Figure 1.6. The phase mismatch factor drops to zero when the crystal length is equal to $2/\Delta k$, i.e., the "coherence length" L_{coh}. This denomination was chosen because it corresponds to the limit where the output wave is not yet out of phase with its driving field. Beyond this distance, it can be shown that the power is coupled back into the incident fields. In order to intuitively grasp the situation, one can consider an array of antennae as introduced in Section 1.1. Figure 1.7 depicts it in the case of a spacing between dipoles inducing a Δk of $\pi/4$. The periodic behavior with length is clearly

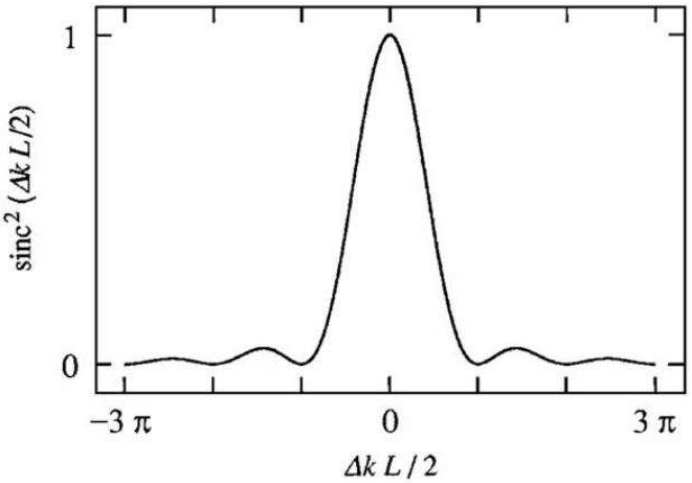


Figure 1.6: Effects of momentum mismatch on the efficiency of sum-frequency generation. Taken from Ref. [12].

visible and relates directly to the oscillation of the sinc function. Each of the waves generated by the dipoles must thus add constructively (in the forward direction) to the resulting field from all the preceding dipoles if the best efficiency is to be achieved. This limitation implies

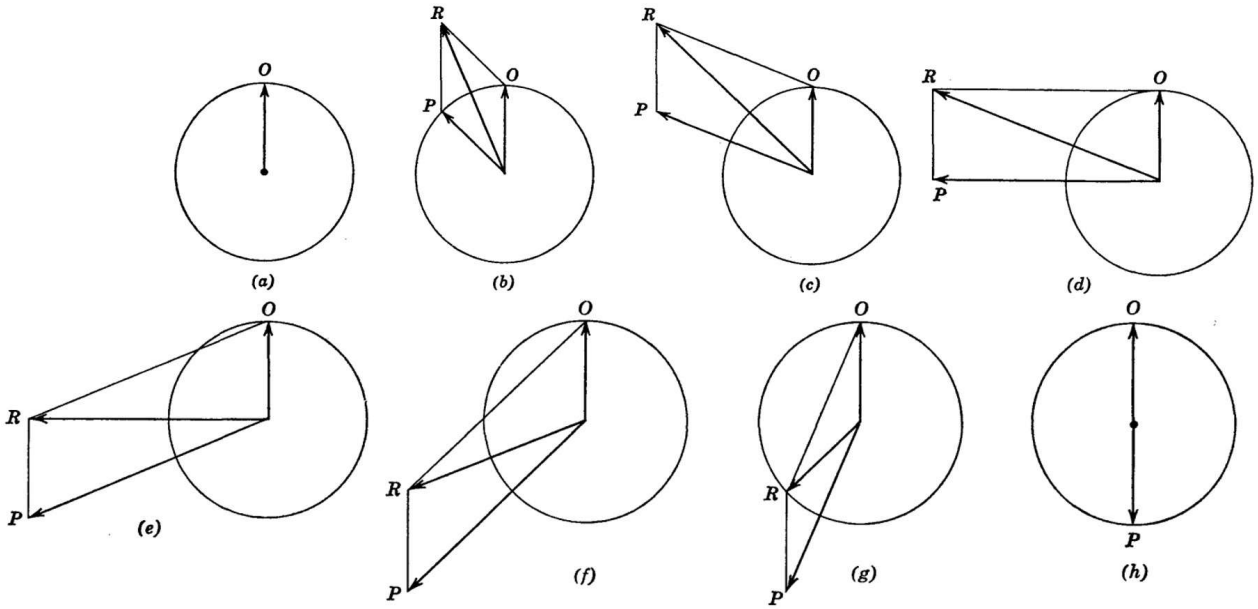


Figure 1.7: Intuitive illustration of a phase mismatch between the driving and generated waves on the amplitude and phases of the resulting waves. O, P, and R respectively denote the waves generated by the individual dipole, the resultant of the waves from the previous dipoles, and the resultant of O and P. The spacing between dipoles corresponds to a phase mismatch of $\pi/4$. Adapted from Ref. [11].

that the best conversion efficiency is attained when the waves are phase matched ($\Delta k = 0$) so that the phase mismatch factor does not matter anymore. It must be kept in mind that this condition corresponds to the case of collinear waves as mentioned before, otherwise it would write vectorially as $\Delta\mathbf{k} = 0$. This more general expression should be preferred but, in

practice, the waves are assumed to be collinear because it is experimentally more convenient. This perfect phase matching criterion can be expressed in terms of the refractive indices [23]

$$k_3 = k_2+k_1 \Leftrightarrow \frac{\omega_3 n_3}{c} = \frac{\omega_1 n_1}{c}+\frac{\omega_2 n_2}{c} \Leftrightarrow n_3 = \frac{\omega_1 n_1+\omega_2 n_2}{\omega_3} \overset{\text{SHG}}{\Rightarrow} n(2\omega) = n(\omega) \quad (1.27)$$

where the last equality was obtained in the specific case of SHG. Unfortunately, without any tricks, this equation does not possess a solution because of the normal dispersion of the refractive index, i.e., the increase of n with the frequency. Theoretically speaking, the anomalous dispersion (decrease of n with the frequency) near an absorption band could be exploited in order to achieve Eq. (1.27). This strategy would however be in contradiction with the first requirement of Section 1.4. Alternatives exist thanks to specific phenomena which are able to modify the refractive index. The three possibilities to achieve phase matching are the angular, temperature-tuned [25], and quasi phase matching [26]. Only the first option is developed hereafter since it is the most common technique.

1.9 Angular phase matching

In order to attain true phase matching as expressed in Eq. (1.27), it has been discovered that using the birefringence of the crystal is key [27, 28]. The two components (ordinary and extraordinary) of a beam do not "see" the same refractive index as explained in Section 1.5. This means that one could choose a configuration such that, for example, $n_\mathrm{e}(\theta_m,\omega)$ be equal to $n_\mathrm{o}(2\omega)$ [29]

$$n_\mathrm{o}(2\omega) = n_\mathrm{e}(\theta_m,\omega) = \frac{n_\mathrm{o}(\omega)n_\mathrm{e}(\omega)}{\sqrt{n_\mathrm{o}^2(\omega)\sin^2(\theta)+n_\mathrm{e}^2(\omega)\cos^2(\theta)}} \quad (1.28)$$

$$\Rightarrow \sin^2(\theta_m) = \frac{n_\mathrm{e}^2(\omega)}{n_\mathrm{o}^2(2\omega)}\frac{n_\mathrm{o}^2(\omega)-n_\mathrm{o}^2(2\omega)}{n_\mathrm{o}^2(\omega)-n_\mathrm{e}^2(\omega)} \quad (1.29)$$

This is equivalent to stating that the extraordinary component of both incident fields at the fundamental frequency interact to generate an ordinary wave at the second-harmonic. This example works only in the case of a positive uniaxial crystal. Indeed, the feasibility of such method relies on the combination of normal dispersion and birefringence as shown by Figure 1.8a. The necessary condition is that the index at the second-harmonic be larger than the one at the fundamental so that one can find θ_m, i.e., the angle for which the effective extraordinary index respects the phase matching condition. In other words, the way to achieve phase matching is to find the loci of points in space where the refractive index surfaces of second-harmonic and fundamental cross one another. The 2D-projection in Figure 1.8b illustrates this statement as it defines the phase matching angle at the intersection, which corresponds to a cone in 3D, of the index surfaces. Figure 1.8a also implies that, depending on the type of crystal and thus of birefringence, different combinations of e- and o-waves are possible. Instead of the e+e→o combination that was chosen above, a

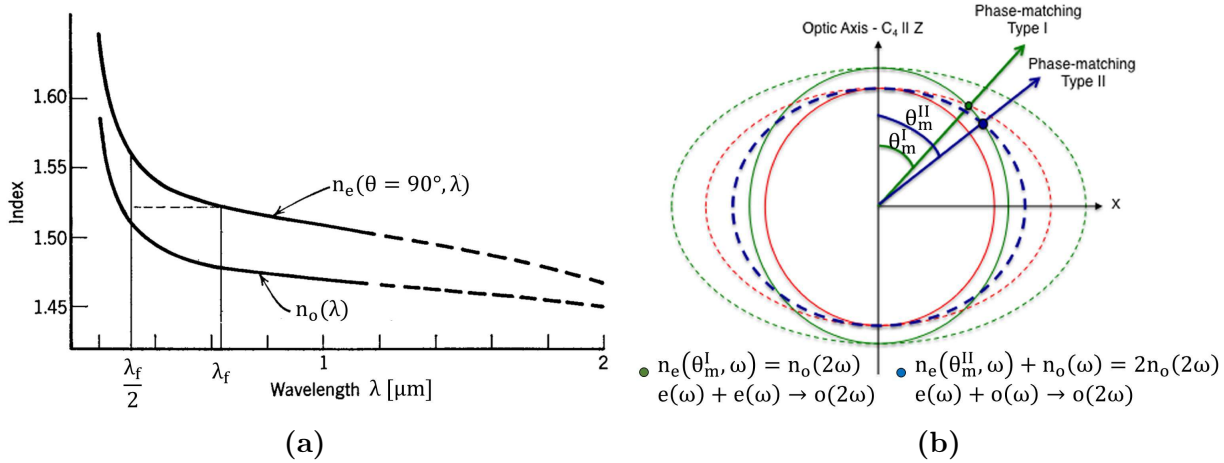


Figure 1.8: (a) Illustration of the feasibility of angular phase matching thanks to normal dispersion and birefringence in the case of a positive uniaxial crystal. Adapted from Ref. [11]. (b) 2D-projection of the refractive index surfaces. The fundamental and second-harmonic correspond to the red and green lines while the dotted and solid ones refer to $n_e(\theta)$ and n_o, respectively. The dotted blue ellipse is the mean of the surfaces at the fundamental. Adapted from Ref. [19].

negative uniaxial crystal allows the o+o→e interaction as one can understand from Figure 1.8.

Up to now, only type I phase matching, i.e., when incident waves are similarly polarized, has been presented. Type II phase matching is also a possibility and corresponds to the interaction between perpendicularly polarized driving fields: e+o→o or e+o→e. In this case, Eq. (1.27) results in the following condition

$$n_o(2\omega) = \frac{n_e(\omega) + n_o(\omega)}{2} \tag{1.30}$$

This new condition is graphically represented in Figure 1.8b as the intersection of the green solid line and the dotted blue line which respectively correspond to the left and right hand side of the above equation.

In conclusion, uniaxial crystals possess four different possibilities to be phase matched:

- Positive type I: e+e→o
- Negative type I: o+o→e
- Positive type II: e+o→o
- Negative type II: e+o→e

The choice of combination must be reflected in Eq. (1.23) and thus in d_{eff}. All possibilities for the latter have been tabulated and can be found in the literature [23, 24]. This also implies that, for a given $d_{\text{eff}}(\theta, \varphi)$, only specific couples of angles can lead to phase matching and practical SHG. As mentioned previously, the case of biaxial crystals was not developed due to its complexity in comparison with uniaxial crystals. However, the same reasoning can be applied for both types of crystals. The interested reader can find more information

on angular phase matching in both uni- and biaxial crystals in Refs. [11, 12, 19].

Although angular phase matching seems like a perfect method, it comes with some limitations. The first one is due to the so-called "walkoff effect." It can be shown that, for extraordinary rays, the propagation direction is not aligned with the Poynting vector when θ differs from 0° or 90°. This implies that the spatial overlap of ordinary and extraordinary waves is limited by their quick divergence from one another. This leads to a decrease of the conversion efficiency. Type II phase matching particularly suffers from such phenomenon. The second drawback is a consequence of the nature of light. A high output intensity requires a large one at the fundamental frequency as can be seen from Eq. (1.26). This is achieved by focusing the incident beam which in turn causes its fast divergence as predicted by Gaussian optics [17]. However, it can be demonstrated that the variation of phase mismatch is linearly dependent with the variation of θ. The level of convergence thus restricts the crystal length over which efficient conversion occurs. For these two reasons, it can be more advantageous to switch to temperature-tuned or quasi phase matching, which are both out of the scope of the present work [11, 30].

1.10 Materials screening

As mentioned in Section 1.4, the numerous and sometimes contradicting conditions to be considered as a potential NLO crystal have prevented the discovery of many new ones. For a long time, characterization techniques required single crystals of good quality which was the major experimental issue. The efforts were thus focused on the crystal-growing task which is time-consuming and expensive. Even after having grown a correct sample, it was not certain beforehand that phase matching was possible. A fast experimental screening of the materials was therefore not conceivable. Fortunately, the Kurtz-Perry (KP) powder technique was developed, thereby eliminating the need for large single crystals and simplifying the process of screening materials before growing them properly to test them further. It allows experimentalists to determine both the magnitude of the NLO response and whether or not phase matching is possible.

The principle is as follows. Incident waves are sent onto a powder sample in order to measure its SHG response integrated over the whole solid angle as a function of the particles size inside it. Two very different behaviors are then observed depending on whether the material allows phase matching or not as illustrated in Figure 1.9 [11, 31]. In the first part of the graph, both responses increase with increasing average particles size, $\hat{r}$. Indeed, since the particles are smaller than the coherence length (averaged on every orientation, $\hat{l}_c$), the phase error is not large enough to cause any decrease in the conversion efficiency. However, once $\hat{r} > \hat{l}_c$, the phase-mismatch effects become non-negligible for most orientations and non-phase-matchable materials see their response decrease because the increase in crystal

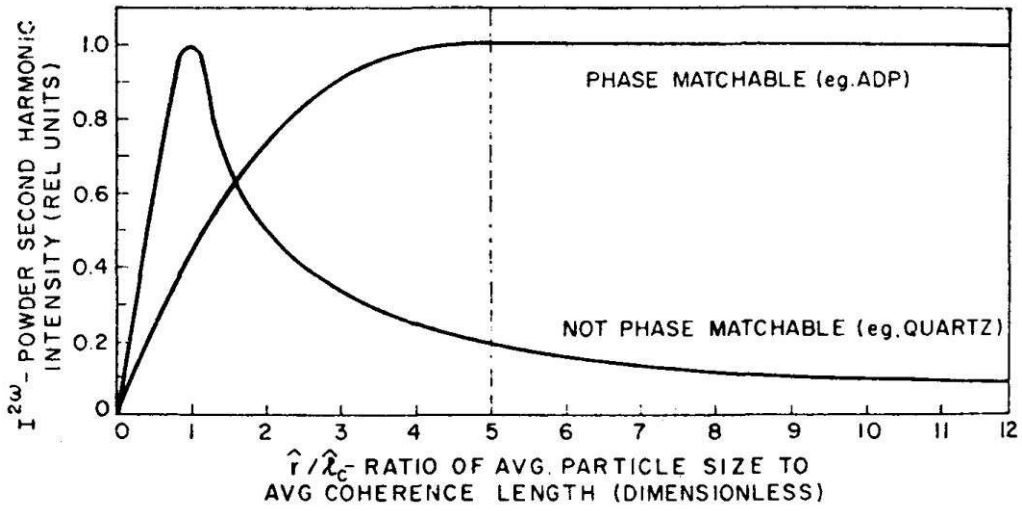


Figure 1.9: Typical graph of the SHG response as a function of the average particles size obtained with the Kurtz-Perry powder method. Taken from Ref. [11].

length does not compensate for the decrease of grain (particle) number anymore. On the contrary, phase-matchable materials reach saturation in their response because the increase of crystal length is enough to counterbalance the decreasing number of particles. In other words, one can determine if a compound is phase-matchable or not by looking at its behavior with this powder technique. However, in practice, the details of this interpretation should be nuanced following the critics expressed in Ref. [32].

Kurtz and Perry also derived expressions for the SHG intensity emitted with their powder technique. For the sake of clarity, only the relevant dependencies are present in the following equations which are only valid for grain sizes far larger than the average coherence length [33]

$$\text{Non-phase-matchable: } I(2\omega) \propto d^2_{\text{eff,KP}} \frac{\hat{l}_c}{\hat{r}} \tag{1.31}$$

$$\text{Phase-matchable: } I(2\omega) \propto d^2_{\text{eff}} \sin(\theta_m) \tag{1.32}$$

where d_{eff} is evaluated at the phase-matching angles and the only new quantity is $d_{\text{eff,KP}}$ which is the angular average of d_{ijk}, i.e., its rotational average over all random orientations in a fixed lab frame. It originates from the random orientations of the particles in the powder sample. Recent methods exist to efficiently compute this value in the case of high-rank tensor [34]. However, it is not necessary for a third-rank tensor and a simple analytical formula was found [31, 35]

$$\begin{aligned} d^2_{\text{eff,KP}} &= \frac{19}{105}\sum_i (d_{iii})^2 + \frac{13}{105}\sum_{i\neq j} d_{iii}d_{ijj} \\ &+ \frac{44}{105}\sum_{i\neq j} (d_{iij})^2 + \frac{13}{105}\sum_{ijk,cyclic} d_{iij}d_{jkk} + \frac{5}{7}(d_{ijk})^2 , \end{aligned} \tag{1.33}$$

where each component of d_{ijk} refers to the second-harmonic. This new effective nonlinear coefficient thus allows a rapid determination of the magnitude of the NLO response in the case of non-phase-matchable materials. The difference with phase-matchable crystals is due to their response being dominated by the phase matched directions, thus eliminating the need for an angular average.

If one discards at first the problem of phase matching, $d_{\text{eff,KP}}$ can be considered as an effective measure of the magnitude of the SHG response in both cases. If this approximation is not made, then the high-throughput screening of large materials database becomes more complex since a distinction between non- and phase-matchable materials must be realized as a first step. Moreover, the agreement is far from perfect for both $d_{\text{eff}}(\theta, \varphi)$ and (θ_m, φ_m) obtained via experimental and *ab initio* quantities rescaled on the density functional perturbation theory (DFPT) static values. This issue is illustrated in Figure 1.10 for the case of $KTiOPO_4$ (KTP) and is still a work in progress. The main message is thus that the rotational average of the nonlinear tensor can serve as a first approximation of the SHG magnitude that would be obtained via the KP powder technique.

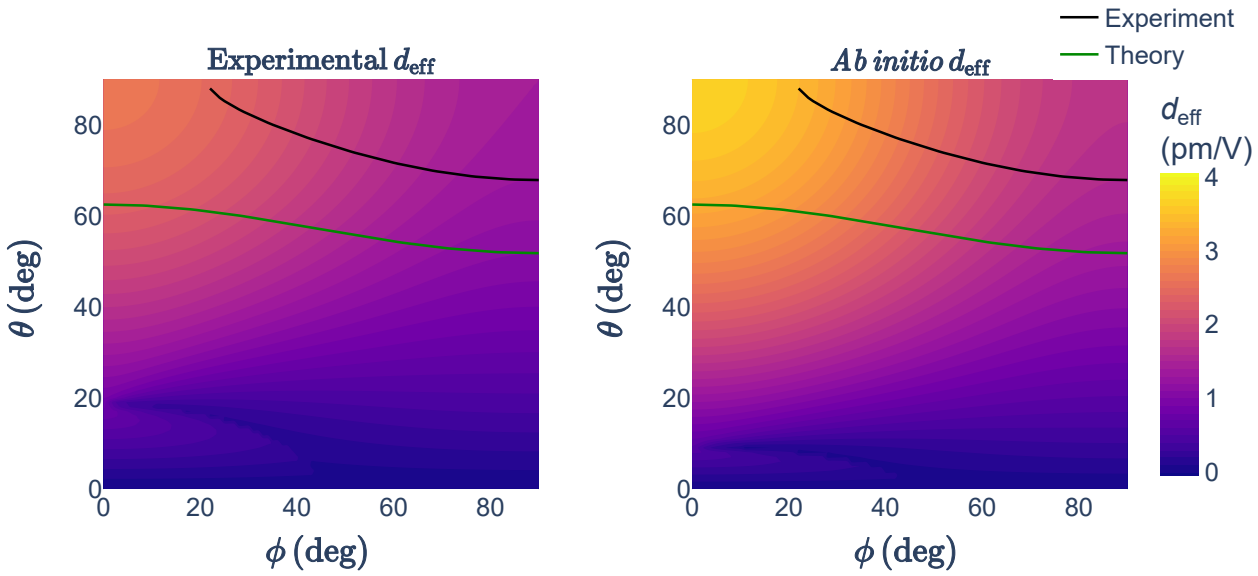


Figure 1.10: Type II phase matching for SHG in KTP at the fundamental wavelength $\lambda =$ 1064 nm. The solid curves correspond to the phase-matching combinations of (θ, φ) derived in Ref. [36]. The black and green curves refer to the use of experimental [36] and *ab initio* (Optic utility of ABINIT) refractive indices (which were rescaled on static DFPT results), respectively. Thanks to the analytical formula of Ref. [24] as introduced in Section 1.7, the colorbar indicates the value of d_{eff} for each (θ, φ) computed from an experimental SHG d_{ijk} [37] (left) and an *ab initio* one via static DFPT (right). Unpublished results from Naccarato *et al.*

1.11 Conclusion

This first chapter introduced useful concepts of nonlinear optics. The second-order polarization was derived within a simple classical framework in Section 1.1 and was shown to cause SHG among others in Section 1.2. A link between linear and nonlinear optics, called Miller's rule, was established in Section 1.3 and promises to greatly help the search for new materials. Indeed, the requirements to be a NLO crystal were found to be contradictory since both a high E_g and refractive index are needed as explained in Section 1.4. To fulfill one of these conditions, the property of birefringence, i.e., the anisotropy of refractive indices, was developed in Section 1.5. Moreover, simplifications of the SHG tensor such as the Kleinman's condition were adopted based on symmetry and physical considerations in Section 1.6. This also led in Section 1.7 to a single scalar equation when taking into account the decomposition of the propagating ray into an ordinary and an extraordinary component. The concept of phase matching was shown in Section 1.8 to play a major role in the conversion efficiency. Section 1.9 presented the method of angular phase matching which combines the birefringence with the normal dispersion of the refractive indices. Finally, the KP powder technique was presented in Section 1.10. It is a common experimental method to assess the potential of a compound to perform phase matching. It also provides an effective SHG coefficient that can characterize the strength of the SHG response of any material.

Chapter 2

Investigation of promising compounds

As mentioned, the availability of appropriate NLO crystals is limited due to the numerous requirements introduced in Section 1.4. The mid-IR region particularly suffers from this lack of materials. Indeed, the current ones all deal with various drawbacks such as small SHG coefficients, low LDT, or two-photon absorption. This has become an important issue considering the vast spectrum of applications that this range has to offer, for instance, in the medical field. To that end, the family of II-IV-V_2 compounds is being investigated in order to determine their potential as NLO crystals [38]. The present work thus focuses on the Mg-IV-V_2 (IV = Si,Ge,Sn ; V = P, As) materials by studying their SHG response from first-principles calculations. The analysis of the results is greatly inspired by the one done in Ref. [38]. Section 2.1 details the approach employed throughout. After discussing the structural, electronic and linear optic results in Section 2.2.1, Section 2.2.2 and Section 2.2.3, emphasis is put on retrieving trends between their NLO response and their structure or elements in Section 2.2.4. To go further, the consequences of structural modifications as well as the behavior of their binary analogs are investigated in Section 2.2.5 and Section 2.2.6, respectively.

2.1 Methods

The calculations were realized with ABINIT unless stated otherwise [39, 40]. The procedure that was adopted is the following.

1. A relaxation step modifies the structure that was downloaded beforehand from the MP [41, 42, 43, 44, 45, 46]. The maximal absolute force tolerance on each atom was set to 5×10^{-5} Ha/Bohr (2.5×10^{-3} eV/Å) and a 10×10×10 k-point grid was used. This step was repeated twice: with PBEsol and with LDA exchange-correlation functionals. The first iteration gave the structural input of the next step with VASP while the result of the second iteration was used for all the other steps [47].

2. Using the projector augmented wave (PAW) method with Perdew-Burke-Ernzerhof (PBE) functionals as implemented in the software VASP, the ground-state (GS) density was first computed followed by the bandgap with Heyd–Scuseria–Ernzerhof (HSE06) hybrid functionals [47]. In the case of the GS, the k-point grid and the tolerance in the

difference of total energy are a 8×8×8 and 10^{-10} eV respectively while it is a 4×4×4 and 10^{-6} eV for the bandgap. In both cases, the energy cutoff is 600 eV.

3. The bandstructure, Projected Density of States (PDOS) and Joint Density of State (JDOS) were then computed with a 10×10×10, 16×16×16 and 16×16×16 k-point grid, respectively. The latter, as well as the parameters of the integration, is the result of a convergence study for the case of $MgSiAs_2$ (see Appendix B). The tolerance on the wavefunction squared residual was set to 10^{-12} Ha^2.

4. Optical properties were then computed in the framework of the independent-particle approximation (IPA) as implemented in the Optic utility of ABINIT [48, 49]. This step was realized with a scissor shift to match the HSE bandgap of step 2 and a 16×16×16 kpoint grid.

5. Finally, the static nonlinear properties were computed with DFPT, both with and without a scissor shift to match the HSE bandgap of step 2 [50, 51, 52]. The k-point grid used was a 16×16×16 and the tolerance on the difference of total energy was 10^{-22} Ha. The k-point sampling was decided based upon a convergence study where the threshold was set to a change of 0.02% in the value of d_{36} for the case of $MgSiP_2$.

Except for the second step, all calculations were realized using Scalar Relativistic (SC) Optimized Norm-Conserving Vanderbilt Pseudopotentials (ONCVPSP) with Local Density Approximation (LDA) as exchange-correlation functionals [53]. They were downloaded directly from the PseudoDojo [54]. The energy cutoff of 42 Ha that was used throughout is also taken from the latter (normal cutoff energy hint). The 10×10×10 k-point grid as well as the grid used with VASP were chosen based on previous experiences of colleagues. However, the former was tested afterwards in a convergence study of the lattice constants. All the parameters and convergence studies used throughout this work can be found in Appendix B.

2.2 Results and discussion

2.2.1 Crystal structures

The six compounds of interest are II–IV–V_2 ternary compounds in chalcopyrite structure (space group $I\bar{4}2d$). They are ternary analogs of III-V zincblende semiconductors and have thus closely related crystalline structure. The link between the zincblende and the chalcopyrite structure is illustrated in Figure 2.1. One can start from the diamond cubic structure such as displayed by silicon. The cubic zincblende structure presents the same atomic arrangement as the diamond cubic one but with two different types of atoms forming two different sublattices. This is shown by AlP where each specie is surrounded by the other. The final step is to replace the element of the third column by ones of the second and fourth such as Mg and Si in Figure 2.1. This actually induces a distortion in the anion sublattice.

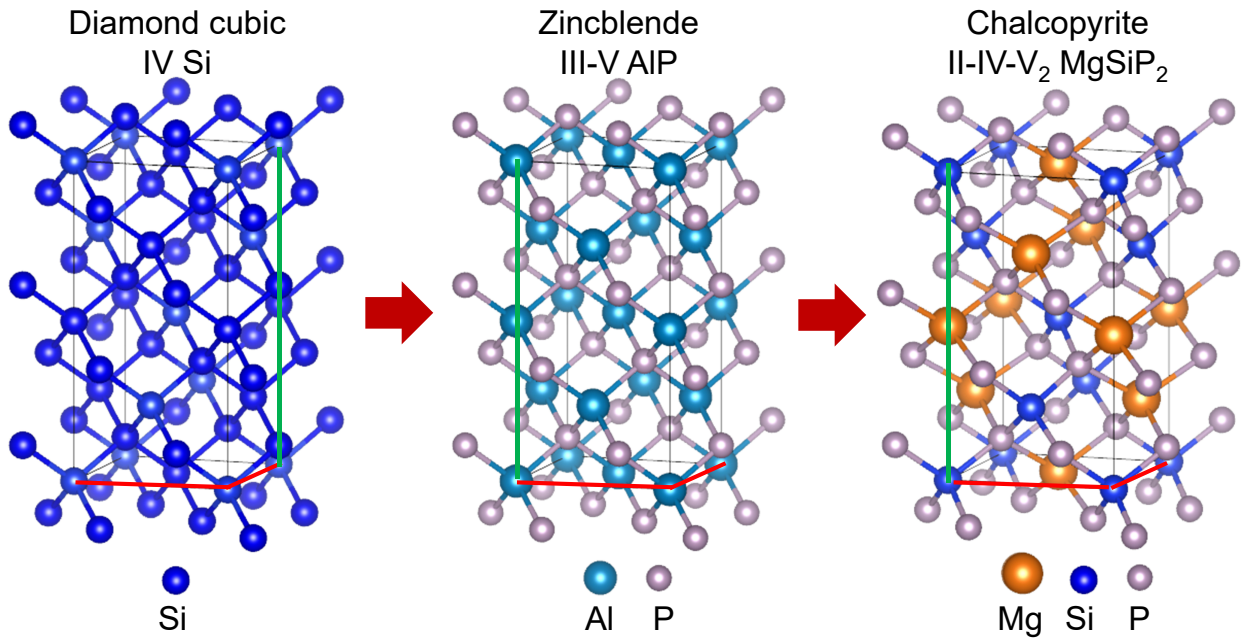


Figure 2.1: Illustration of the evolution to the chalcopyrite structure from the diamond cubic one. The red and green lines are respectively the lattice parameters a and c. Both the diamond cubic and zincblende are doubled along their c axis in a supercell 1×1×2 configuration such that $c/a = 2$. This eases the visualization of their link with the conventional unit cell of the chalcopyrite structure for which $c/a \neq 2$.

The lattice parameter a is thus not the only quantity to characterize the structure anymore. The lowering of the symmetry is indeed reflected in two new parameters: the ratio c/a, now different than 2, and the distortion parameter u, which has an ideal value of 1/4 [55, 56]

$$u = \frac{1}{4} + \frac{d_{\mathrm{II-V}}^2 - d_{\mathrm{IV-V}}^2}{a^2}, \tag{2.1}$$

where d is the distance between two different atomic groups indicated by the subscripts. This parameter is actually involved in the spatial arrangement of the group V atom. Indeed, from the Bilbao Crystallographic Server, the symmetry of the group $I\bar{4}2d$ allows the following atomic positions (relative coordinates) in the conventional unit cell [57, 58, 59]:

- Group II atom (resp., IV): (0,0,0) $(\frac{1}{2},0,\frac{3}{4})$
- Group IV atom (resp., II): $(0,0,\frac{1}{2})$ $(\frac{1}{2},0,\frac{1}{4})$
- Group V atom: $(x,\frac{1}{4},\frac{1}{8})$ $(-x,\frac{3}{4},\frac{1}{8})$ $(\frac{1}{4},-x,\frac{7}{8})$ $(\frac{3}{4},x,\frac{7}{8})$

where x can vary depending on the species of the compound. It takes the value 1/4 in the case of a zincblende 1×1×2 supercell as represented in Figure 2.1. The above positions are only half of the total ones such that the others are obtained by a shift of (1/2,1/2,1/2). It comes that, if a group II atom occupies the (0,0,0) spot, its tetrahedron shares a group V

atom at (x,1/4,1/8) with the tetrahedral unit of the group IV atom placed at (1/2,0,1/4). This implies the following equation:

$$d_{\mathrm{II-V}}^2 - d_{\mathrm{IV-V}}^2 = \left(a^2x^2 + \frac{a^2}{16} + \frac{c^2}{64}\right) - \left(\frac{a^2}{4} - a^2x + a^2x^2 + \frac{a^2}{16} + \frac{c^2}{64}\right) \tag{2.2}$$

$$= a^2\left(x - \frac{1}{4}\right). \tag{2.3}$$

By comparison with Eq. (2.1), it comes that the distortion parameter u corresponds to x and $u - 1/4$ thus indicates the deviation from the zincblende structure. Moreover, the volume of the tetrahedra formed by the four anions around the cations now depends on the specie of its central atom. Table 2.1 lists the results of the present work as well as some other values taken for comparison when considering the structure.

Compound	a (Å)	c (Å)	c/a	$d_{\mathrm{Mg-V}}$ (Å)	$d_{\mathrm{IV-V}}$ (Å)	u	V_{Mg}^T (Å^3)	V_{IV}^T (Å^3)
$MgSiP_2$	5.678	10.037	1.768	2.517	2.233	0.292	7.963	5.709
[38]	5.744	10.257	1.786	2.553	2.265	0.292		
[56]	5.733	10.242	1.787	2.549	2.261	0.292		
[60]	5.720	10.120	1.769					
$MgGeP_2$	5.714	10.481	1.834	2.511	2.312	0.279	8.017	6.340
[38]	5.787	10.750	1.858	2.549	2.354	0.278		
[56]	5.787	10.740	1.856	2.546	2.355	0.278		
$MgSnP_2$	5.836	11.376	1.949	2.513	2.499	0.252	8.143	8.002
[38]	5.933	11.572	1.950	2.554	2.542	0.252		
[56]	5.923	11.586	1.956	2.550	2.543	0.251		
$MgSiAs_2$	5.869	10.545	1.796	2.589	2.339	0.285	8.726	6.563
[38]	5.964	10.825	1.815	2.637	2.385	0.286		
[56]	5.954	10.800	1.814	2.633	2.380	0.286		
$MgGeAs_2$	5.908	10.963	1.855	2.586	2.416	0.274	8.785	7.235
[38]	6.015	11.284	1.876	2.635	2.473	0.273		
[56]	6.009	11.270	1.876	2.632	2.471	0.273		
$MgSnAs_2$	6.025	11.794	1.957	2.587	2.594	0.249	8.881	8.959
[38]	6.145	12.089	1.967	2.639	2.654	0.248		
[56]	6.138	12.057	1.964	2.637	2.648	0.249		

Table 2.1: Calculated lattice parameters (a and c), bond length (d), distortion parameter (u, dimensionless), and tetrahedron volume (V^T). The first, second, and third row of each material correspond respectively to the theoretical results of the present work (LDA), of Ref. [38] (PBE), and of Ref. [56] (PBE). The fourth row of $MgSiP_2$ corresponds to experimental values from Ref. [60].

The first observation is the difference between the lattice parameters and interatomic distances of the present work and the ones of other works, which are constantly larger. It makes sense since LDA and GGA tend to overbind and soften the interatomic bonds, respectively [61]. In fact, this justifies why the input for the HSE bandgap calculations was

relaxed with PBEsol functionals. However, it was not necessary to realize all calculations with such functionals because the resulting optical quantities were similar with both type of functional when including a scissor shift.

The second observation is the quantitative agreement, in the case of $MgSiP_2$, of the c/a ratio of the present work with the experimental one in contrast to the theoretical values of other works. This can be due to the higher accuracy of the convergence of the structure.

The third observation is the quantitative agreement of the distortion parameter for all works despite their structural differences. This implies that the relation between a, $d_{\mathrm{II-V}}$, and $d_{\mathrm{IV-V}}$ stays the same no matter the type of functional.

The fourth observation is the increase of both lattice parameters and interatomic distance when changing the group V atom from P to As. The c/a ratio behaves similarly even though the augmentation is relatively smaller. However, this replacement decreases the distortion parameter although relatively less than the interatomic distance. The substitution of the group IV element in the order Si-Ge-Sn seems to increase both lattice constants, to increase the ratio c/a, to decrease the Mg-V interatomic distance (with the exception of $MgGeAs_2$-$MgSnAs_2$) and to decrease u.

2.2.2 Electronic structures

The bandgaps of the present work are compared to other references in Table 2.2. As

Compound	LDA	HSE	GGA [38, 56]	HSE [38]	GW_0 [56]	Exp. [62]
$MgSiP_2$	1.13	1.94	1.38 1.37	2.04	1.83	2.20
$MgGeP_2$	1.25	2.07	1.54 1.52	2.19	2.28	
$MgSnP_2$	1.28	2.14	1.21 1.20	2.00	2.05	
$MgSiAs_2$	1.05	1.75	0.91 1.26	1.86	1.40	
$MgGeAs_2$	0.73	1.51	0.57 0.57	1.29	1.27	
$MgSnAs_2$	0.49	1.25	0.42 0.42	1.09	1.07	

Table 2.2: Comparison of the bandgaps in eV obtained through different means. The two first columns correspond to the results of the present work. All of the bandgaps are direct.

expected, the predictions of LDA and GGA largely underestimate the more precise values obtained via HSE or GW_0. Since E_g is linked to the SHG response via the refractive index and Miller's relation, this supports the need for a scissor shift. However, the HSE gaps of

the present work do not match the one reported in Ref. [38]. After testing some hypothesis, it was determined that the disagreement is only due to the different convergence of the cell parameters. The latter should be more accurate in the present work since PBEsol functionals were used with a more stringent convergence threshold than in Ref. [38]. Unfortunately, only one experimental value is available and is thus not enough to verify this assumption. In the present work, the predicted gaps show the following order: $MgSnAs_2 < MgGeAs_2 < MgSiAs_2 < MgSiP_2 < MgGeP_2 < MgSnP_2$ which differs from the one in Ref. [38] where $MgSnP_2 < MgSiP_2 < MgGeP_2$. Replacing P by As decreases the bandgap while increasing the atomic number of the group IV atom increases it when the group V atom is P and decreases it when the group V atom is As. It is also worth remembering that the bandgap gives a good indication of the LDT which means that the materials containing P should be able to sustain stronger fields.

The bandstructures of the six compounds are reported in Figure 2.2a. It shows that all of the gaps are situated at the Γ point. The high symmetry k-points were chosen in order to ease the comparison with the graphs of Ref. [38]. As in the latter, the gaps can be characterized as illustrated in Figure 2.2b. A pseudo-direct transition corresponds to a

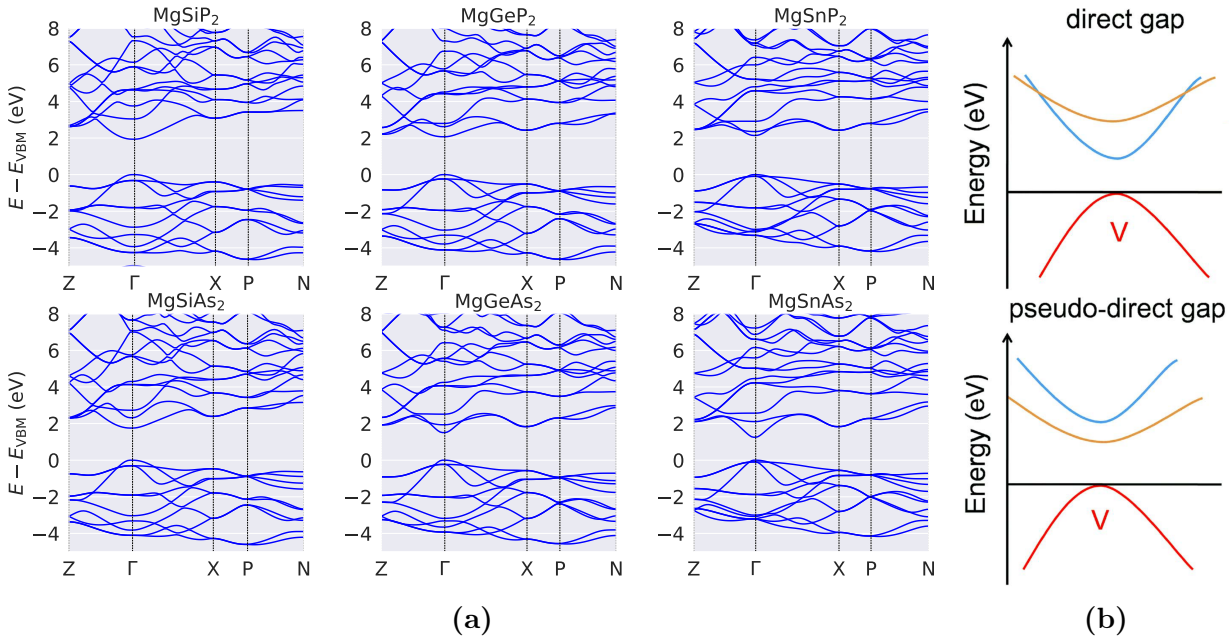


Figure 2.2: (a) Bandstructures of the six compounds of interest. VBM stands for valence band maximum. (b) Schematic diagram of bandgaps of chalcopyrites. Taken from Ref. [38].

direct gap which originates from an indirect one in the zincblende binary analog. Indeed, the transition from zincblende to chalcopyrite is accompanied by the doubling of the unit cell along one direction which is equivalent to reducing the Brillouin zone. This implies that many points in the zincblende Brillouin zone can be mapped onto the same point which results in indirect gaps being mapped to direct ones in the chalcopyrite ternary compound. However, this type of gap corresponds to a forbidden transition by symmetry selection rules

and thus to a very low optical absorption [63, 64, 65]. It can be seen from Figure 2.2a that $MgSiP_2$, $MgGeP_2$, and $MgSiAs_2$ present a pseudo-direct gap derived from an indirect transition in their II-V binary analog.

Figure 2.3 depicts the projected density of states (PDOS) for the six compounds of interest. Four different regions can be defined to ease the analysis (see caption of Figure 2.3). Overall, the orbital decomposition is qualitatively similar for the six materials. In VB2, the mixture of IV-s, V-s, and V-p components dominates. Although it is also present in VB2, the V-p component far surpasses all others in the VB1 region with the second largest one being IV-p. The CB1 section is mostly composed of the V-p and IV-s orbitals. Their strong hybridization points out the strong covalent nature of the IV-V bonds. Meanwhile, the largest component in the CB2 region is the IV-p one which is mixed with the V-p one. This last observation must be nuanced by the MgSn-V_2 materials for which the V-p orbital still dominates. Except for this last feature, Figure 2.3 qualitatively agrees with its analog from Ref. [38].

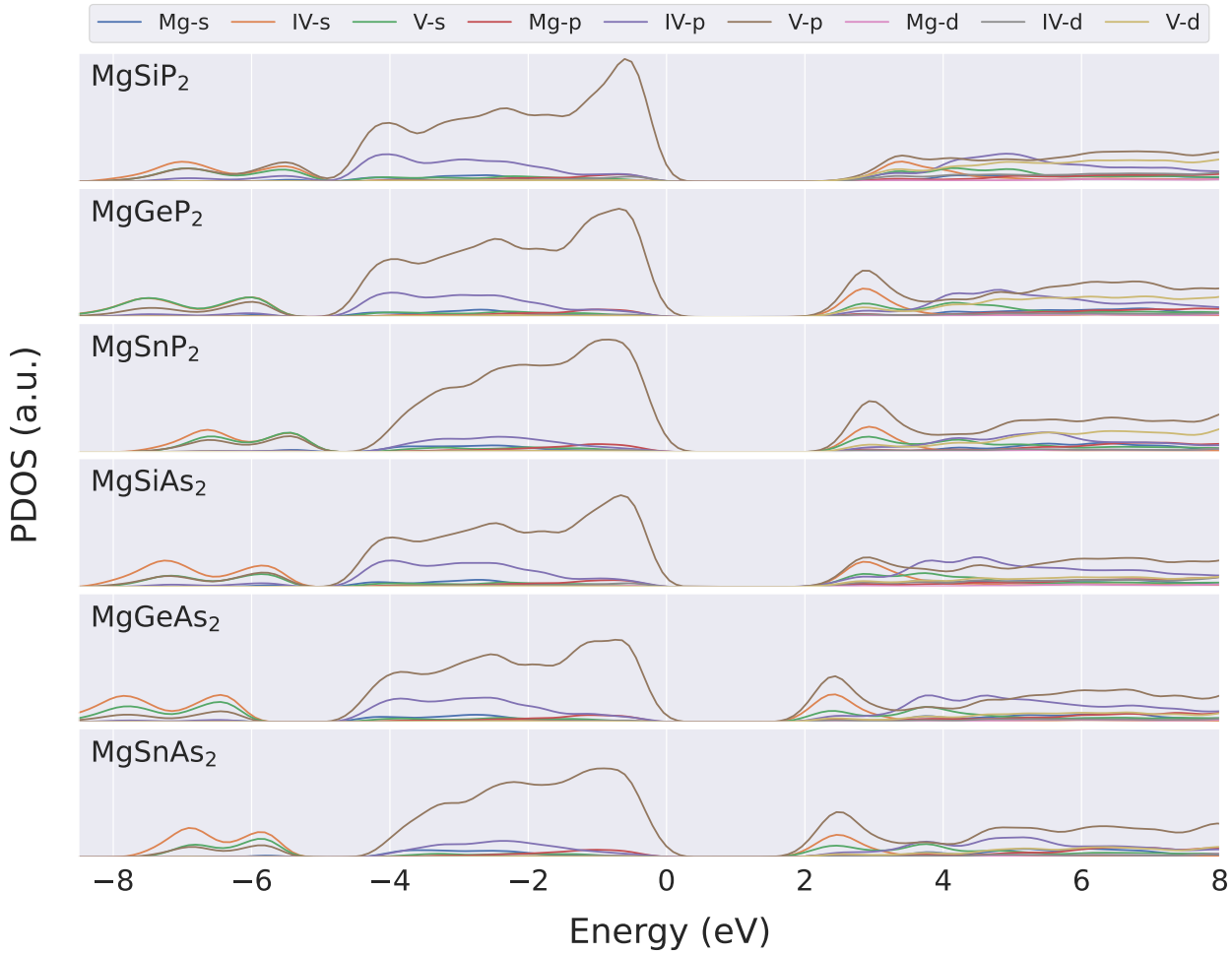


Figure 2.3: Calculated projected density of states with the Fermi levels set to 0 eV for the Mg-IV-V_2 compounds of interest. Four regions can be defined to ease the description (see text): VB2 from -8 to -4.5 eV, VB1 from -4.5 to 0 eV, CB1 from 1.5 to the intersection of V-p and IV-p, and CB2 for the rest.

2.2.3 Linear optical properties

The low optical absorption related to the pseudo-direct gap can be seen from the imaginary part of the dielectric constants as illustrated in Figure 2.4. Indeed, the blue arrows indicate small shoulders which correspond to the transitions from the valence band maximum (VBM) to the conduction band minimum (CBM). This feature is directly present after the direct bandgap in the case of $MgSnP_2$, $MgGeAs_2$, and $MgSnAs_2$. However, it is absent in the case of the pseudo-direct gaps. To be more precise, in the case of $MgGeP_2$ and $MgSiAs_2$, a very slight shoulder is present far from the gap. It corresponds to the (VBM)-(CBM+1) transitions.

Figure 2.4 also shows that the dielectric tensor contains only two independent principal dielectric constants, namely $\epsilon_{xx} = \epsilon_{yy}$ and ϵ_{zz}. The chalcopyrite structure can thus be classified as an uniaxial crystal as could have been predicted by its tetragonal nature. This implies that, in the desired range of frequency, angular phase matching may be feasible depending on the magnitude of the birefringence. Figure 2.5 shows the latter as a function of the energy. As mentioned, the region of interest is the mid-infrared one which runs

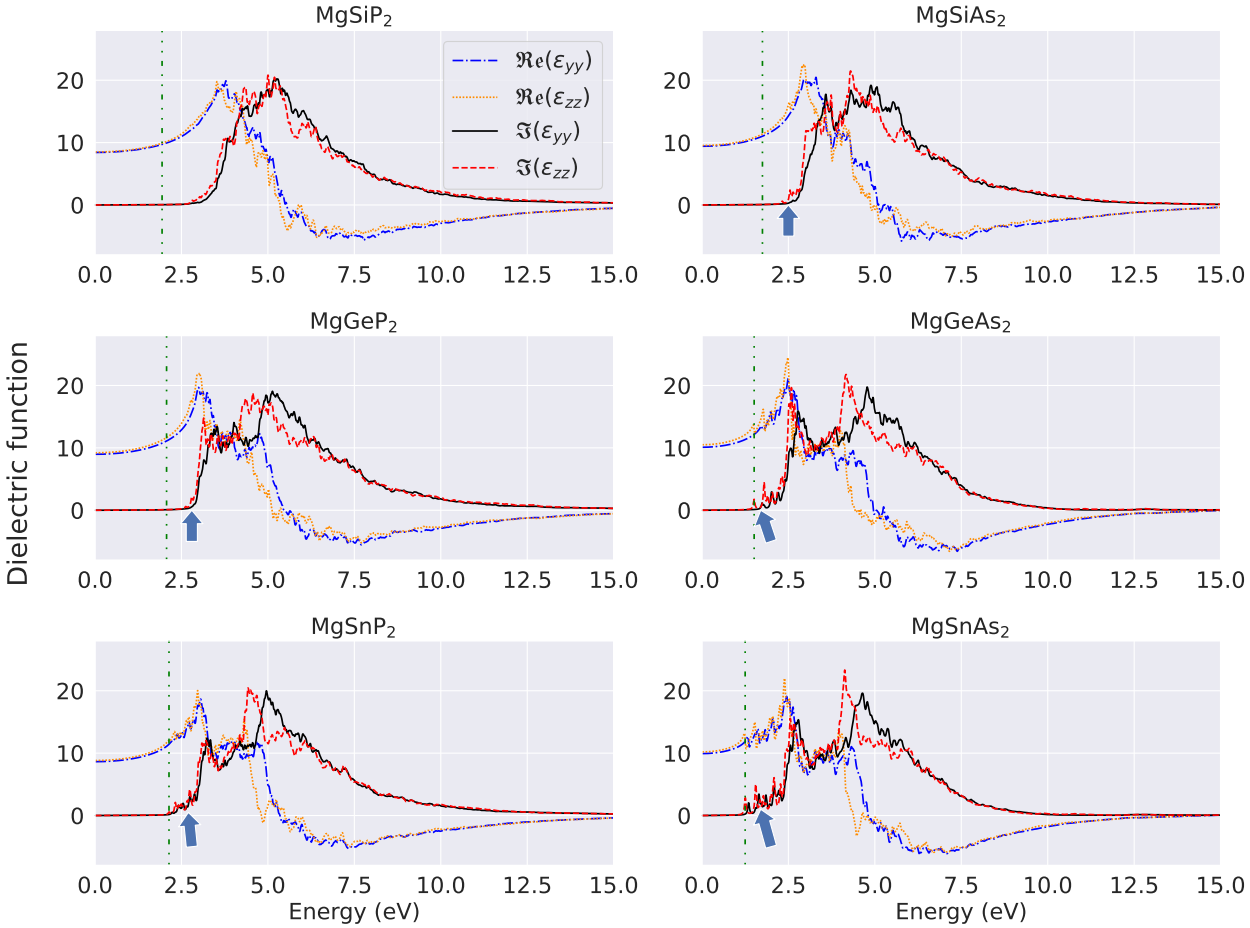


Figure 2.4: Calculated dielectric constants (dimensionless). The green dashdotdotted line indicates the bandgap. The blue arrows indicate the first VBM to CBM transitions.

from 2 to 25 μm (0.05 to 0.62 eV). In the latter, the birefringence increases slightly with increasing energy. After this region, the augmentation becomes slightly stronger, especially for $MgGeAs_2$ and $MgSnAs_2$. Qualitatively, one could consider that $MgSiP_2$ and $MgSiAs_2$ are not suited for angular phase matching due to their low value of birefringence. Indeed, the birefrigence allowing angular phase matching ranges from 0.03 to 0.1 in practice [66].

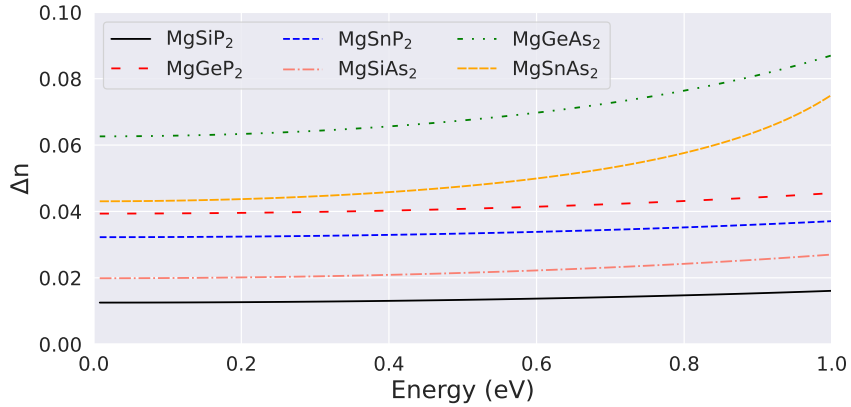


Figure 2.5: Calculated birefringence (dimensionless) as a function of energy in the mid-infrared region (2-25 μm$\leftrightarrow$0.05-0.62 eV).

The static ($\omega \to 0$) values of the dielectric constants as well as the static refractive index are shown in Table 2.3. It can be seen that the agreement between the IPA and the more accurate DFPT values is quite good. Although the refractive index is inversely proportional

Compound	$\epsilon_{s,yy}$	$\epsilon_{s,yy}^{DFPT}$	$\epsilon_{s,zz}$	$\epsilon_{s,zz}^{\mathrm{DFPT}}$	n_s^{DFPT}
$MgSiP_2$	8.42	7.67	8.50	7.86	2.78
$MgGeP_2$	8.97	8.20	9.21	8.65	2.89
$MgSnP_2$	8.63	7.84	8.82	8.30	2.83
$MgSiAs_2$	9.41	8.73	9.54	9.04	2.97
$MgGeAs_2$	10.10	9.74	10.50	10.50	3.16
$MgSnAs_2$	9.93	9.53	10.20	10.21	3.12

Table 2.3: Calculated static dielectric constants and the related refractive index via either the independent-particle approximation as implemented in the Optic utility or via DFPT.

to the bandgap, their respective orderings are not perfectly inverted. Indeed, the static refractive index follows the order: $MgSiP_2 < MgSnP_2 < MgGeP_2 < MgSiAs_2 < MgSnAs_2 < MgGeAs_2$. Replacing P by As increases n_s while increasing the atomic number of the group IV atom does not seem to be related to a particular effect. Indeed, going from Si to Ge and then to Sn, the index increases and then decreases. From Miller's relation, the best NLO crystals should be $MgGeAs_2$ and $MgSnAs_2$. Indeed, small differences in the refractive index have a large impact on the SHG response since Miller involves ϵ^3 which corresponds to n^6. In general, it is very interesting, for optical materials, to present simultaneously a high

refractive index and a wide bandgap. This led to the development of several models that describe the inverse relationship between these two quantities. A recent addition to this list of models consists in a mapping procedure of the electronic structure onto a two-state system which results in the following equation [15]

$$n_s^2 = 1 + 8\pi \frac{KJ}{\omega_g^3} = 1 + \left(\frac{\omega_{\text{eff}}}{\omega_g}\right)^3, \tag{2.4}$$

where, in the new system, ω_g is the transition energy, K its probability, and J is the product of the degeneracy of the two states. In the original system, they correspond respectively to the weighted average of the optical transitions (an average optical gap), the average probability of the transitions and the integral of the joint density of states. An effective frequency, ω_{eff}, is defined in order to ease the analysis. The average optical gap is linked to the bandgap by

$$\omega_g = E_g + \Delta, \tag{2.5}$$

where Δ thus depends on the distribution in energy of the valence and conduction bands as well as their dispersion. Indeed, ω_g is, among others, inversely proportional to the effective mass of the transitions. Figure 2.6 illustrates these dependencies.

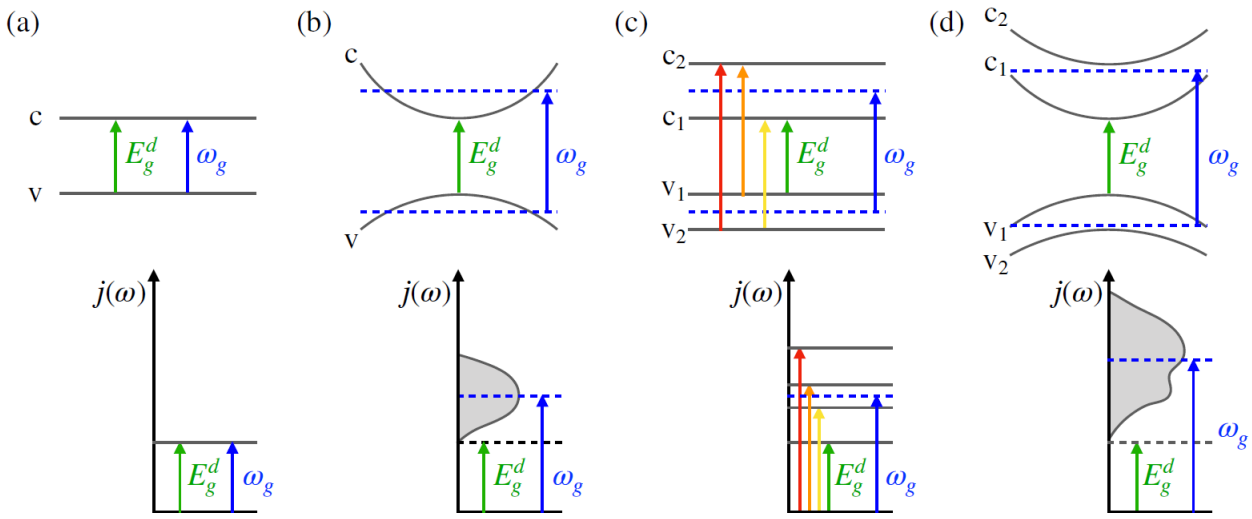


Figure 2.6: Schematic illustration of the dependence of the average optical gap ω_g on the width of the JDOS $j(\omega)$. Starting from a two-state system (a) with flat bands for which ω_g coincides with the direct band gap E_g^d and $j(\omega)$ is a Dirac peak, the graphs shows how ω_g is affected by (b) the dispersion of the bands which increases the width of the JDOS, (c) the band distribution in energy when new bands are added leading to new Dirac peaks in $j(\omega)$, and (d) the combination of both. Taken from Ref. [6].

The calculation of the different quantities introduced above (which will be called "optical coefficients") can give a new insight into the electronic structure and the refractive index and thus into the SHG coefficients that are yet to be presented. To that end, the JDOS were computed as presented in Figure 2.7. The ratio JDOS/ω^3 is shown as well since it

appears in the definition of K and ω_g. One can directly observe that the JDOS curves of

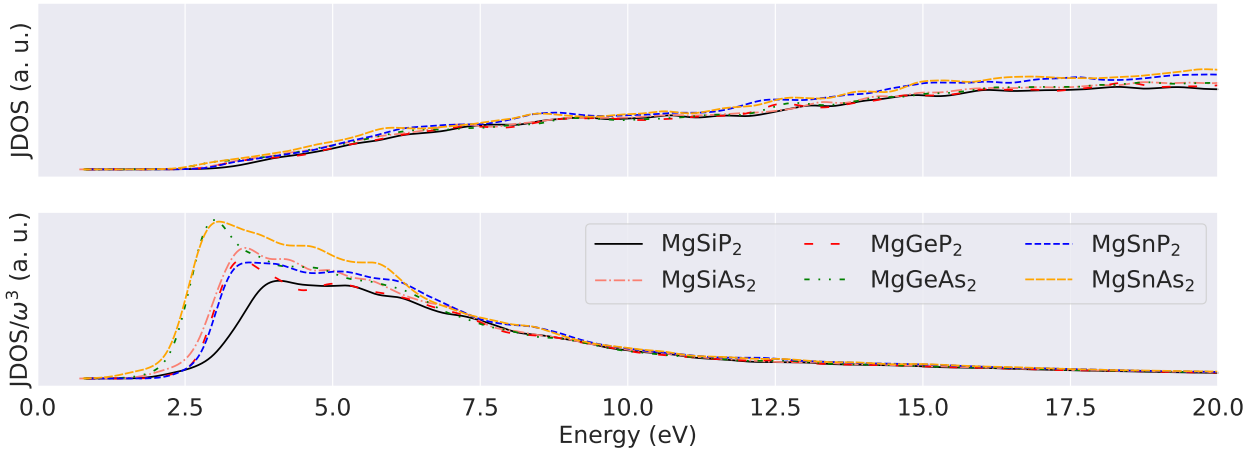


Figure 2.7: Calculated joint density of states.

the six materials follow each other such that they almost overlap. On the contrary, the ratio JDOS/ω^3 of $MgGeAs_2$ and $MgSnAs_2$ stands out by exhibiting the highest peak among all of them. They also start earlier than the others since they have the lowest bandgaps. Overall, they present the JDOS/ω^3 with the largest maximum and smallest width while $MgSiP_2$ displays the lowest ratio.

Table 2.4 contains the optical coefficients that were extracted from the electronic structure of each material following the equations in Ref. [15]. It is important to note that the value of K (and thus of ω_{eff}) were rescaled in order to agree with the more accurate DFPT static refractive index. The average probability of transition follows the same trend as the static

Compound	K (a.u.)	J (a.u.)	ω_{eff} (eV)	ω_g (eV)	Δ (eV)	E_g (eV)	n_s
$MgSiP_2$	0.486	272.563	14.932	7.908	5.968	1.94	2.78
$MgGeP_2$	0.497	241.111	14.437	7.426	5.356	2.07	2.89
$MgSnP_2$	0.428	271.219	14.292	7.473	5.333	2.14	2.83
$MgSiAs_2$	0.501	192.151	13.420	6.758	5.008	1.75	2.97
$MgGeAs_2$	0.564	130.071	12.264	5.897	4.387	1.51	3.16
$MgSnAs_2$	0.492	138.306	11.960	5.802	4.552	1.25	3.12

Table 2.4: Optical coefficients calculated as in Ref. [15]. The HSE bandgap and the static refractive index (dimensionless) were added for the sake of the reader.

refractive index, i.e., it increases when replacing P by As while it increases and then decreases when switching from Si to Ge and then to Sn. The integral of the JDOS adopts the inverse trend as the one just described. Replacing P by As and increasing the atomic number of the group IV atom both lower the value of the effective frequency. The same trend is observed for the average optical gap and its difference from the bandgap, Δ, except for the transition from Ge to Sn when the group V atom is P and As, respectively. $MgGeAs_2$ and $MgSnAs_2$ stand out from the other materials with their very low J, low ω_{eff}, ω_g, and Δ. Their value of

J is particularly low and corresponds to a decrease of about 30% from the lowest value of the other compounds. This is explained by their low upper frequency limit of the JDOS integral which is based on the imaginary part of the dielectric constant as defined in Ref. [15]. Indeed, $\mathfrak{I}(\varepsilon_{ii})$ decreases faster for the Mg-IV-As_2 materials as shown in Figure 2.4. Moreover, their small Δ indicates the low dispersion and distribution of their bands. $MgGeAs_2$ displays the highest K among all of them. Section C.1 contains figures representing each JDOS and their corresponding optical coefficients.

2.2.4 Nonlinear optical properties

As mentioned in Section 1.6, the symmetry of the crystal directly impacts the form of the d_{ijk} tensor. In the present case, only two components are independent and non-zero, namely $d_{14} = d_{25}$ and d_{36} [67]. However, they are both equal under Kleinman symmetry condition. It follows that the SHG tensor has one unique component: $d_{14} = d_{25} = d_{36}$. This condition is assumed to be valid considering the present interest for the static value. The results of the IPA calculations are shown in Figure 2.8 for the six compounds. The DFPT values are also

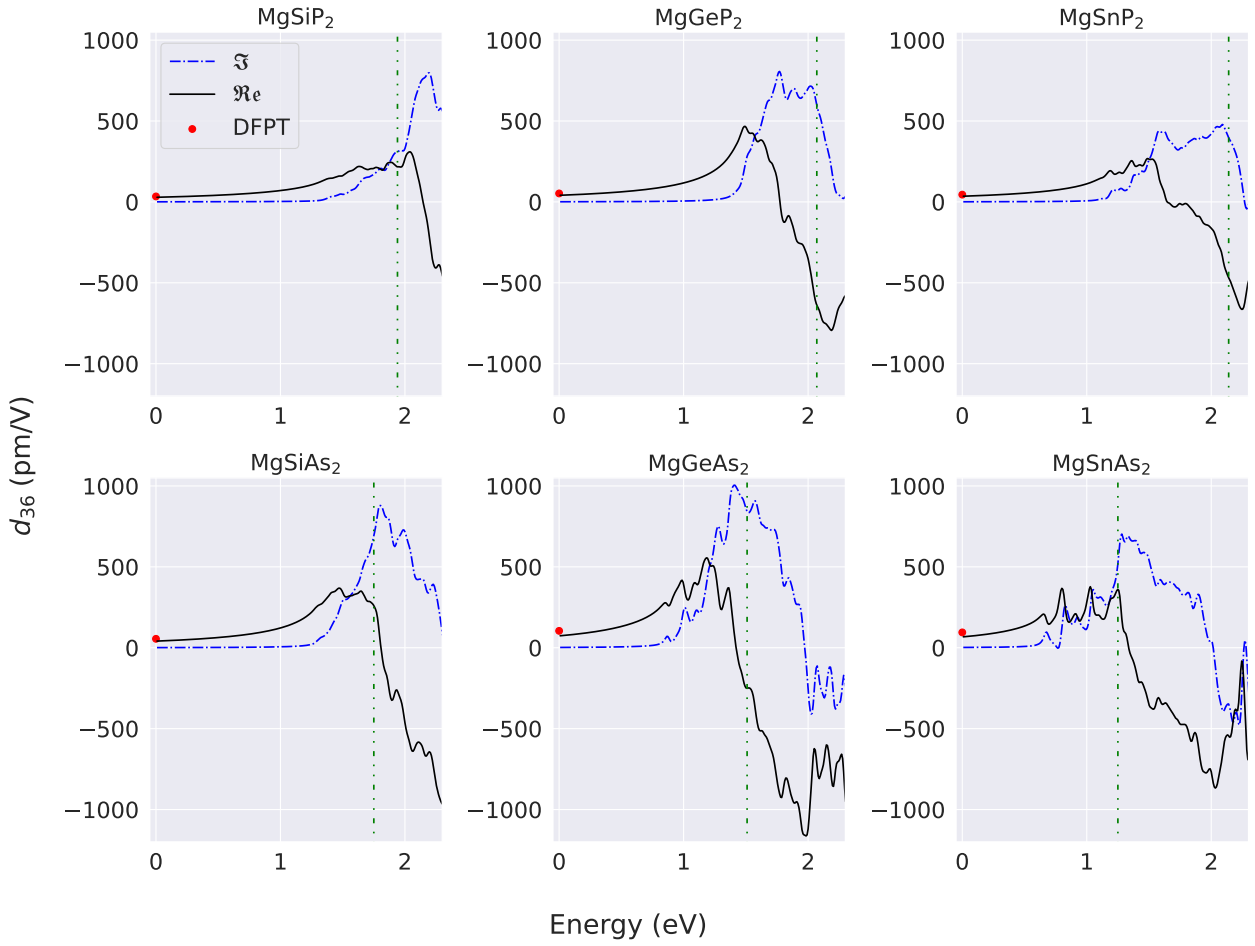


Figure 2.8: Calculated dispersion of the only non-zero component of the SHG tensor. The bandgap of each compounds is represented by the green dashdotdotted line. The energy range corresponds to the range of interest for the incident wave, i.e., the mid-IR.

represented in order to observe the agreement between both approach. Unfortunately, it has been brought to light that the Optic utility of ABINIT might contain some errors related to the calculation of the SHG tensor and its dispersion. Considering the good agreement with the DFPT values, it can be hypothesized that these errors do not affect either the low energy range in an important way or this specific crystallographic symmetry. This matter should be investigated in future works. Until then, the reader should keep this in mind when analyzing Figure 2.8.

Qualitatively, identical trends can be identified among the compounds. After a peak, the real part of the SHG coefficient drops and changes sign before reaching a new peak. Each compound is different regarding the energy where it becomes negative. The bandgap does not seem to be a factor. $MgSnAs_2$ displays three individual peaks with the same magnitude instead of a single maximum. This feature is absent from the other materials although $MgGeAs_2$ presents a relatively similar behavior. At first, the imaginary part is null as it should be when far from any absorption band. Then, it reaches a maximum around the bandgap before dropping towards negative values as can be seen in the case of $MgGeAs_2$ or $MgSnAs_2$. As mentioned, the IPA values seem to match the more accurate DFPT values. However, the latter strongly differ from the ones reported in Ref. [38] as Figure 2.9 shows. $GaAgS_2$ and $GaAgSe_2$ were used in order to assess the quality of the results. The static d_{36} value of these two benchmark materials were calculated to be 9.46 pm/V and 24.46 pm/V respectively which is relatively close to the experimental coefficients (12.2 ± 2.5 and 39.5 ± 1.9 pm/V [68]) and the ones of Ref. [38] (18.88 and 36.97 pm/V). The experimental values are underestimated and this may become worst at higher values of SHG coefficient.

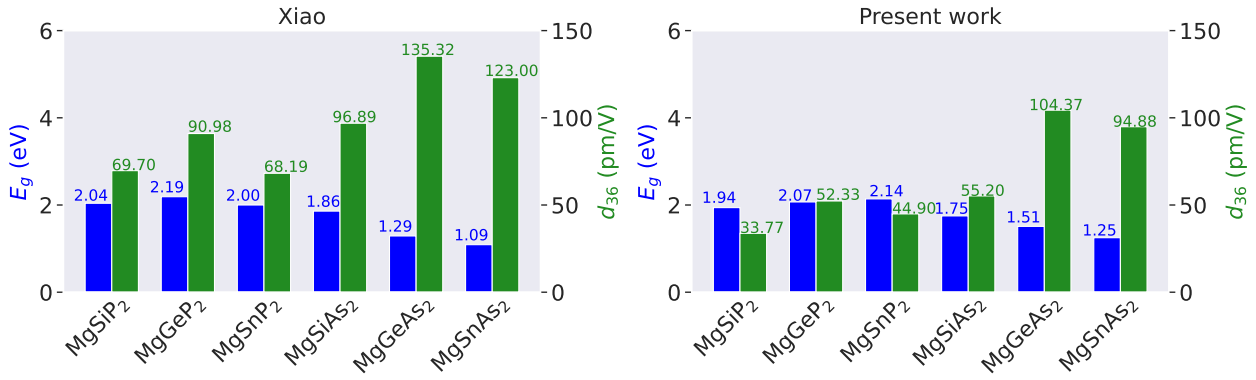


Figure 2.9: Comparison with Ref. [38] (left) of the static d_{36} obtained via DFPT. The HSE bandgaps are indicated as well. Adapted from Ref. [38].

Indeed, the results of the present work are all smaller than the ones of Ref. [38] as shown in the first column of Table 2.5. The difference is particularly important for the first four compounds. The bandgap clearly plays an important role since a lower one implies a higher refractive index and thus a higher SHG response for a fixed Miller's coefficient. However, the

present work presents lower bandgaps for some compounds, meaning that this parameter is not the only factor. If the structural differences are responsible for this disagreement, then the latter does not originate from the distortion parameter since it is the same as in Ref. [38] (see Table 2.1). Moreover, Table 2.5 is not able to highlight any direct relations between the differences of the SHG coefficient and of the structural parameters. Unfortunately, Ref. [38] does not detail enough how its SHG response was computed which limits the present understanding of this quantitative disagreement.

Compound	$\frac{d_{36}}{d_{36}^{\text{Xiao}}}$ (%)	$\frac{a}{a^{\text{Xiao}}}$ (%)	$\frac{c}{c^{\text{Xiao}}}$ (%)	$\frac{(c/a)}{(c/a)^{\text{Xiao}}}$ (%)	$\frac{d_{\text{Mg-V}}}{d_{\text{Mg-V}}^{\text{Xiao}}}$ (%)	$\frac{d_{\text{IV-V}}}{d_{\text{IV-V}}^{\text{Xiao}}}$ (%)
$MgSiP_2$	48.45	98.85	97.86	98.99	98.59	98.59
$MgGeP_2$	57.52	98.74	97.50	98.71	98.51	98.22
$MgSnP_2$	65.85	98.37	98.31	99.95	98.39	98.31
$MgSiAs_2$	56.97	98.41	97.41	98.95	98.18	98.07
$MgGeAs_2$	77.13	98.22	97.16	98.88	98.14	97.70
$MgSnAs_2$	77.14	98.03	97.56	99.49	98.03	97.74

Table 2.5: Comparison of the SHG coefficient and the structural parameters with the ones of Ref. [38].

Setting the quantitative aspect of Figure 2.9 aside, the relative trends between SHG response of the compounds are the same as in Ref. [38]: $MgSiP_2 < MgSnP_2 < MgGeP_2 < MgSiAs_2 < MgSnAs_2 < MgGeAs_2$. This ordering is the same as the one of the static refractive index which emphasizes the importance of Miller's relationship. Adopting similar arguments as in Ref. [6], Miller's coefficient can be naively seen as the proportionality coefficient between the linear and the NL optical properties. This allows to bypass the tensor notation of Eq. (1.15) by rewriting it as the following approximation

$$d_{\text{eff,KP}} = \delta\, n_s^6, \tag{2.6}$$

where the Kurtz-Perry effective nonlinear coefficient and the static refractive index were chosen to represent the NLO and linear optical properties, respectively. The link between them is provided via an effective Miller coefficient, δ. Table 2.6 contains the values of these new quantities for the six compounds of interest. Since both d_{36} (and thus $d_{\text{eff,KP}}$) and n_s have the same ordering, it follows that δ displays a similar trend. However, the increase and decrease when respectively replacing Si by Ge and Ge by Sn are relatively weaker for δ than for $d_{\text{eff,KP}}$.

In an effort to analyze the present results, Figure 2.10 eases the identification of trends relating different parameters by displaying the most relevant ones. In Figure 2.10a, K follows the same trend as n_s, δ, and $d_{\text{eff,KP}}$ while J adopts the inverse trend. With the exception of $MgGeP_2 \rightarrow MgSnP_2$, Δ also behaves similarly. In Figure 2.10b, an identical

Compound	d_{IV-V} (Å)	V_{IV}^{T} (Å^3)	d_{36} (pm/V)	$d_{eff,KP}$ (pm/V)	n_s	δ (pm/V)
$MgSiP_2$	2.517	5.709	33.77	17.394	2.78	0.038
$MgGeP_2$	2.511	6.340	52.33	26.954	2.89	0.046
$MgSnP_2$	2.513	8.002	44.90	23.129	2.83	0.045
$MgSiAs_2$	2.589	6.563	55.20	28.434	2.97	0.041
$MgGeAs_2$	2.586	7.235	104.37	53.757	3.16	0.054
$MgSnAs_2$	2.587	8.959	94.88	48.869	3.12	0.053

Table 2.6: Kurtz-Perry effective nonlinear coefficent and effective Miller coefficient for the six compounds under investigation. The IV-V bond length, the volume of the group IV atom tetrahedral unit, the only non-zero SHG component and the static refractive index (dimensionless) were added for the sake of the reader to ease the comparison later on.

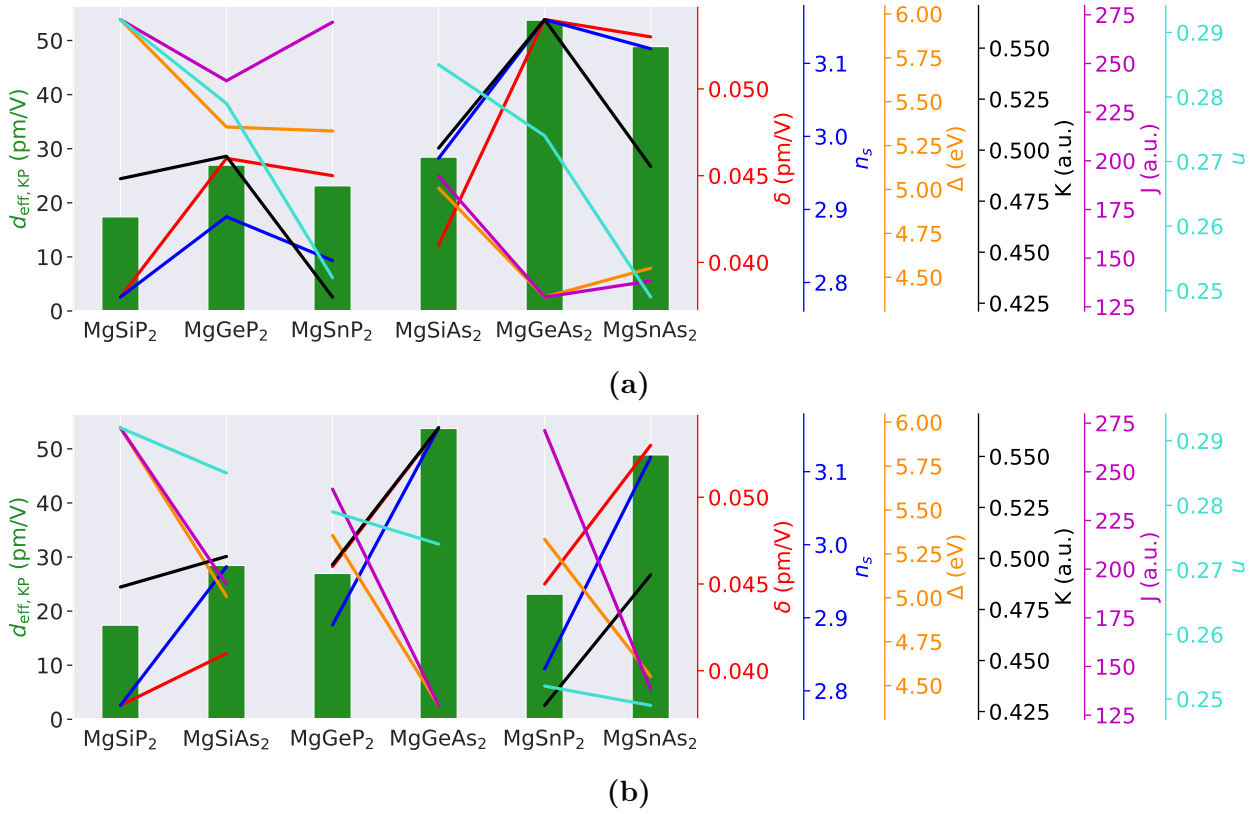


Figure 2.10: Comparison of several parameters to ease the visualization of possible trends when changing the group IV atom (a) and the group V atom (b). Both n_s and u are dimensionless.

analysis can be made. Replacing P by As increases the SHG coefficient as well as the refractive index, Miller's coefficient and K while it decreases u, J, and Δ. In both figures, the distortion parameter appears to be inversely linked to the variation of the SHG response. This seems counter-intuitive. Indeed, u, among others, describes the reduction of symmetry characterizing the passage from the zincblende to the chalcopyrite structure. Since centrosymmetric materials do not display any SHG response, one could intuitively expect that the "most non-centrosymmetric" compounds possess the highest NLO coefficient.

Based on Figure 2.10, this does not seem to be the case. Further investigations are required as will be done below.

Although not explored in the present work, a band-resolved analysis would allow to identify the orbitals contributing significantly to the SHG process [69]. In Ref. [38], $MgSiP_2$ was shown to have its SHG response mainly coming from the transitions of occupied Si-p and P-p orbitals to unoccupied Si-s and P-p orbitals. Overall, it was determined that the orbitals of the IV-V tetrahedra were mostly responsible for the SHG. This can be linked to a stronger hybridization of the orbitals (stronger covalent character) which has been proven to enhance the transitions between valence and conduction bands [70]. In the present work, it is thought that the average transition probability K reflects this phenomenon. However, replacing P by As does not change the strength of the covalent character of the bonds since their electronegativity are basically the same. Thus, this fails to explain the large increase of the SHG response as seen in Figure 2.10b.

In an attempt to quantify the relations between the various parameters of Figure 2.10, their correlation (in the sense of any dependence) matrices can be calculated. Four types of coefficient were tested in order to cover a broader range of properties. It might also compensate the small size of the dataset. In Figure 2.11, the adopted correlation coefficient, r, follows Pearson's definition such that it corresponds to [71]

$$r = \mathrm{cor}(X,Y) \equiv \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{\langle (X-\mu_X)(Y-\mu_Y)\rangle}{\sigma_X \sigma_Y} \tag{2.7}$$

$$\Rightarrow \hat{r} = \frac{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)\left(Y_i-\bar{Y}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2 \sum_{i=1}^{n}\left(Y_i-\bar{Y}\right)^2}}, \tag{2.8}$$

where $\mathrm{cov}(X,Y)$, $\sigma_{X(Y)}$, and $\mu_{X(Y)} = \langle X(Y)\rangle$ respectively indicate the covariance of two variables X and Y, their standard deviation, and their means. The second line of the above equations corresponds to the estimator of the coefficient, $\hat{r}$, when dealing with a sample of finite size n. Its means for both variables are given by $\bar{X}$ and $\bar{Y}$. The linear character of the relationship is given by r such that a value of 1 or -1 amounts to a perfect linear relation. Its sign determines the one of a regression slope between the two parameters. It is important to note that a value of 0 does not imply the independence of the variables. Indeed, it only points out the absence of any linear relationship. The limited number of data in the present case also limits the correctness of any interpretation [72]. These two last comments should be kept in mind for the following interpretation.

Unsurprisingly, most of the coefficients in Figure 2.11a are high (in magnitude). This could have been predicted since, except for u, they are all related one way or another by Miller's rule and Eq. (2.4). Since the latter is not linear, it is assumed that the small size of

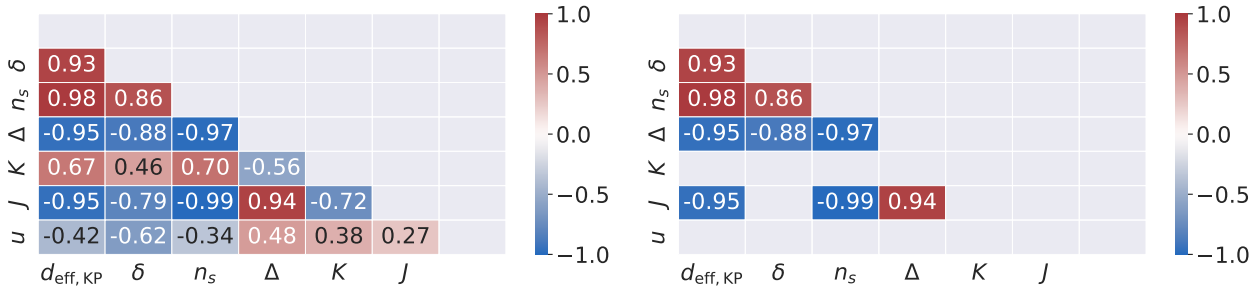


Figure 2.11: (a) Correlation matrix of the variables of Figure 2.10. The coefficient follows Pearson's definition. Since such matrix is symmetric with a trivial diagonal by construction, only its lower half has been plotted for the sake of clarity. (b) Correlation matrix of (a) without the statistically nonsignificant coefficients.

the dataset is responsible for incorrectly misleading r. Indeed, it is known that $d_{\text{eff,KP}} \propto n_s^6$ by the approximation of Miller's rule although their correlation coefficient seems to indicate a strong linear relation. The signs in this matrix also confirm the various trends identified above. Several remarks can be formulated to develop the present analysis. As previously observed, the SHG response seems to be inversely linked to the distortion parameter. Moreover, if a relationship exists between the two of them, then it is not a linear one since their r is low. However, it seems that the linear dependence of the KP coefficient on u is contained in Miller's coefficient since the latter is more linearly correlated to u although the linear character of their relationship is not particularly strong. It is also interesting to note that δ seems to be highly linear with the static refractive index which could imply that it is not strictly a spatial factor. If δ is indeed linearly related to n_s, then its strong linear dependence with the optical coefficients (especially with Δ) is either an artifact due to the number of data or originates independently from n_s. One can also notice that J is inversely linked to K and n_s while it varies similarly to Δ. The only electronic parameter, which seems to deviate from a linear dependence on $d_{\text{eff,KP}}$, is K. Moreover, most of their relation appears to originate from δ.

A more rigorous statistical point of view can be adopted by testing the null hypothesis H_0 against the alternative one H_1:

$$\mathrm{H}_0 : \; r \; = \; 0 \quad ; \quad \mathrm{H}_1 : \; r \neq 0.$$

To that goal, the statistic t is calculated such that

$$t \; = \; \frac{r}{S_r} \quad \text{with} \quad S_r \; = \; \sqrt{\frac{1-r^2}{n-2}}, \tag{2.9}$$

where $n-2$ corresponds to the degrees of freedom of the Student distribution followed by t under H_0. By a statistical two-tailed test, it comes that the null hypothesis can be safely

rejected if

$$|t| > t_{\alpha/2,n-2}, \tag{2.10}$$

where the right hand side is called the critical value. It can be found in the Student's t table. The significance level, α, is the probability of rejecting H_0 when it is true. A value of 0.05 was chosen by convention [73]. This led to a critical value of 2.776 which allowed to filter out the nonsignificant coefficients, i.e., the ones that do not allow to safely reject H_0, as shown in Figure 2.11b [74]. This implies that the previous interpretation of the now absent coefficients is likely to be a pure speculation since r might actually be null in these cases.

Figure 2.12a pushes the analysis one step further by considering Spearman's coefficient, ρ. The latter is defined by Eq. (2.7) with X and Y being replaced by their ranks R_X and R_Y such that R_{X_i} (R_{Y_i}) is 1 if X_i (Y_i) takes the smallest value, 2 if it takes the second smallest one, etc. This coefficient can be estimated for a finite sample of size n by

$$\hat{\rho} \;=\; 1 - \frac{6\sum_{i=1}^{n}\left(R_{X_i} - R_{Y_i}\right)^2}{n\left(n^2-1\right)} \;=\; 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n\left(n^2-1\right)}. \tag{2.11}$$

If multiple instances are equal, then they are given a non-integer average rank. The definition of the estimator must also be corrected accordingly [73]. Contrarily to Pearson's r, Spearman's ρ characterizes the degree to which the relationship between two variables is monotonic. Its sign is thus positive when large (small) values of X tend to be associated to large (small) Y. Similarly, it is negative when large (small) X are linked to small (large) Y. It ranges from -1 to 1 which respectively indicate a perfectly inverse and similar ordering of the ranks (perfect monotonic relation).

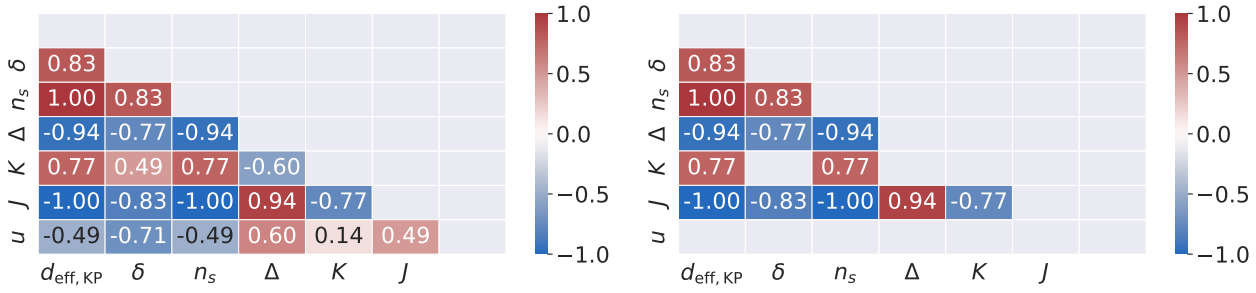


Figure 2.12: (a) Correlation matrix of the variables of Figure 2.10. The coefficient follows Spearman's definition. Since such matrix is symmetric with a trivial diagonal by construction, only its lower half has been plotted for the sake of clarity. (b) Correlation matrix of (a) without the statistically nonsignificant coefficients.

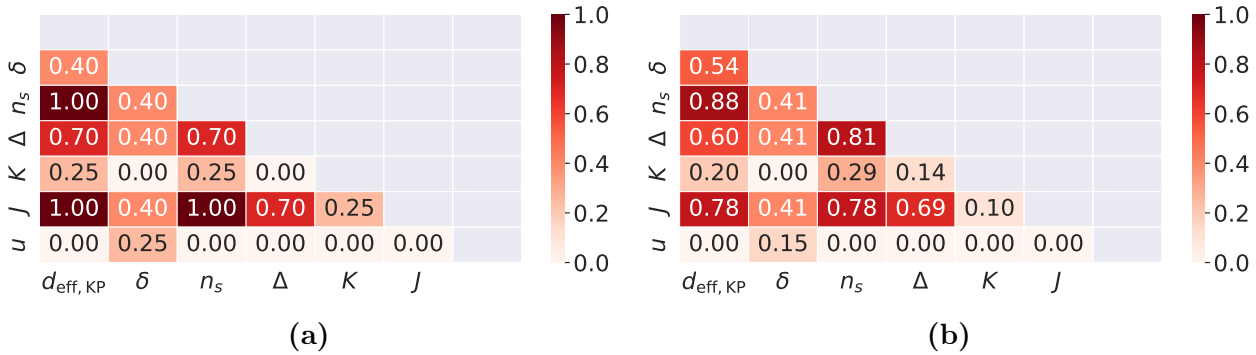
As expected, the coefficients in Figure 2.12a present the exact same signs as their Pearson's analog. The overall picture is also the same. An interesting feature is the perfect monotonic dependence between the KP coefficient, the static refractive index, and the

integral of the JDOS. The parameter Δ also seems to be strongly linked to $d_{\text{eff,KP}}$ (via n_s) while K is once more the electronic parameter with the minimum correlation to the SHG response. Once again, Δ and J seem to be highly correlated. The coefficients of the distortion parameter are relatively low except for the one with δ. As before, the null hypothesis, H_0: $\rho = 0$, can be tested against the alternative one, H_1: $\rho \neq 0$, by rejecting the former when $|\rho| > \rho^*_{n,\alpha/2}$, where the last quantity is a critical value given by the Spearman table (see Ref. [75]) for a significance level α. The latter was chosen to be 0.05 by convention which resulted in $\rho^*_{n,\alpha/2} = 0.738$. Figure 2.12b shows the correlation matrix of the Spearman's ρ without the nonsignificant values. Contrarily to Figure 2.11b, the coefficients of δ-J, $d_{\text{eff,KP}}$-K, and n_s-K were not removed.

Recently, Ref. [76] introduced a new correlation coefficient, ξ, to measure the degree of dependence between two variables. After rearranging the sample of (X_i, Y_i) by increasing X_i and thus working with a new sample of $(X_{(i)}, Y_{(i)})$, ξ can be estimated by

$$\hat{\xi}(X,Y) \;=\; 1 - \frac{3\sum_{i=1}^{n-1} |r_{i+1} - r_i|}{n^2 - 1}, \tag{2.12}$$

where r_i is the rank of $Y_{(i)}$. If the initial sample presents ties between its instances (either in X or in Y), then the definition is slightly modified. When the variables are fully independent, ξ is 0 while it becomes 1 when they are related by an unknown function. Figure 2.13a shows the corresponding correlation matrix for the present parameters. Compared to the other



(a) **(b)**

Figure 2.13: Correlation matrix of the variables of Figure 2.10. (a) The coefficient, ξ, was normalized and follows the definition of Ref. [76]. The matrix was rendered symmetric as suggested in Ref. [76]. Since its diagonal is trivial, only its lower half has been plotted for the sake of clarity. (b) The coefficient is the normalized mutual information as described in Section 3.1.2. Since such matrix is symmetric with a trivial diagonal by construction, only its lower half has been plotted for the sake of clarity.

coefficients, ξ seems to be more consistent. Indeed, δ is seen to be equally dependent on Δ and J which makes sense since they display a high $\hat{\xi}$ together. Moreover, this link with δ is perfectly transmitted to n_s as it should be through Eq. (2.4). Similar observations can be made with other parameters. A striking difference with the previous cases is the presence of null values. They all match the nonsignificant Spearman's coefficients of Figure 2.12b

except for the δ-u one. This could definitely indicate an existing relation between the two quantities as mentioned by this concept of "degree of centrosymmetry." Unfortunately, the computation of the p-values related to ξ as introduced in Ref. [76] with a significance level of 5% showed that only the coefficient of 1 ($d_{\text{eff,KP}}$, J, n_s) is statistically significant (allows to safely reject the hypothesis that $\xi = 0$).

To end this statistical analysis, the last coefficient of correlation to be considered is the normalized mutual information as described later in Section 3.1.2. It is similar to ξ in the sense that it measures a possible relation between two variables such that 0 and 1 respectively mean no relation at all and a clear dependence. Figure 2.13b shows the corresponding correlation matrix. The overall picture is very similar to the one with ξ although it seems to present finer details. The consistency mentioned above is indeed recovered for most of the parameters but not all. Moreover, Δ and K seem to be slightly dependent on one another here. Again, u and δ appear to be slightly related.

2.2.5 Effect of the distortion parameter

As mentioned above, an intuitive hypothesis is to consider that the "most non-centrosymmetric" material should display the largest SHG response. This requires to quantify the degree of centrosymmetry which is not an easy task nor concept. In the case of the zincblende structure, the centrosymmetry is broken by the presence of two different species. Going to the chalcopyrite structure, the symmetry is further reduced by the introduction of a third specie and by the distortion of the tetrahedral units. The latter is characterized by the distortion parameter, u, as defined above. This section aims at studying its effect on the NLO response. To that end, the structures of the six compounds were modified in order to recover the zincblende structure ($c/a \neq 2$ being neglected) or more precisely a pseudo-zincblende structure. In practice, this is equivalent to displacing the atoms as illustrated by the red arrows in Figure 2.14. The applied displacements all have the same magnitude, i.e., u-0.25. It was also decided to apply a scissor such that the HSE bandgap of the previous sections is recovered for each compound.

The results of this structural modification are presented in Table 2.7. The latter is structured in the same way as Table 2.6 in order to facilitate the comparison. When adopting a (pseudo-)zincblende structure, the Mg-V and IV-V bonds have henceforth the same length which is shorter than the previous Mg-V and longer than the previous IV-V. A similar observation can thus be made for the tetrahedral units. Their volume is now smaller and larger than their chalcopyrite counterpart around Mg and the group IV atom, respectively. This analysis must be nuanced by the case of $MgSnAs_2$ which had the inverse transformation. This makes sense since its distortion parameter was lower than 0.25. It is indeed the only compound with an opposite displacement of its atoms compared to the others.

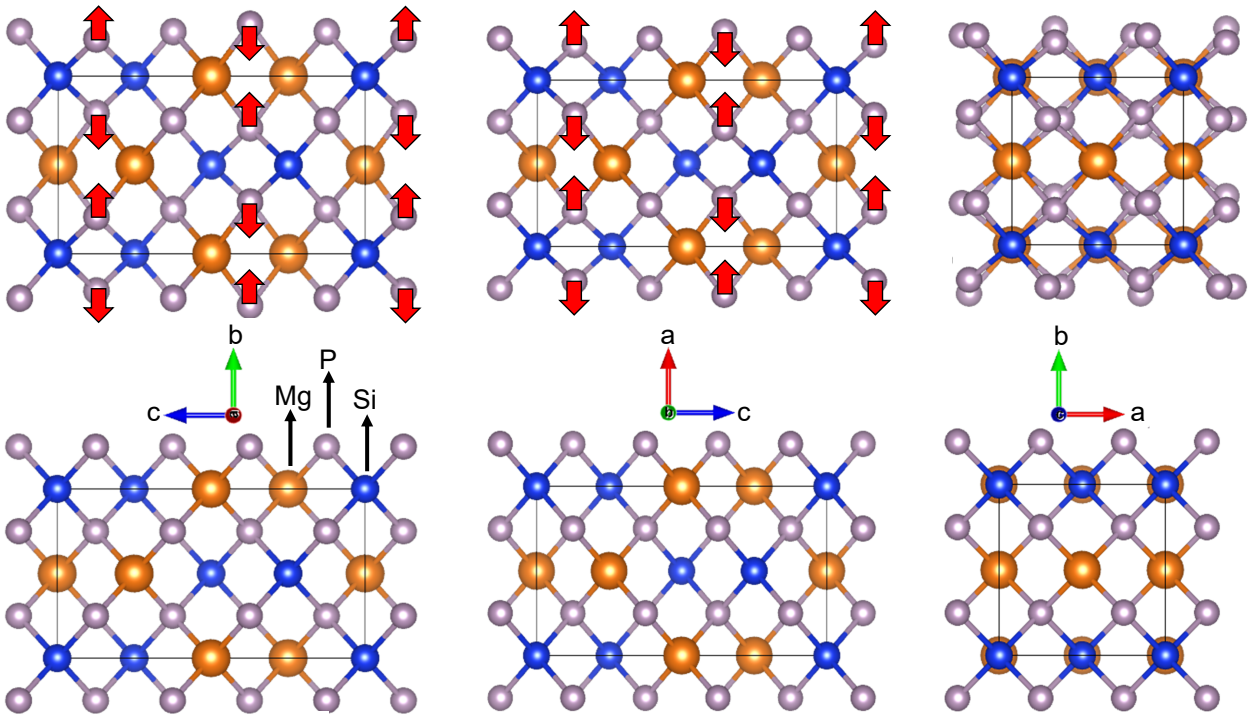


Figure 2.14: Illustration of the atomic displacements (red arrows) needed to transform the chalcopyrite structure (first row) of $MgSiP_2$ into a zincblende one (second row) when neglecting the c/a ratio and the presence of three species instead of two. The red arrows were omitted from the third rotation of the axes (on the far right) for reasons of clarity. The displacements all have the same magnitude which corresponds to u-1/4 with u the distortion parameter.

Compound	$d_{\mathrm{Mg/IV-V}}$ (Å)	V^T (Å^3)	d_{36} (pm/V)	$d_{\mathrm{eff,KP}}$ (pm/V)	n_s	δ (pm/V)
$MgSiP_2$	2.367	6.742	28.101	14.474	2.796	0.030
$MgGeP_2$	2.409	7.129	58.207	29.981	2.952	0.045
$MgSnP_2$	2.506	8.072	45.190	23.277	2.831	0.045
$MgSiAs_2$	2.458	7.567	50.302	25.910	3.002	0.035
$MgGeAs_2$	2.498	7.972	113.047	58.228	3.225	0.052
$MgSnAs_2$	2.591	8.920	94.896	48.879	3.123	0.053

Table 2.7: Calculated properties of the pseudo-zincblende structure. The volume of the tetrahedral units is designated by V^T and n_s is dimensionless.

Before discussing the other properties, Figure 2.15 illustrates the difference between the bandstructure of the original and modified crystal. $MgSnP_2$ and $MgSnAs_2$ stand out due to the superposition of both bandstructure. This was to be expected considering the weak magnitude of their atomic displacement (0.002 and 0.001 respectively). Regarding the other compounds, clear differences can be seen between both structures. The case of $MgGeP_2$ offers indeed a striking change since it presents an indirect bandgap $\Gamma - Z$ in its pseudo-zincblende form. For this material, it must thus be noted that the scissor was applied so that the new nearly direct bandgap at Γ matches the previously direct one.

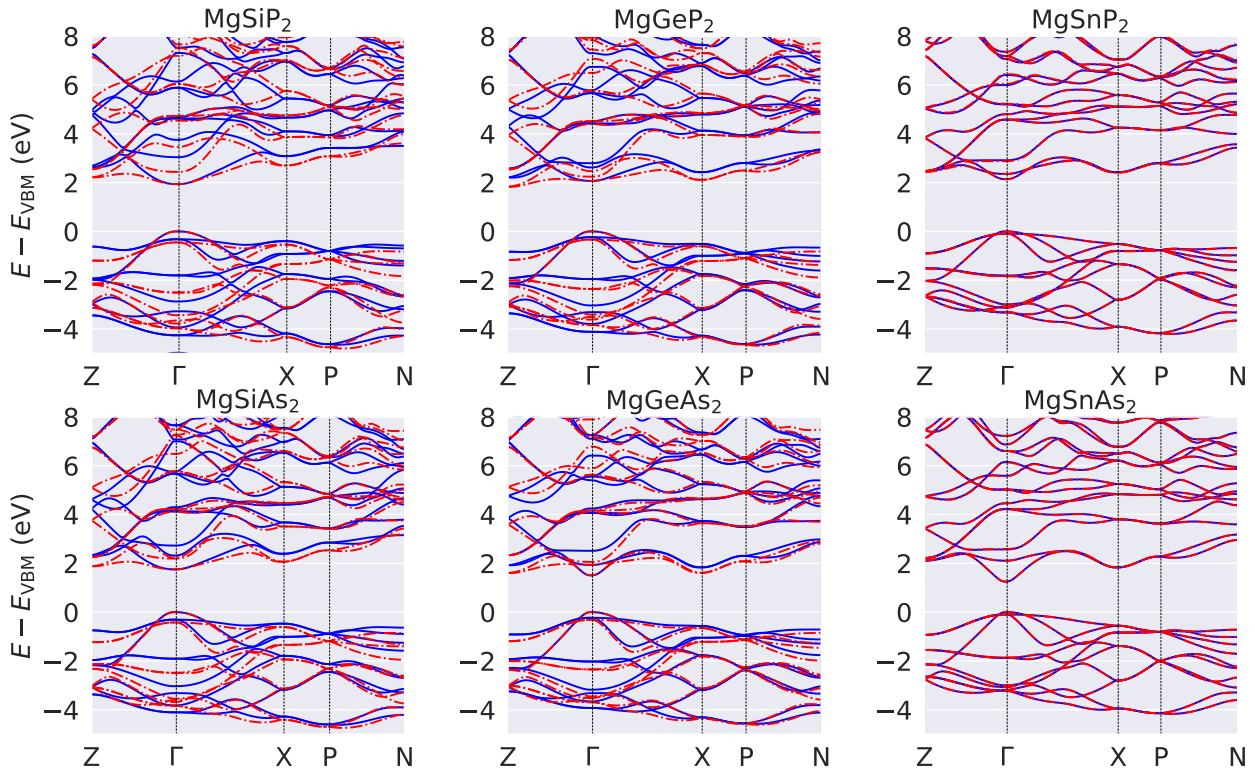


Figure 2.15: Bandstructures of the six compounds of interest. The red dotted and blue solid lines correspond to the pseudo-zincblende and chalcopyrite structure, respectively. VBM stands for valence band maximum.

Regarding the optical properties of Table 2.7, their variation is graphically illustrated in Figure 2.16 which aims at simplifying the identification of trends and their comparison with the ones of Figure 2.10. It can indeed be observed from Figure 2.16a and Figure 2.16b that separately increasing the atomic number of the group IV and group V atom results in the same variation of the SHG response than in the chalcopyrite case. The only noticeable difference is the ordering of the SHG coefficient in the pseudo-zincblende case such that $MgGeP_2 > MgSiAs_2$. Similarly, Miller's coefficient and the static refractive index both adopt the same tendencies as previously when switching the atoms. The transition from $MgGeAs_2$ to $MgSnAs_2$ is the only exception since it causes a slight increase of δ instead of a decrease as with the chalcopyrite structure.

As depicted in Figure 2.16c, the new crystals, except $MgSnAs_2$, have a higher refractive index and a lower Miller's coefficient. The SHG coefficient of $MgSiP_2$ and $MgSiAs_2$ is lower than before while the one of the other compounds is higher. Since the hybridization in the IV-V units should be weaker in the new geometry (longer bonds) and since the latter should be "more centrosymmetric" (based on our understanding), the results seem to indicate that an important factor is still missing from the picture. It also seems that MgSi-V_2 compounds are not affected by it such that they might be more influenced by their "degree of centrosymmetry." However, the latter might be a major component of Miller's coefficient

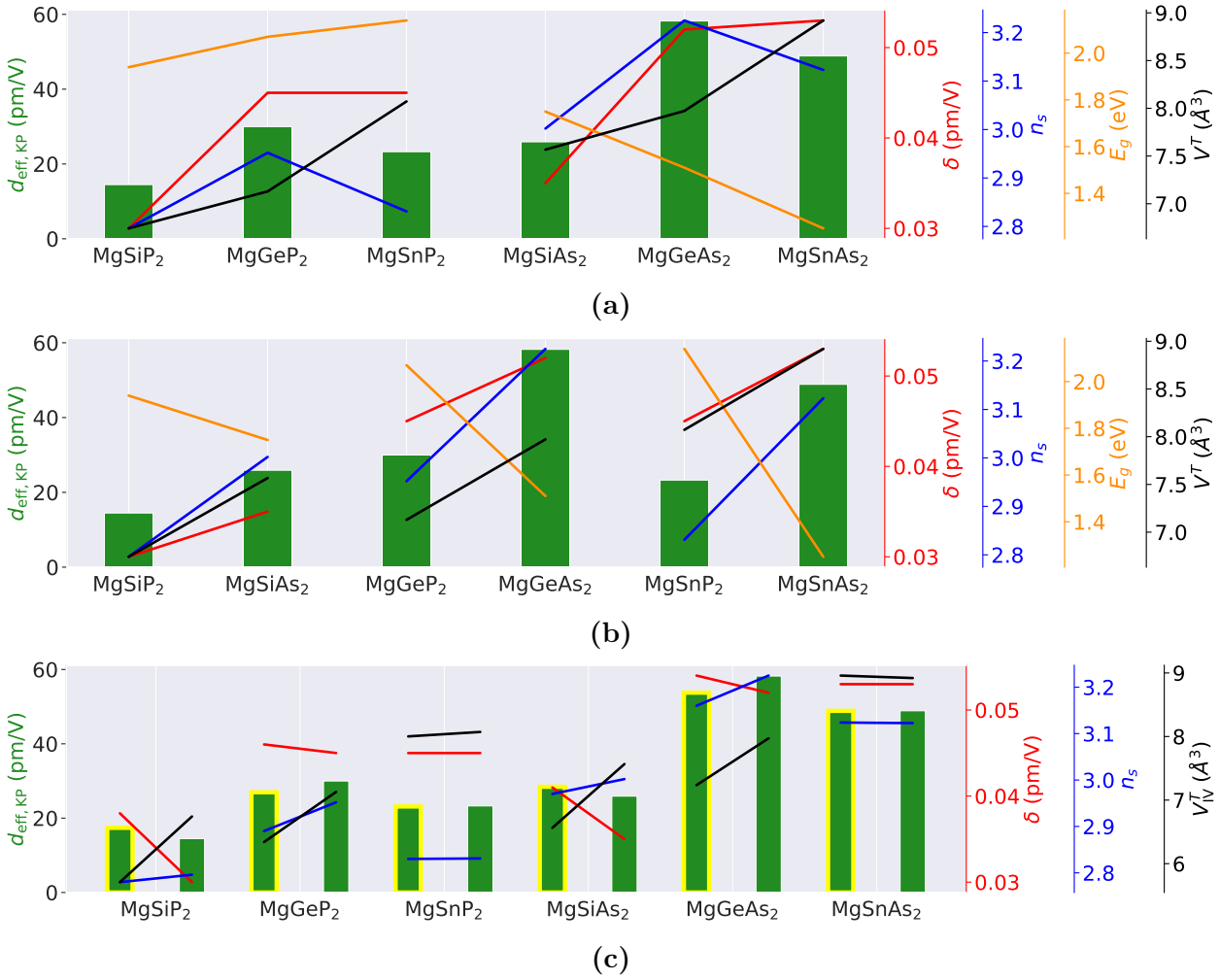


Figure 2.16: Comparison of several parameters to ease the visualization of possible trends for the pseudo-zincblende structure when changing the group IV atom (a), the group V atom (b), or when switching from a chalcopyrite structure (left bars with yellow edges) to a pseudo-zincblende one (right bars without edges) (c). The static refractive index n_s is dimensionless.

instead. This was already suggested in Ref. [6] where structures were interpolated between a non- and a centrosymmetric one. The new structures showed an increase of δ as they were further away from the centrosymmetric one. $MgSnAs_2$ thus constitutes an exception and could indicate either that the "most centrosymmetric" structure might be different than the pseudo-zincblende used in the present work or that the hybridization in the IV-V unit is more important. In order to have additional information, it would be interesting to calculate the JDOS and the related optical coefficients of the new crystals.

In order to further study the effect of the distortion parameter, "inverse u" structures were investigated. They were obtained though modification of the chalcopyrite ones such that the new distortion parameter is $u' = 0.25 - (u - 0.25)$. This corresponds to inverting the atomic displacement $u - 0.25$ which differentiate the chalcopyrite structure from the pseudo-zincblende one. The relevant results are displayed in Table 2.8. Comparing the latter with Table 2.1, it is obvious that the bond length and tetrahedral volume are also inverted

(II↔IV) as expected. Figure 2.17 graphically illustrates the effect of changing the group IV or V atom on the present results.

Compound	$d_{\mathrm{IV-V}}$ (Å)	V_{IV}^{T} (Å^3)	d_{36} (pm/V)	$d_{\mathrm{eff,KP}}$ (pm/V)	n_s	δ (pm/V)
$MgSiP_2$	2.517	7.963	29.661	15.278	2.796	0.032
$MgGeP_2$	2.511	8.017	58.264	30.010	3.001	0.041
$MgSnP_2$	2.513	8.143	45.484	23.428	2.834	0.045
$MgSiAs_2$	2.589	8.726	50.169	25.841	3.115	0.028
$MgGeAs_2$	2.586	8.785	133.842	68.939	3.378	0.046
$MgSnAs_2$	2.587	8.881	94.574	48.713	3.120	0.053

Table 2.8: Calculated properties of the "inverse u" structure. The volume of the tetrahedral units around the group IV atom isdesignated by V_{IV}^{T}. The static refractive index n_s is dimensionless.

In Figure 2.17a, it can be seen that increasing the atomic number of the group IV atom causes its tetrahedral volume and Miller's coefficient to monotonically increase. The increase/decrease trend of both the static refractive index and $d_{\mathrm{eff,KP}}$ is the same as in Figure 2.10 with the chalcopyrite structure. The consequence of modifying the group V atom is represented in Figure 2.17b. Replacing P by As increases n_s. It also decreases δ in the case of MgSi-V_2 but increases it otherwise. The chalcopyrite compounds displayed the opposite trend while it behaved in the same manner (augmentation) as the inverse structure regarding the tetrahedral volume and the nonlinear optical coefficient. The ordering of the latter is however different since $MgGeP_2 > MgSiP_2$ as seen in Figure 2.17. Except for MgSi-V_2, the inverse structure presents higher values of $d_{\mathrm{eff,KP}}$. $MgSnAs_2$ is different since its chalcopyrite form had a distortion parameter lower than 0.25 to begin with.

The evolution of $d_{\mathrm{eff,KP}}$, δ, n_s, and V_{IV}^{T} with respect to the direction of the atomic displacement, i.e., to $u <, =, > 0.25$, is illustrated in Figure 2.17c. Except for the decrease of the tetrahedral volume and the static refractive index, increasing u does not cause consistent changes of the two other parameters among the compounds. The NLO coefficient decreases except for $MgSiP_2$ and $MgSnAs_2$ where it increases in the second and first step, respectively. However, $MgSiAs_2$ only shows an increase of $d_{\mathrm{eff,KP}}$. Miller's coefficient adopts larger values with the exception of $MgSiP_2$ where it first decreases for the pseudo-zincblende structure before increasing when $u > 0.25$. Therefore, n_s does not seem to be their only other dependence considering these inconsistent variations. These observations go against the assumption of an inverse link between a "degree of centrosymmetricity" and the SHG response or Miller's coefficient. Their minimum was indeed expected in the case of the pseudo-zincblende structure. This implies that an important element is still missing. It could also be due to an improper understanding of the quantification of centrosymmetry.

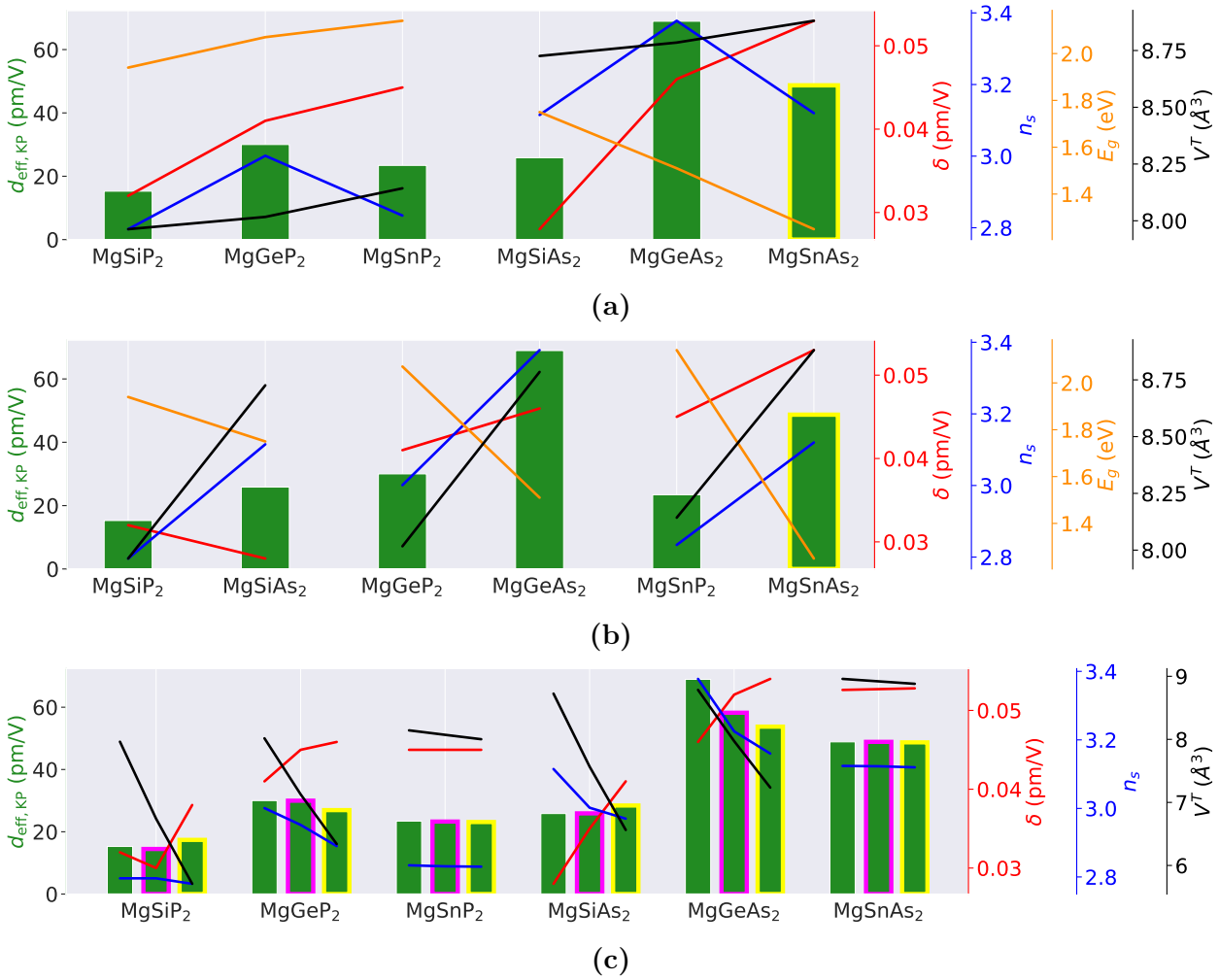


Figure 2.17: Comparison of several parameters to ease the visualization of possible trends for the "inverse u" structures when changing the group IV atom (a), the group V atom (b), or when switching from a $u < 0.25$ structure (left bars without edges) to a $u = 0.25$ one (central bars with pink edges) to a $u > 0.25$ one (right bars with yellow edges) (c). The static refractive index n_s is dimensionless while V^T corresponds to the volume of the tetrahedral units around the group IV atom.

2.2.6 Investigating binary analogs

In an attempt to further investigate this concept of "degree of centrosymmetry", binary analogs (BA) of the Mg-IV-V_2 compounds were defined and examined with regard to their NLO response. The latter is expected to be smaller than the one of the ternary chalcopyrites and pseudo-zincblende. Indeed, two different species should (based on our understanding) be "more centrosymmetric" than three of them. These new crystals possess a zincblende structure. They are comprised of III-V (BA1) or II-VI (BA2) atoms by taking the cation or anion next to the group IV atom, i.e., on the same period of the periodic table of elements. In the present case, this resulted in nine binary compounds:

II-VI	$\rightarrow$	II-IV-V_2	$\rightarrow$	III-V
MgS	$\rightarrow$	$MgSiP_2$	$\rightarrow$	AlP
MgSe	$\rightarrow$	$MgGeP_2$	$\rightarrow$	GaP
MgTe	$\rightarrow$	$MgSnP_2$	$\rightarrow$	InP
MgS	$\rightarrow$	$MgSiAs_2$	$\rightarrow$	AlAs
MgSe	$\rightarrow$	$MgGeAs_2$	$\rightarrow$	GaAs
MgTe	$\rightarrow$	$MgSnAs_2$	$\rightarrow$	InAs

When the atoms of the II-IV-V_2 crystals are not from the same period, defining binary analogs of the form $III^{a}_{0.5}$-$III^{b}_{0.5}$-V (for example) becomes a possibility [55]. However, it was not explored here as a first approach. Most of the above BA are well known materials in the field of optics. Therefore, the HSE bandgaps were recovered from the literature instead of computed with VASP. Except for this last comment, the same methods of computation as in Section 2.1 were used although only the steps 1, 3 (for the bandstructure), and 5 were performed. Table 2.9 contains the results of interest as well as some other values for comparison. Unfortunately, InAs was found to be a metal in the LDA framework. This material was thus removed from the following results and analysis due to its lack of SHG. By using GaP and GaAs as benchmark, one can once again observe that the SHG response is underestimated with respect to the experimental values.

Similarly to the previous analysis, the identification of trends is made easier thanks to Figure 2.18. In the case of the BA1, the effect of increasing the atomic number of the group III atom for a fixed group V atom is shown in Figure 2.18a. It causes the bandgap to decrease while the SHG response, the static refractive index and Miller's coefficient all increase. The volume of the tetrahedral unit around the group III atom is first lowered before being greatly augmented. Moreover, the behavior of the KP coefficient is different from the pyramidal trend it adopted with the chalcopyrite and pseudo-zincblende structure. However, as with the latter, replacing P by As leads to an increase of the SHG response as shown in Figure 2.18b. Miller's coefficient, n_s, and V^T adopt a similar tendency while E_g follows an inverse one. Raising the atomic number of the group VI atom for the BA2 also intensifies the SHG phenomenon although its magnitude is relatively low due to their large bandgap and small refractive index. It is nonetheless interesting to note the considerable enhancement of δ and V^T that accompanies the slight increase of n_s. The bandgap does not vary monotonically for this type of transition since it is larger for MgTe than for MgSe.

The impact of the evolution BA2 $\rightarrow$ Mg-IV-V_2 (pseudo-zincblende) $\rightarrow$ BA1 on the various quantities is illustrated in Figure 2.18c. It shows that, except for $MgSnP_2 \rightarrow$ InP, the SHG coefficient increases and then decreases. The static refractive index also follows this trend for the MgSi-V_2 crystals although it only increases for the other compounds. The tetrahedral

Compound	d (Å)	V^T (Å^3)	d_{36} (pm/V)	$d_{\text{eff,KP}}$ (pm/V)	n_s	δ (pm/V)	E_g (eV)
AlP	2.350	6.659	12.047 14.000[a]	6.205	2.650	0.018	1.42* 2.52*[b]
GaP	2.333	6.516	28.808 21.000[a] 41.000[c]	14.838	3.008	0.020	1.39* 2.30*[d]
InP	2.521	8.227	48.391	24.925	3.065	0.030	0.59 1.55[d]
AlAs	2.436	7.414	21.429 21.000[a]	11.038	2.828	0.022	1.32* 2.24*[b]
GaAs	2.426	7.324	78.176 45.000[a] 90.000[c]	40.267	3.388	0.027	0.49 1.50[d]
MgS	2.423	7.304	1.366	0.704	1.917	0.014	3.35 4.78[b]
MgSe	2.544	8.448	3.736	1.925	2.067	0.025	2.53 3.54[e]
MgTe	2.758	10.771	8.414	4.334	2.166	0.042	2.37 3.74[b]

Table 2.9: Calculated properties of the zincblende binary analogs. The volume of the tetrahedral units around the group III or II atom is designated by V^T while d is the distance from the group III or II atoms to the closest group V atoms. The static refractive index n_s is dimensionless. The first row of E_g corresponds to the LDA value of the present work before applying the scissor shift to match the HSE values of the second row. The asterisks indicate indirect bandgaps. The superscript [a] refers to the LDA value of Ref. [77], [b] to the HSE value of Ref. [78], [c] to the experimental value of Ref. [79], [d] to the HSE value of Ref. [80], and [e] to the HSE value of Ref. [6].

volume V^T decreases monotonically except for $MgSnP_2$ and $MgSiAs_2$ which see its increase when switching respectively to the BA1 compound and to the pseudo-zincblende structure. One of the most interesting features of Figure 2.18c is that Miller's coefficient actually follows the above expectations based on the concept of "degree of centrosymmetry." Indeed, the ternary compounds in the pseudo-zincblende structure all display a higher value of δ than their binary analogs. As touched upon in the previous sections, δ seems to be linked to this notion of quantifying the centrosymmetry based on the species and their structural arrangement. The SHG response is thus also related to it by Miller's rule although it seems to greatly depend on something else hidden by n_s as is indicated by the behavior of $MgSnP_2$ in Figure 2.18c. Further investigations are required and a perspective for future works might be to understand the drastic variation of δ in the BA2 compounds.

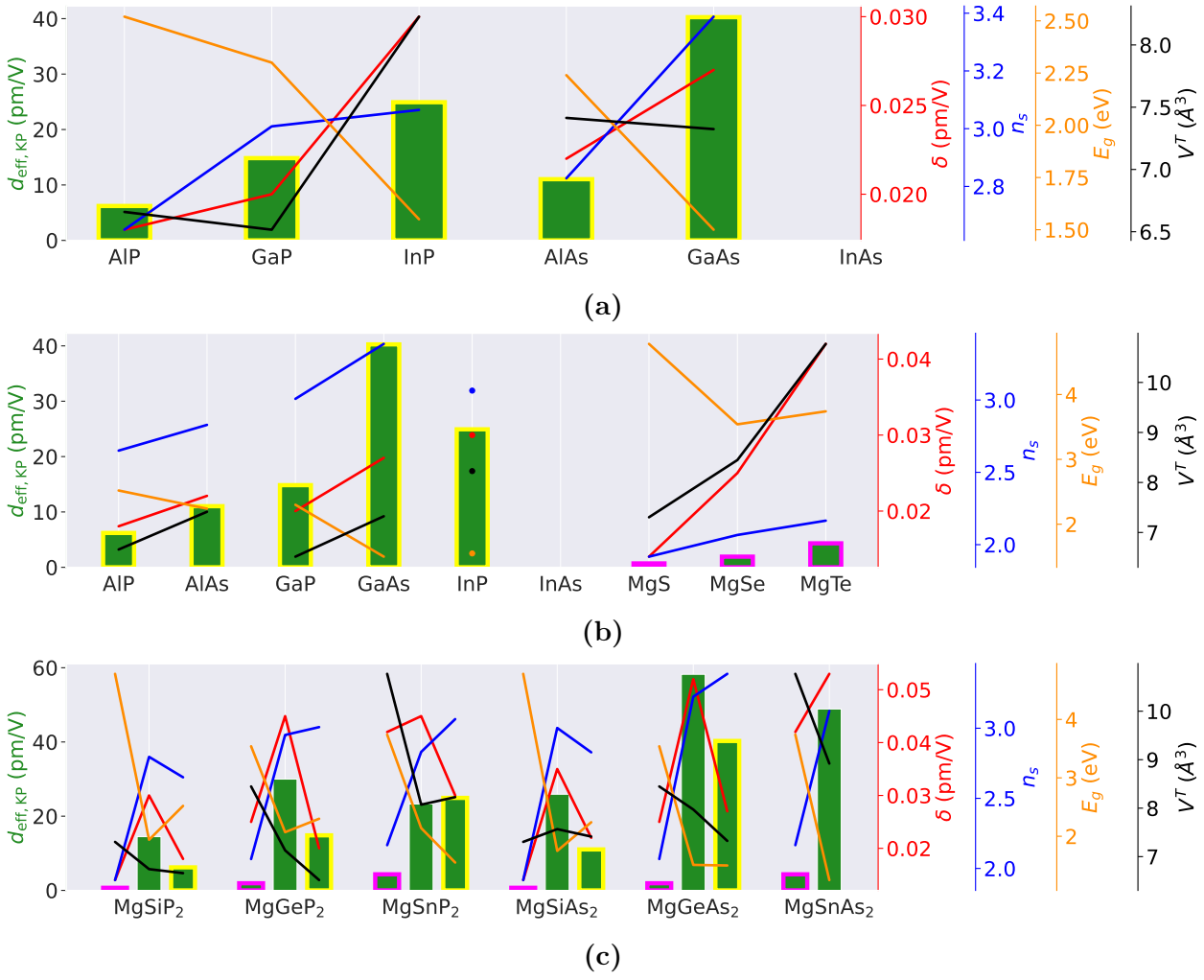


Figure 2.18: Comparison of several parameters to ease the visualization of possible trends for the binary analogs when changing the group III atom of the BA1 (a), the group V or VI atom of the BA1 or BA2 (b), or when switching from a BA2 (left bars with pink edges) to a ternary pseudo-zincblende structure (central bars without edges) to a BA1 (right bars with yellow edges) (c). The static refractive index n_s is dimensionless.

2.3 Conclusion

This second section presented the first-principles results of this work by investigating six potential NLO crystals for the mid-IR region, namely MgSi(-Ge,-Sn)P_2(-As_2). The overall method and parameters of computation were introduced in Section 2.1. In Section 2.2.1, the six compounds were found to adopt a chalcopyrite structure whose deviation from the zincblende one is characterized, among others, by the distortion parameter, u. As explained in Section 2.2.2, their link to zincblende binary analogs was also discovered in their bandstructure since MgSi(-Ge)P_2 and $MgSiAs_2$ all display a pseudo-direct gap which is characterized by a lower absorption as reflected by their absorption spectrum. All the PDOS showed the strong covalent nature of the IV-V bond while the birefrigence of $MgSiP_2$(-As_2) was found to be insufficient to allow phase matching. Section 2.2.3 also introduced "optical coefficients" to describe their electronic structure. The electronic bands of $MgGeAs_2$ and $MgSnAs_2$ display the smallest dispersion. The resulting SHG responses were analyzed in Section 2.2.4. The static DFPT results agreed well with the IPA ones. Compared to the other four compounds, MgGe(-Sn)As_2 displays a significantly higher SHG response. Correlation coefficients as well as graphical visualization of the trends both concluded that the SHG coefficient may not be strongly linked to a "degree of centrosymmetry" as assumed initially but rather to the electronic structure hidden behind the refractive index. However, Miller's coefficient might be more influenced by this quantification of the centrosymmetry. Indeed, this was more or less confirmed both when forcing the structure to a pseudo-zincblende one in Section 2.2.5 and when investigating zincblende binary analogs in Section 2.2.6 although the investigation of "inverse u" structures was less positive about it.

Chapter 3

Predicting the SHG response

The research in materials science has recently been accelerated thanks to an innovative and efficient tool, machine learning. However, the first hurdle before diving into this new paradigm is to obtain DBs with the relevant properties for the task at hand. Recently, two databases have been developed via high-throughput screening of the MP in the framework of linear and nonlinear optics respectively [6, 81]. The "linear" DB consists of 4040 semiconductors, mostly oxides, that were selected based on their thermodynamical stability ($E_{hull} \leq 25$ meV/atom) and the availability of their electronic properties. Indeed, the bandgap was directly taken from the MP, although it is strongly underestimated compared to experimental values since it results from DFT as explained in Appendix A. The static refractive index was calculated using DFPT. The "nonlinear" DB is a subset of the linear one. To begin with, the centrosymmetric compounds were filtered out due to their lack of SHG response. Then, materials with low values of both the refractive index and the bandgap were discarded. For practical reasons (exchange-correlation functionals, low accuracy, etc.), other candidates had to be removed. In the end, the nonlinear DB was made up of 457 instances (materials) for which DFPT calculations were performed to obtain their d_{ijk} tensor. More details on the full process can be found in Ref. [6]. Since the emergence of this dataset, the number of instances was increased up to 579 materials thanks to unpublished results of Naccarato *et al.*

The first section of this chapter deals with the basics of ML. The model used in the present project, MODNet, is then introduced in Section 3.1.2 with the goal of being trained on the nonlinear dataset to predict an adequate quantity as explained in Section 3.2. The analysis of the DB, the choice of features, the overall approach, and its results are respectively described in Section 3.3, Section 3.4, Section 3.5, and Section 3.6. The latter is completed by a simulation of the application of the final model as presented in Section 3.7.

3.1 Machine learning approach

Materials informatics is an emerging field which combines the expertise of materials scientists with the empirically proven benefits of machine learning (ML). This association was made possible through the development of large materials databases built from either quantum computations or experiments. They provide the necessary data to train and test the ML models which can then predict the properties of interest. The advantage is the gain

in computational resources and time. Indeed, those models can reach similar accuracy as DFT results, but at a fraction of the computational cost [82]. Supervised ML encompasses all methods that aim at predicting a certain target (e.g., a specific property) based on certain inputs, also called features (such as the crystalline structure). The basic principle is that the ML model adjusts internal parameters that relate, often in a nonlinear way, the inputs to the output in order to decrease the difference between the prediction and the known target. Throughout this training step, all instances of a training dataset are considered, hence the need for large DBs. Once all parameters have been optimized, the model can predict the target property for any compounds by using the corresponding inputs. This approach takes advantage of the capability of the machine to translate NL relationships between features into a computationally simple model.

3.1.1 General principles

Supervised learning and ANN ML tasks can be categorized as either unsupervised or supervised learning. In the former, the goal of the system is to identify patterns in datasets without the use of any targets. For example, by modifying their parameters, clustering algorithms derive implicit similarity criteria between instances of the DB in order to separate them in relevant groups. On the contrary, supervised learning aims at approaching an unknown mathematical function f which relates inputs $\mathbf{X}$, also called features or descriptors, to a specific output or target t

$$t_i \;=\; f(\mathbf{X}_i) \approx \hat{f}(\mathbf{X}_i) \;=\; y_i, \tag{3.1}$$

where the subscript indicates one of the instances of the DB and $\hat{f}$ is the function found by the system. The denominations "classification" and "regression" refer to problems where the target can take discrete and continuous values, respectively. The error between the output and the target allows to define an objective or loss function, L, such as the common mean-squared-error (MSE) for instance:

$$L_{\text{MSE}} \;=\; \frac{1}{N}\sum_{i}^{N}(t_i - y_i)^2 \;=\; \frac{1}{N}\sum_{i}^{N}(e_i)^2, \tag{3.2}$$

where N is the number of instances. The choice of this loss function is based on assumptions on the distribution of the errors and can influence the performance of the model. Indeed, the latter is fed with instances of a training dataset and modifies some of its parameters, i.e., its weights, iteratively in order to reach a minimum of this loss. Different functions can then lead to different minima. The full process is illustrated in Figure 3.1 in the case of an artificial neural network (ANN). This type of model consists of several layers of individual units, also called "neurons" or "nodes." Each one of these units performs a linear combination of its inputs and feeds it to an "activation function" which introduces some nonlinearity such as the common hyperbolic tangent. The representation of such model is given in Figure 3.2.

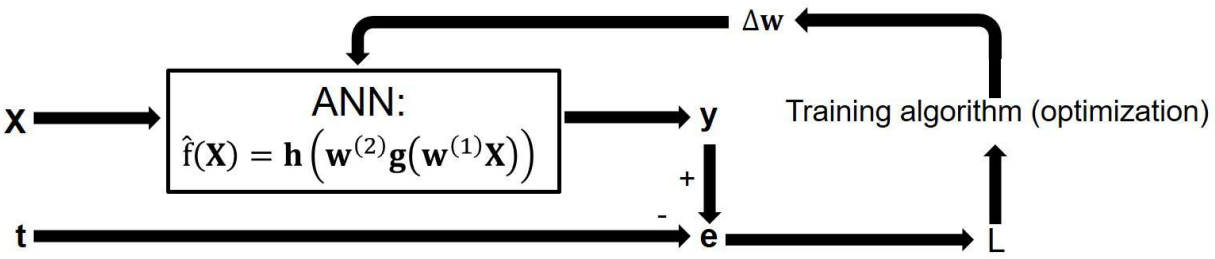


Figure 3.1: Schematic of the training process for a two-layer ANN. $\mathbf{X}$, $\mathbf{t}$, $\hat{f}$, $\mathbf{w}^{(l)}$, $\mathbf{h}$, $\mathbf{g}$, $\mathbf{y}$, $\mathbf{e}$, and L are respectively the training inputs, the targets, the ML model in training, the weights of the l-th layer, the activation function of the second layer, the one of the first layer, the predicted target, the error, and the loss function. This loop is repeated until the loss function is minimized for this given training dataset. Adapted from Ref. [83].

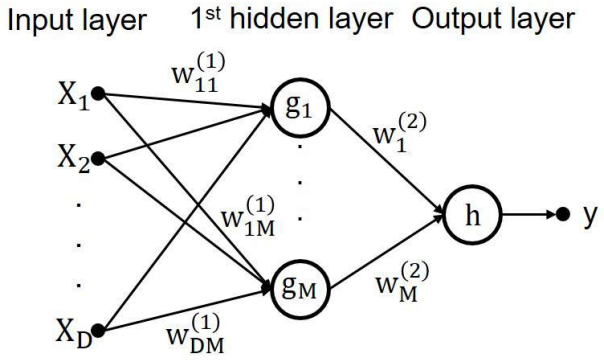


Figure 3.2: Two-layer artificial neural network with X_d the d-th feature, D the number of features, $w_{ij}^{(l)}$ the weight associated to the i-th node from the $(l-1)$-th layer and to the j-th node from the l-th layer, g and h the activation functions of the first hidden and output layer, and y the output. Adapted from Ref. [84].

Different methods to update the weights exist. The most basic one is the backpropagation of the error which makes use of the chain-rule of derivatives to derive the following equation

$$w_{ij}^{(l)}(t+1) \;=\; w_{ij}^{(l)}(t) - \alpha \frac{\partial L}{\partial w_{ij}^{(l)}(t)}, \tag{3.3}$$

where α is called the "learning rate" and t refers to the epoch, i.e., the number of passes (iterations) the algorithm has completed during training. The choice of optimization algorithm greatly influences the computational time as well as the final performance because some methods are more easily stuck in local minima for instance [83, 84].

Feature selection The inputs can be represented as a table where each row and each column corresponds to an instance of the DB (a specific material in the present work) and a feature (e.g., the mean band gap of the elements), respectively. One of the first step is thus to generate these descriptors. To that end, the Python library "Matminer" was developed and allows one to "featurize" materials based on their chemical compositions and/or structures, and optionally on their electronic bandstructure and/or density of states [85]. A large number

of features can thus be generated and are hopefully related in some way to the target property. In theory, the number of descriptors is directly proportional to the quality of the model since a new column can be seen as additional information. In practice, too many of them can actually be detrimental to a good performance on unseen data and this phenomenon is called the "curse of dimensionality." One of its aspects is explained hereafter. One can think of the instances of the training DB as coordinates in a high-dimensional space, i.e., the feature space. The goal of the ML is thus to approximate a hidden function by interpolating it based on this sampling of the feature space. However, increasing the number of features, and thus the dimensionality of the feature space, causes the training instances to be located "further away" from each other. The sampling of the space becomes sparse instead of uniform. As a result, the interpolation becomes less possible until it can be considered as extrapolation between more populated regions of the space. This implies that the hidden function is well approximated in these regions but is poorly estimated in the emptier ones. This is called "overfitting" which means that the ML model is not able to generalize its predictions to unseen samples. The model is actually too complex and possesses too many degrees of freedom. Fortunately, reducing the number of features by selecting the most relevant ones can prevent this effect. One of the basic feature selection is to keep the features that are the most correlated (in the sense of linear correlation) to the target. This method is not ideal as it does not allow to measure nonlinear relations. Other criteria can be developed to that end [86].

Model selection The next step is to select the best model for the given task. The term "best model" refers here to the model with the optimal set of hyperparamaters. The latter are parameters that are not optimized during the training as, for example, the number of hidden layers and nodes for each of them. The usual process to assess the quality of the models is to divide randomly the dataset into a training and a validation set. Once a configuration of the model is trained on the first, it is used to predict the targets of the second. The error on this validation set thus gives an estimation of the performance, for this type of model, on unseen data. In order to prevent a potential bias of the validation set, the usual procedure is to repeat this training/validation split k times so that, in the end, each instance of the whole dataset is part of only one validation set. The mean error of the k validation sets is the final score of the model. This process is called "k-fold cross-validation" (KCV) and is illustrated in Figure 3.3 [87, 88].

Performance evaluation The issue with the KCV method is that the score of the best model is too optimistic as a generalization score, i.e., as a measure of how well the final model will perform on unseen data. Indeed, the whole dataset played a role in the performance evaluation and in the choice of the best hyperparameters, thus in the training so to speak. This implies that the score is biased towards this DB which results in overly optimistic performance. In order to prevent this undesirable consequence, one can realize a "nested

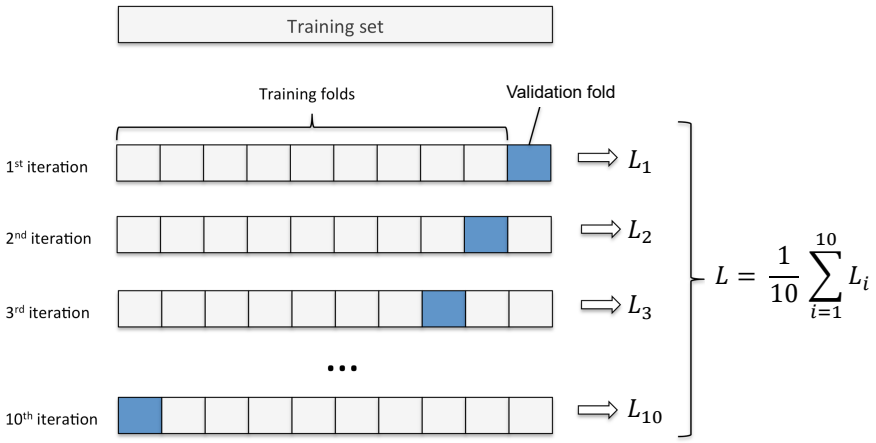


Figure 3.3: Illustration of the procedure of a 10-fold cross-validation with a loss function, L. Adapted from Ref. [89].

cross-validation." This procedure consists in two KCV loops. The outer one divides the DB in a test and a training set T_o while the inner loop is the KCV as explained above but applied only on T_o. Once the best hyperparameters are selected for T_o, the generalization accuracy of the resulting model is measured on the corresponding test set of the outer loop. This means that the final performance estimation is an average of scores obtained with potentially different models since each training set of the outer loop might result in a different choice of hyperparameters. The final score now evaluates the strategy (i.e., the inner KCV optimization) used to obtain the best model and the optimism of the KCV is thus eliminated. Once found, the best strategy can be re-applied on the full dataset in order to make use of all the available instances [90, 91].

3.1.2 MODNet

Considering the limited available data in natural sciences, the choice of the ML model for the intended application is important as it impacts the resulting accuracy. The present work makes use of the material optimal descriptor network (MODNet) model which currently leads 4 out of the 13 tasks of Matbench [92, 93]. This test suite ensures a fair comparison of ML models based on their performance on various tasks. MODNet consists of a feedforward neural network based on two concepts: feature selection and joint-learning. The latter is omitted since it is not used in the present work. The architecture of MODNet is illustrated in Figure 3.4.

Feature selection MODNet is designed to reduce the high-dimensional (HD) feature space to a subset, $\mathcal{F}_s$, of the most physically relevant descriptors ("Feature selection" step in Figure 3.4). These are chosen based on a relevance-redundancy criterion written as follows in the case of a feature f

$$\mathrm{RR(f)} = \frac{\mathrm{NMI(f,y)}}{[\max_{\mathrm{f_s}\in\mathcal{F}_s}(\mathrm{NMI\,(f,f_s)})]^{\mathrm{p}} + \mathrm{c}} \quad \text{with} \quad \mathrm{NMI(X,Y)} = \frac{\mathrm{MI(X,Y)}}{(\mathrm{H(X)} + \mathrm{H(Y)})/2}, \tag{3.4}$$

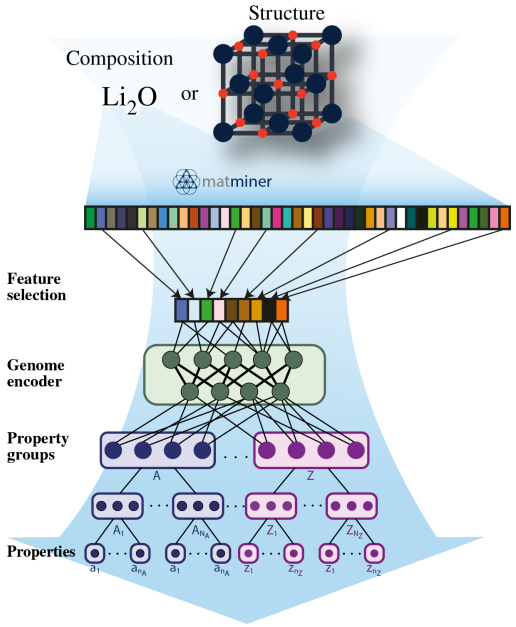


Figure 3.4: Schematic of the MODNet model. The feature selection on Matminer is followed by a hierarchical tree-like neural network. Various properties $A_1, \ldots, A_{N_A}, \ldots, Z_1, \ldots, Z_{N_Z}$ are gathered in groups from A to Z of similar nature. Each of these may depend on a parameter: A(a), ..., Z(z). The properties are available for various values of the parameters $a_1, \ldots, a_{n_A}, \ldots, z_1, \ldots, z_{n_Z}$. The first green block of the neural network encodes a material in an appropriate all-round vector, while subsequent blocks decode and re-encode this representation in a more target-specific nature. Taken from Ref. [92].

where RR is the relevance-redundancy score, NMI is the normalized mutual information, y the target, MI an estimator of the mutual information (see Ref. [94]) and H the information entropy, i.e., H(X)=MI(X,X). X and Y are two random variables and the mutual information is defined as

$$\mathrm{I(X,Y)} = \iint dxdy\, \mu(x,y) \log \frac{\mu(x,y)}{\mu_x(x)\, \mu_y(y)}, \tag{3.5}$$

where $\mu(x)$ is the marginal density of X. The mutual information allows to evaluate the relation of a random variable with another one by measuring how the uncertainty on the first one is reduced when the second one is known (and inversely). In this sense, it goes beyond the usual linear correlation coefficient which is unable to detect nonlinear dependence. In Eq. (3.4), p and c are two hyperparameters that can be varied to favor relevance over redundancy or vice versa.

The first selected feature has the largest NMI with the target. Afterwards, each added descriptor is the one with the highest RR. The best ranked descriptors are chosen in a quantity that is either fixed arbitrarily or in such a way that the error is minimized. Reducing the feature space helps reducing the "curse of dimensionality" and thus the potential danger of overfitting. This is particularly important when dealing with small datasets. In addition to an increased performance, this reduction to optimal descriptors and its associated global

ranking give the researcher an additional way of inspecting the physics and can guide him into further developing the theory.

Ensemble MODNet model MODNet offers the possibility to realize bagging. This concept is a simple and effective way to improve the predictive power of ML models as well as furnish an estimation of the corresponding error. It consists in bootstrapping the DB, i.e., choosing multiple instances with replacement in order to create several artificial datasets, and training one MODNet model per new set. Together, they all form an "Ensemble model." The prediction on an unseen instance is then the mean of the predictions of the individual models and the uncertainty can be defined as their standard deviation [95, 96, 97].

Hyperparameters optimization Usually, the hyperparameter optimization is realized with a gridsearch approach. This means that multiple combinations of values of hyperparameters are given in advance by the user. Each possibility is then tested by the program and assigned a score, via KCV for example. In the MODNet package, several well-working combinations are already implemented as "presets." One can then realize a gridsearch on all these candidates. However, this procedure is highly time-consuming and does not guarantee that the optimal configuration is among these presets. The newest implementation of hyperparameters optimization in MODNet is a genetic algorithm (GA) [98, 99]. It is a type of random search of the optimization space. Figure 3.5 illustrates the general principle in the case of four binary hyperparameters. The first step is to randomly draw an initial population of individual candidates from the hyperparameter space. Then, the best individuals should be selected based on an objective function (typically via KCV on the training set here). Once the population has been thinned, random (probabilistic) modifications can happen to the remaining individuals. For example, Figure 3.5 shows a cross-over where two individuals exchange two of their "genes", i.e., their values, and a mutation where genes simply switch values. At the end of these random processes, a new "generation" is obtained and the procedure can be repeated by adding new random samples from the whole space. The GA stops either when it has converged to the best population or when a certain generation is reached. MODNet can then create an Ensemble model of tunable size S where the choice of the individual models is either the optimal combination of the final generation individuals or the best S ranked.

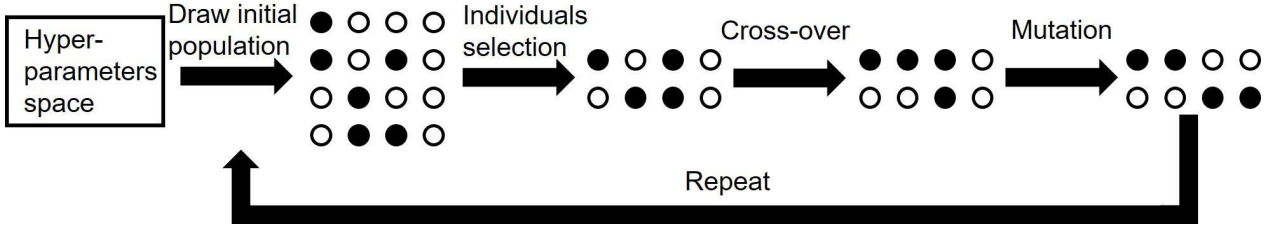


Figure 3.5: Schematic representation of the general principle of a genetic algorithm. See text for more details. Adapted from Ref. [99].

3.2 Choice of target property

3.2.1 Inadequacy of the Kurtz-Perry effective coefficient

Thanks to the nonlinear DB, it is possible to develop a ML model that aims at predicting the SHG response. At first, the Kurtz-Perry effective nonlinear coefficient of Eq. (1.33) was chosen as target property since it presents the advantage of being a scalar quantity as well as a direct link to experimental works. Surprisingly, it was discovered that $d_{\text{eff,KP}}$ was not an invariant of the SHG tensor, which naturally depends on the orientation of the system relative to the choice of coordinates. Indeed, it was initially assumed that Eq. (1.33) was free from such dependence since it does not make sense to characterize a material in a global manner based on a property that depends on the orientation of the system. Moreover, intuition dictates that a rotational average should not present such dependence.

In a 3D-space, for the case of a rotation of α, followed by one of β, and then one of γ around the x-, y-, and z-axis respectively, the matrix of rotation writes [100]

$$\begin{aligned} R &= R_z(\gamma)R_y(\beta)R_x(\alpha) \\ &= \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix} \\ &= \begin{bmatrix} \cos\beta\cos\gamma & \sin\alpha\sin\beta\cos\gamma - \cos\alpha\sin\gamma & \cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma \\ \cos\beta\sin\gamma & \sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma & \cos\alpha\sin\beta\sin\gamma - \sin\alpha\cos\gamma \\ -\sin\beta & \sin\alpha\cos\beta & \cos\alpha\cos\beta \end{bmatrix}, \end{aligned}$$

where a positive angle corresponds to an anti-clockwise rotation when the axis of rotation points towards the observer. The three angles above are called Euler angles. This rotation matrix is more precisely a transformation one, $\mathbf{Q}$. It allows to express tensors in a new coordinate system following a rigid rotation of the first one. Its components come directly from the relation between the base vectors of the two different systems:

$$\mathbf{e}_i \cdot \mathbf{e'}_j = Q_{ij} \Leftrightarrow \mathbf{e}_i = Q_{ij}\mathbf{e'}_j, \tag{3.6}$$

where $\mathbf{e}_i$ and $\mathbf{e'}_j$ refer to the i-th base vector of the first system and the j-th of the second one. The subscripts indicate different vectors instead of different components of vectors. The above equation leads to the tensor transformation rule:

$$T_{ijk...n} = Q_{ip}Q_{jq}Q_{kr}...Q_{nm}T'_{pqr...m} \Leftrightarrow \mathbf{T} = \mathbf{Q}\mathbf{T'}\mathbf{Q}^T, \tag{3.7}$$

where T is a tensor of arbitrarily high order and the subscripts now refer to the components of the tensors [101]. The matrix form of the equation is also present on the far right. Eq. (3.7) was applied on the d_{ijk} tensor of $MgSiP_2$ for a set of Euler angles before calculating

the corresponding KP effective nonlinear coefficient. The latter is shown in Figure 3.6 in the form of a map with respect to the Euler angles. A periodicity of 90° was found which justifies the maximum value of α in Figure 3.6. First-principles computations were also realized as a verification of the tensor transformation rule. Although not investigated further, this dependence might be an error in the analytic form of Eq. (1.33). This matter ought to be verified in a future work by computing the rotational average in an alternative way such as the one introduced in Ref. [34].

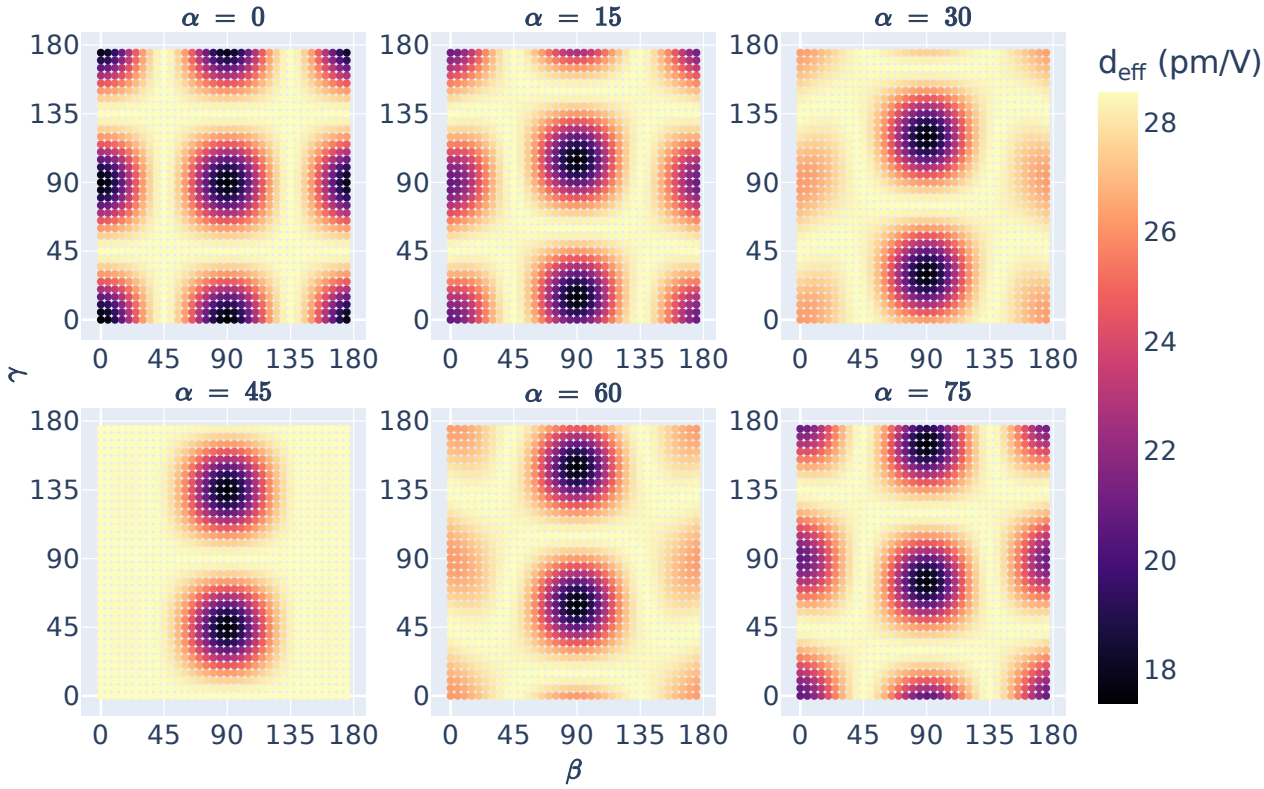


Figure 3.6: Map of the KP effective nonlinear coefficient of $MgSiP_2$ for a set of Euler angles.

The target property should be independent of the orientation of the system. Indeed, the materials features generated by Matminer are related to the compound itself and not to its orientation in an arbitrarily chosen coordinates system. This implies that the performance of the model can greatly suffer from such dependence. As mentioned, it also does not make sense to rank potential NLO crystals based on such variable. For these reasons, the idea of using KP effective nonlinear coefficient had to be abandoned. It was thus necessary to define a new scalar quantity to characterize, in a first approach, the magnitude of the SHG response of crystals.

3.2.2 Invariants of tensors

Since the invariance with respect to the orientation of the system is an important aspect, the obvious choice is to consider invariants (under rotation at least) of the tensor, i.e., scalars derived from its components and that do not change when the frame is rotated. The case of second-order tensors is well known and is reminded hereafter for the sake of extending the discussion to third-order tensors afterwards [102, 103]. By expressing the matrix representation of a second-order tensor in its principal frame, its off-diagonal components are set to 0 and its diagonal ones are called eigenvalues. The latter can be considered as invariants of the tensor. They can be computed by solving its characteristic polynomial

$$\det(\lambda \mathbf{I} - \mathbf{D}) \;=\; \lambda^3 - J_1\lambda^2 + J_2\lambda - J_3 \;=\; 0, \tag{3.8}$$

where λ is the variable of the polynomial, $\mathbf{I}$ the identity matrix, and $\mathbf{D}$ a generic second-order tensor in its matrix form. The J_i are called the principal invariants of the tensor and can be expressed from the components of $\mathbf{D}$

$$J_1 \;=\; \mathrm{tr}(\mathbf{D}) \;=\; \lambda_1 + \lambda_2 + \lambda_3 \tag{3.9}$$

$$J_2 \;=\; \frac{\mathrm{tr}(\mathbf{D})^2 - \mathrm{tr}\left(\mathbf{D}^2\right)}{2} \;=\; \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3 \tag{3.10}$$

$$J_3 \;=\; \det(\mathbf{D}) \;=\; \lambda_1\lambda_2\lambda_3 \tag{3.11}$$

$$J_4 \;=\; \mathrm{tr}\left(\mathbf{D}^{\mathrm{t}}\mathbf{D}\right) \;=\; J_1^2 - 2J_2 \tag{3.12}$$

$$\;=\; D_{11}^2 + 2D_{12}^2 + 2D_{13}^2 + D_{22}^2 + 2D_{23}^2 + D_{33}^2 \;=\; \lambda_1^2 + \lambda_2^2 + \lambda_3^2,$$

where J_4 is not independent from the others. It is written because it has proven useful in some fields. These quantities have indeed a particular significance and various applications in many domains of physics and engineering such as expressing the spherical and deviatoric part of tensors in mechanics or the von Mises criterion for the occurrence of plastic yielding [104]. It can thus be interesting to look into their analog in the case of third-order tensors and to adopt one of them as the target property for the ML model. Unfortunately, the case of high-order, i.e., higher than second, tensors is more complex. It has been shown that, in a three-dimensional space, the maximum number of principal invariants amounts to 3 for the second-order while it is lower or equal to $\mathrm{m}^2 + m$ for the m-th-order [103]. This implies that the d_{ijk} may present 12 principal invariants. Since the theory can thus become relatively heavy, some rigor has been left aside in what follows and the interested reader can find more information in Refs. [105, 106].

Based on the case of second-order tensors, it makes sense to search for invariants of third-order ones by considering their eigenvalues. However, the situation is more complex than

previously since multiple definitions exist such as the C-, E-, Z-, and L-eigenvalues. The latter being defined as the real scalar value $\sigma \geq 0$ in

$$\mathcal{A}\mathbf{V} = \sigma\mathbf{x}, \quad \mathcal{A}^{\top}\mathbf{x} = \sigma\mathbf{V}, \quad \mathbf{V} \bullet \mathbf{V} = 1, \quad \mathbf{x} \bullet \mathbf{x} = 1, \tag{3.13}$$

where $\mathcal{A}$, $\mathbf{V}$, and $\mathbf{x}$ correspond to generic third-, second-, and first-order tensors in their three dimensional space (or Hilbert space) while $\bullet$ refers to its inner product. This definition can be expressed in $\mathfrak{R}^3$ under an orthonormal basis:

$$a_{ijk}v_{jk} = \sigma x_i, \quad a_{jki}x_j = \sigma v_{ki}, \quad v_{ij}v_{ij} = 1, \quad x_i x_i = 1, \tag{3.14}$$

where Einstein summation convention is used. It can be demonstrated that any third-order tensor possesses three L-eigenvalues, σ_i, with their corresponding L-eigentensors, $\mathbf{V}_i$, and L-eigenvectors, $\mathbf{x}_i$. As can be understood, L-eigenvalues constitute invariants of $\mathcal{A}$ although their calculation can be tedious. Fortunately, the advantage of this type of eigenvalues is their direct link to the eigenvalues, $\lambda_j = \sigma_j^2$, of the kernel tensor, $\mathbf{U}$. The latter is a symmetric positive semi-definite second-order tensor defined as $\mathbf{U} = \mathcal{A}\mathcal{A}^T$ or $u_{il} = a_{ijk}a_{jkl}$. The transpose of a generic third order tensor $\mathcal{A}$ is unique and is defined by $[(\mathcal{A}^T)^T]^T = \mathcal{A}$ or $a^T_{ijk} = a_{jki}$. This implies that the kernel tensor is also uniquely defined. Therefore, its invariants such as its trace and determinant are invariants of $\mathcal{A}$ as well. Moreover, two additional kernel tensors can be defined, namely $\bar{\mathbf{U}}$ for $\mathcal{A}^T$ and $\hat{\mathbf{U}}$ for $(\mathcal{A}^T)^T$. In total, 7 invariants can thus be found based on the L-eigenvalues of a third-order tensor:

$$\mathrm{tr}(\mathbf{U}), \quad \mathrm{tr}(\mathbf{U}^2), \quad \mathrm{tr}(\mathbf{U}^3), \quad \mathrm{tr}(\bar{\mathbf{U}}^2), \quad \mathrm{tr}(\bar{\mathbf{U}}^3), \quad \mathrm{tr}(\hat{\mathbf{U}}^2), \quad \mathrm{tr}(\hat{\mathbf{U}}^3) \tag{3.15}$$

since $\mathrm{tr}(\mathbf{U}) = \mathrm{tr}(\bar{\mathbf{U}}) = \mathrm{tr}(\hat{\mathbf{U}})$. The three invariants of $\mathbf{U}$ can be re-expressed to find the principal invariants introduced in the previous paragraphs. If one assumes the Kleinman's symmetry condition to be valid such as in the case of the static SHG tensor, then the cyclic symmetry implies that $\mathcal{A} = \mathcal{A}^T$. This reduces the number of independent invariants from 7 to 3:

$$\mathrm{tr}(\mathbf{U}) = \lambda_1 + \lambda_2 + \lambda_3 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2 = a_{ijk}a_{ijk} \tag{3.16}$$

$$\mathrm{tr}(\mathbf{U}^2) = \lambda_1^2 + \lambda_2^2 + \lambda_3^2 = \sigma_1^4 + \sigma_2^4 + \sigma_3^4 = a_{ijk}a_{jkl}a_{klm}a_{lmi} \tag{3.17}$$

$$\mathrm{tr}(\mathbf{U}^3) = \lambda_1^3 + \lambda_2^3 + \lambda_3^3 = \sigma_1^6 + \sigma_2^6 + \sigma_3^6 = a_{ijk}a_{jkl}a_{klm}a_{lmn}a_{mno}a_{noi}. \tag{3.18}$$

Eq. (3.16) has been arbitrarily chosen as the target property of the ML model. It has nonetheless been modified to recover the same units as the SHG coefficient, i.e., pm/V. It was also divided by 27. The target property is thus the root mean square (RMS) of the d_{ijk} components, which we will call the RMS (SHG) coefficient or d_{RMS}. The two other invariants as well as the ones from other types of eigenvalue have not been investigated and could be the focus of a future work. It can indeed be shown that seven independent invariants can be

found in the case of a symmetric third-order tensor such as the SHG tensor. As mentioned, the latter adopts the same form than the piezoelectric one. This implies that, similarly to its piezoelectric analog, the SHG tensor is a so-called right-side symmetric tensor, i.e., $d_{ijk} = d_{ikj}$ for all i, j, and k. Combined with its cyclical symmetry, this makes d_{ijk} a symmetric tensor with seven independent invariants that could prove useful in the NLO domain and crystals characterization.

3.3 Dataset analysis

The RMS coefficient presents the additional advantage of exhibiting values in almost the same range as the KP effective coefficient as can be seen from Figure 3.7a. Since it does not make sense anymore to work with $d_{\text{eff,KP}}$, the latter can be averaged over a large set of Euler angles. This results in $\bar{d}_{\text{eff,KP}}$ whose relationships with the RMS and KP coefficient are illustrated in Figure 3.7b and Figure 3.7c, respectively. These three SHG effective

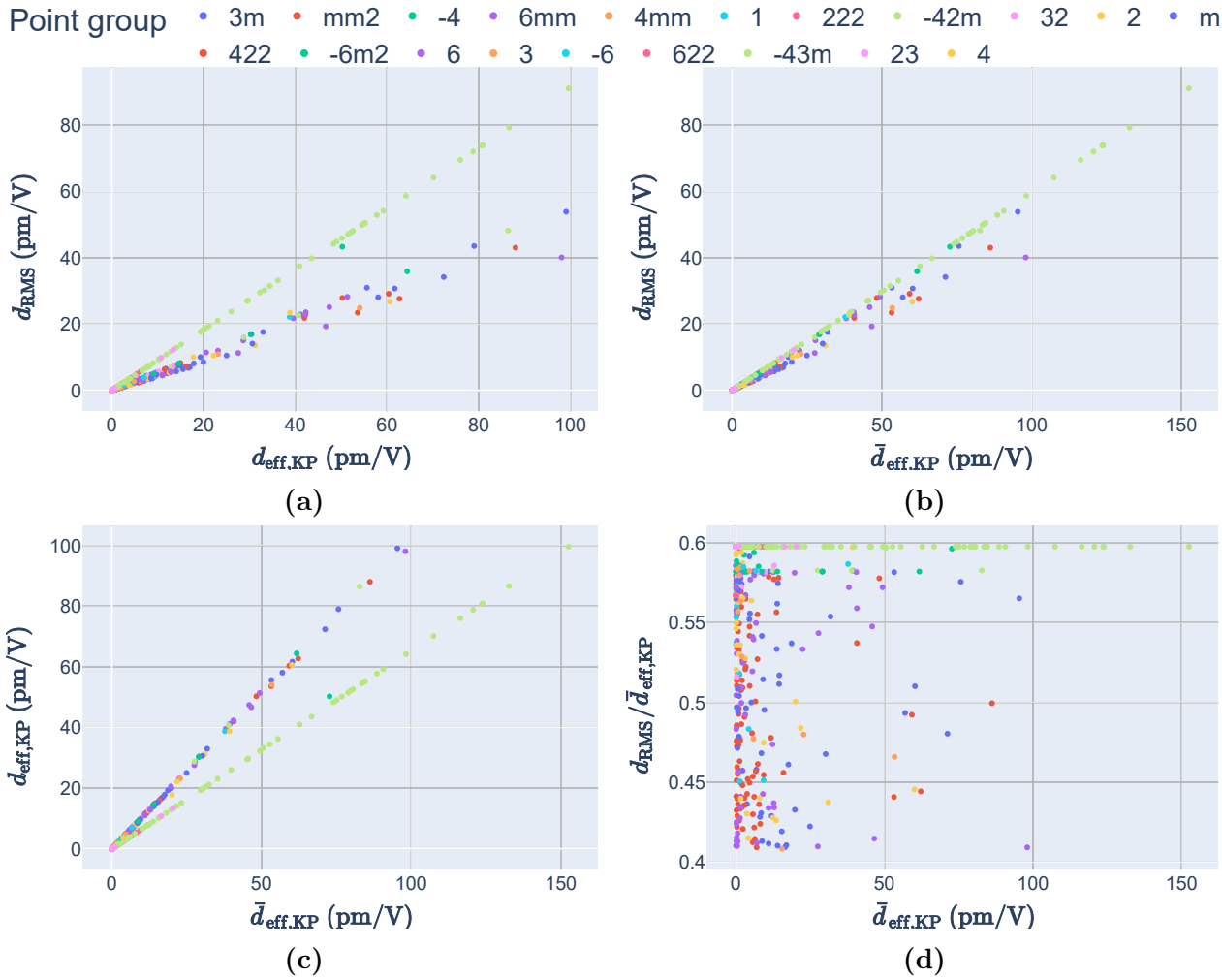


Figure 3.7: Different graphs illustrating the strong linear correlations between d_{RMS}, $d_{\text{eff,KP}}$, and $\bar{d}_{\text{eff,KP}}$. The dimensionless ratio $d_{\text{RMS}}/\bar{d}_{\text{eff,KP}}$ is also represented in (d).

coefficients are highly linearly correlated as shown in Figure 3.7. This implies that the

performance of the ML model should be independent of the choice of the target property. However, the RMS coefficient is mathematically more robust than $\bar{d}_{\text{eff,KP}}$. The former is thus better suited as target property. At first, it was thought that the point group of the material impacted the slope relating the coefficients. However, it is not the case based on Figure 3.7. Figure 3.7d particularly shows that, except for the -43m group, the RMS coefficient is not linearly dependent on the average KP one.

Unfortunately, as the histogram on the left of Figure 3.8 shows, the DB at hand is strongly biased towards low target. A skewed dataset, whether in its features or in its target, can strongly impair the final performance. It can indeed lead to overfitting such that the lower end of the RMS coefficients is well predicted at the expense of their higher values. In principle, the MODNet model should be flexible enough to handle this kind of bias. Figure 3.8b also shows how unbalanced the DB is with respect to the point groups. If the latter are part of the optimal descriptors chosen by the model, then poor performance can be expected when predicting an instance of the minority point groups.

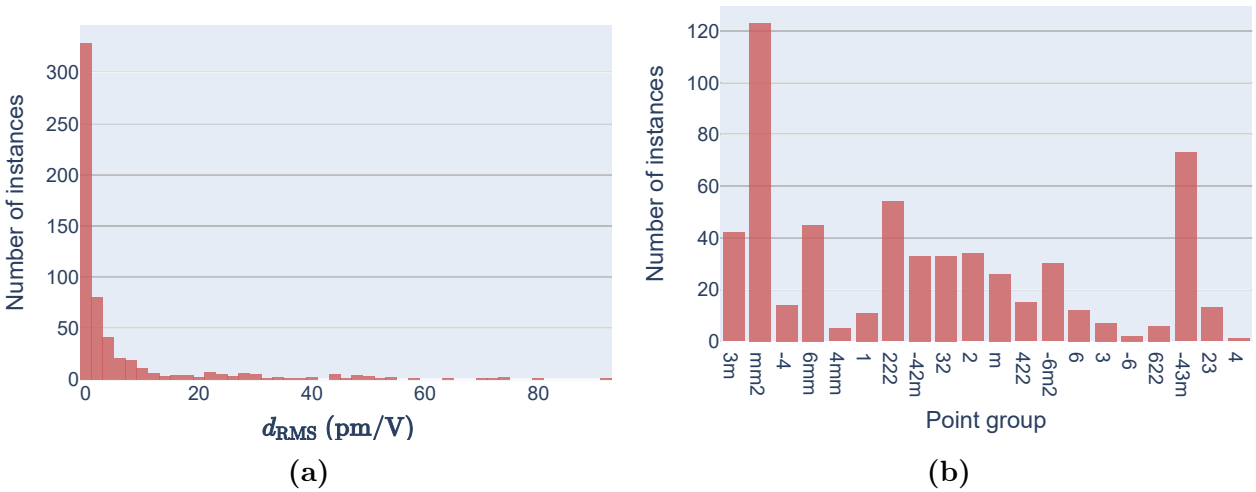


Figure 3.8: Histograms to illustrate (a) the strong bias towards the low values of d_{RMS} and (b) the unbalance of point groups in the DB.

Figure 3.9c shows the inverse proportionality relating E_g and n_s while Figure 3.9a and Figure 3.9b illustrate Miller's rule which can be approximated in the same way as in Eq. (2.6) with the new SHG effective coefficient. The trend "high SHG, high refractive index, low bandgap" can indeed be clearly identified. Their relation can be measured thanks to the normalized mutual information as introduced in Eq. (3.4) such that

$$\begin{aligned} \text{NMI}(d_{\text{RMS}},n_s) &= 0.09758 \\ \text{NMI}(d_{\text{RMS}},E_g) &= 0.08867 \\ \text{NMI}(n_s,E_g) &= 0.13896 \end{aligned}$$

It follows that the refractive index is the first ranked optimal descriptor as it could be

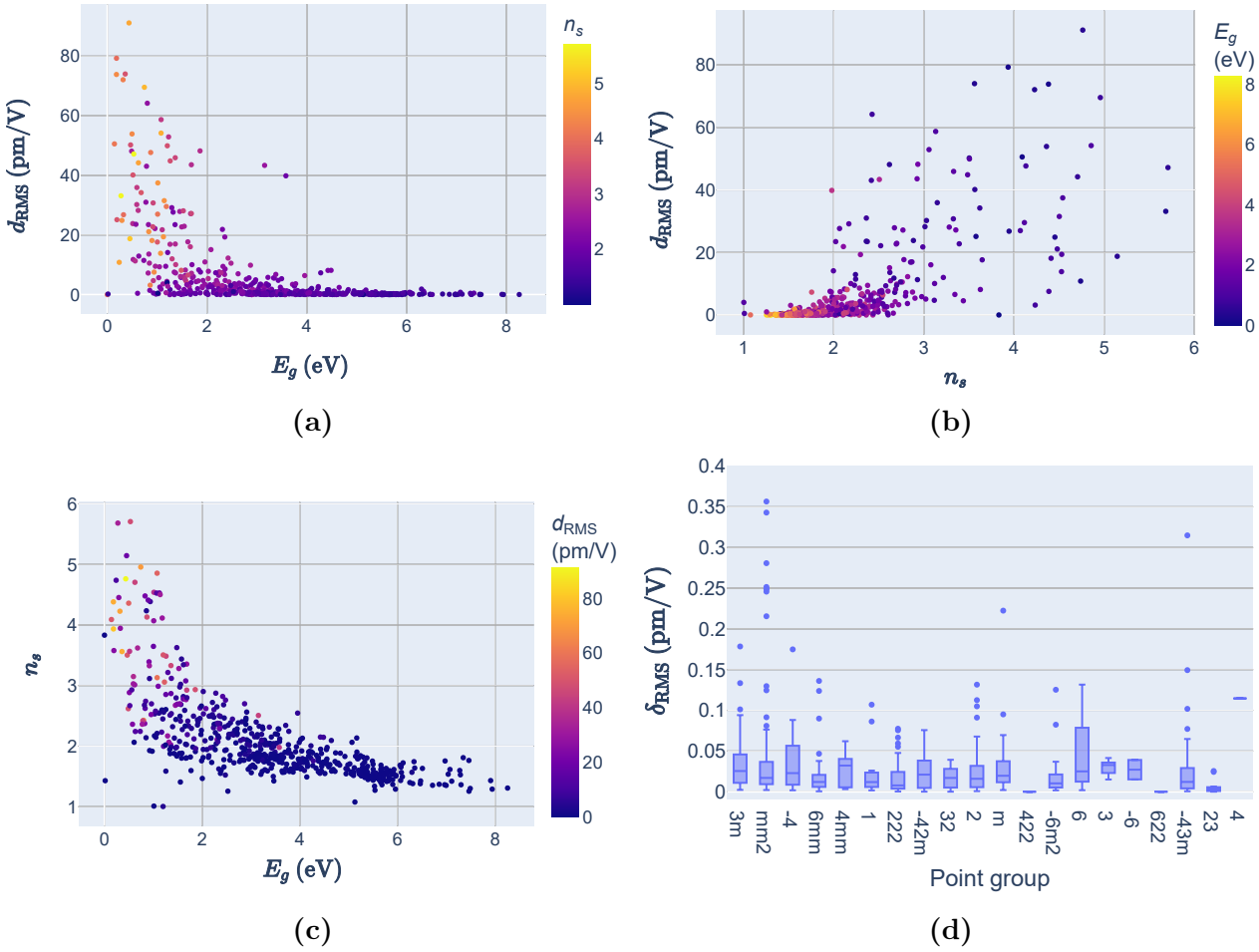


Figure 3.9: Scatter plots illustrating the relationship between the target d_{RMS}, n_s (dimensionless), and the bandgap, E_g. (d) presents box plots of the Miller coefficient from the RMS coefficient as a function of the point groups. A few outliers were cast aside for practical reasons.

predicted from Miller's rule. The new approximated Miller coefficient is displayed in Figure 3.9d with respect to the point group of the crystal. It is difficult to state any hypothesis regarding a possible link between these two parameters since the DB is highly unbalanced as can be seen in Figure 3.8b.

3.4 Features

As mentioned, the features are generated using the Python library "Matminer" based on the chemical composition and structure of the compounds. However, the ones with the lowest variance when considering the entirety of the MP were discarded. More precisely, if 70% of the MP materials adopt the same value for a given feature, then the latter is removed. This can help the ML model in practice. In the end, 1359 features from Matminer were selected. In order to investigate the effect of using the refractive index or bandgap as descriptors, four copies of the DB were created. They are comprised of the Matminer features without n_s and

E_g, with n_s, with E_g, and with both, respectively. The first five optimal variables of each DB are listed in Table 3.1, Table 3.2, Table 3.3, and Table 3.4. These illustrate how features with low target NMI can still be the best ones based on the RR score ranking. It can also be seen that including, or not, n_s and/or E_g affects significantly the ranking. When included, they are always the first ones. Moreover, in Table 3.4, the bandgap is not part of the first five descriptors due to its high redundancy with the refractive index. It is interesting to observe that the space group is an important parameter based on this ranking since it is the third one in Table 3.2, Table 3.3, and Table 3.4. However, it only appears when either n_s or E_g or both are included. It must be noted that the process of calculating the target NMI is partially stochastic which explains the slight difference of value for the same feature in two different datasets.

Optimal descriptors	Target NMI
AtomicOrbitals|HOMO_energy	0.079142
XRDPowderPattern|xrd_51	0.008932
XRDPowderPattern|xrd_45	0.022551
OPSiteFingerprint|std_dev hexagonal pyramidal CN_7	0.033915
OPSiteFingerprint|std_dev q2 CN_11	0.037828

Table 3.1: First five optimal features based on their RR score and their NMI with the target for the first dataset, i.e., without the static refractive index and the bandgap.

Optimal descriptors	Target NMI
Bandgap	0.088522
XRDPowderPattern|xrd_45	0.022534
GlobalSymmetryFeatures|spacegroup_num	0.053508
ElementProperty|MagpieData range GSvolume_pa	0.045232
RadialDistributionFunction|radial distribution function|d_3.70	0.034733

Table 3.2: First five optimal features based on their RR score and their NMI with the target for the second dataset, i.e., without the static refractive index and with the bandgap.

Optimal descriptors	Target NMI
Static refractive index	0.097606
ElementProperty|MagpieData minimum NsValence	0.004828
GlobalSymmetryFeatures|spacegroup_num	0.054504
RadialDistributionFunction|radial distribution function|d_2.70	0.028227
ElementProperty|MagpieData mean GSvolume_pa	0.036434

Table 3.3: First five optimal features based on their RR score and their NMI with the target for the third dataset, i.e., with the static refractive index and without the bandgap.

Optimal descriptors	Target NMI
Static refractive index	0.097592
ElementProperty|MagpieData minimum NsValence	0.005053
GlobalSymmetryFeatures|spacegroup_num	0.054336
GeneralizedRDF|mean Gaussian center=2.0 width=1.0	0.027453
OPSiteFingerprint|std_dev hexagonal pyramidal CN_7	0.032494

Table 3.4: First five optimal features based on their RR score and their NMI with the target for the fourth dataset, i.e., with the static refractive index and the bandgap.

3.5 Methods

Nested cross-validation (5-fold inner and 5-fold outer) with inner feature selection as implemented in the *benchmark* Python script of the *modnet* package was adopted to obtain the performance score of each model. Hyperparameters optimization was realized both with a gridsearch on presets and with the genetic algorithm. While the former was run for both the Vanilla MODNet model, i.e., the usual one, and its Ensemble analog which consists in bagging 100 Vanilla models, the latter was performed three times such that the final model of the inner GA was:

- a 50-individuals Ensemble model for which an optimal bagging was done among the final population (*refit=0*)
- a 50-individuals Ensemble model built from the 50 best individuals in the final population (*refit=50*)
- a Vanilla model corresponding to the best individual in the final population (*refit=1*)

In the rest of the present work, the five approaches will be called Vanilla, Ensemble, GA0, GA50, and GA1. Unless stated otherwise, the default parameters were systematically used.

In addition to these five different approaches to find the best ML model, the effect of adding various proportions of centrosymmetric materials to the DB was investigated. The motivation came from the idea that the SHG response should be linked to the "degree of centrosymmetry" of the crystal as mentioned previously. In practice, the goal was to better sample the feature space in order to prevent overfitting and the curse of dimensionality in general. The following quantities of centrosymmetric instances were randomly taken from the linear DB: 144, 289, 434, 579, 723, 868, i.e., 25%, 50%, 75%, 100%, 125%, 150% of the initial nonlinear dataset. Similarly to the previous section, the first five optimal descriptors and their NMI with the target were listed. The case of the fourth dataset (with n_s and E_g) with 25% of centrosymmetric instances is presented in Table 3.5. It illustrates how the boolean centrosymmetry feature is ranked first with a target NMI higher than the one of the static refractive index. An identical observation was made regarding the three other datasets. The lowest proportion of the new materials is thus enough to profoundly change the choice of

variables. Increasing the quantity of centrosymmetric compounds further consolidates the first place of the centrosymmetry feature by augmenting its target NMI. It also modifies the ranking of the other descriptors but no clear relation could be identified.

Optimal descriptors	Target NMI
GlobalSymmetryFeatures|is_centrosymmetric	0.165672
AtomicOrbitals|gap_AO	0.020896
Static refractive index	0.068343
VoronoiFingerprint|std_dev Voro_index_7	0.041156
OPSiteFingerprint|std_dev tetrahedral CN_4	0.041586

Table 3.5: First five optimal features based on their RR score and their NMI with the target for the fourth dataset, i.e., with the static refractive index and the bandgap, and 25% of centrosymmetric materials with respect to the initial nonlinear dataset.

3.6 Results

The effect of adding n_s and E_g on the performance of each model tested is represented in Figure 3.10. The chosen metric to assess the performance is the coefficient of determination R^2:

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{3.19}$$

with y_i the target values, $\hat{y}_i$ their predictions, and $\bar{y}$ their mean. It portrays the proportion of the target variance that is explained by the independent variables [107]. Although it allows to compare the performance of multiple models, it presents the drawback of lacking a physical interpretation such as the one of the mean absolute error (MAE) or the RMS error (RMSE). The impact of including n_s and E_g depends on the model considered. For each

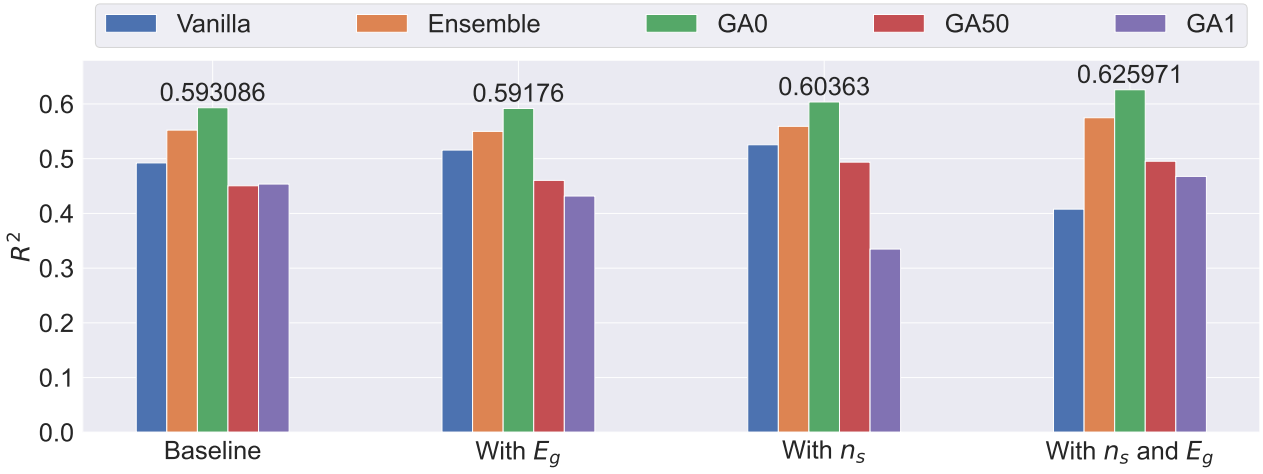
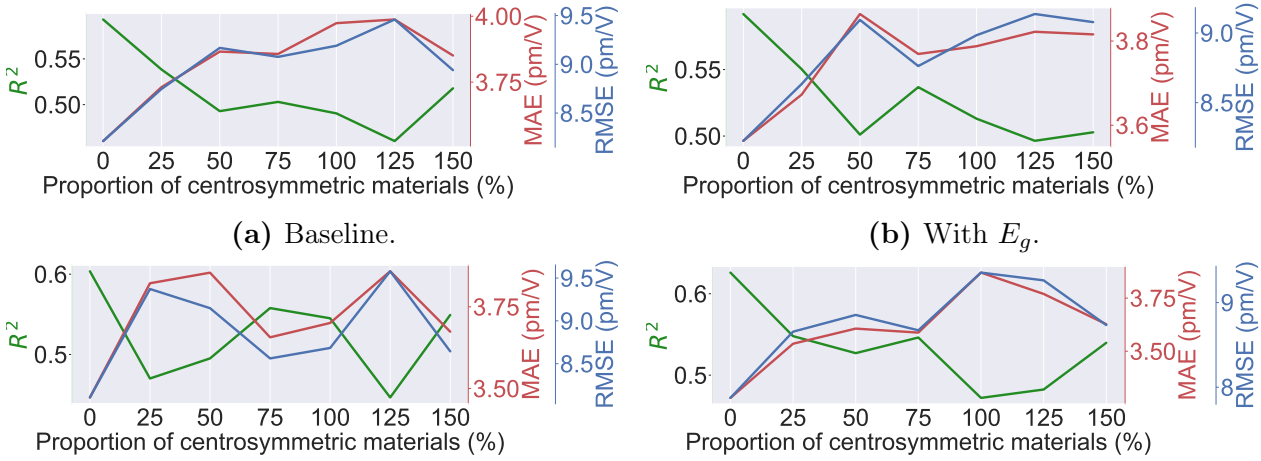


Figure 3.10: Impact of including the static refractive index and/or the bandgap on the R^2 score (dimensionless) of each model.

dataset, the GA0 approach displays the best score. For this reason, the other methods will

not be shown in the rest of this work. However, the disadvantage of evolutionary algorithm such as the GA is their inherent randomness. It implies that the reproducibility of the results is not guaranteed. For example, from Figure 3.10, including the bandgap seems to lower the performance of the GA0 model. This could nonetheless be an artifact of the method since both scores are close to each other. In any case, the separate addition of n_s or E_g does not produce a substantial improvement contrarily to their combined one. It must be noted that a R^2 score of 0.62 is relatively low for the intended application.

Contrarily to what was hoped, the progressive inclusion of centrosymmetric compounds is shown to worsen the results of all the models. Figure 3.11 illustrates this decreasing trend of the performance in the case of the GA0 model. It is important to note that



(a) Baseline. **(b)** With E_g. **(c)** With n_s. **(d)** With n_s and E_g.

Figure 3.11: Evolution of the generalization score of the GA0 model for each dataset as the proportion of centrosymmetric materials with respect to the initial nonlinear DB is increased. The performance metric are the coefficient of determination, the mean absolute error and the root mean squared error. Only the noncentrosymmetric instances are taken into account to evaluate the model.

only the error of the initial noncentrosymmetric instances is taken into account. Three hypothesis can be made: either the selected Matminer features are inadequate to represent the concept of "degree of centrosymmetry" or the latter does not matter or is not enough to compensate the worsening of the feature space sampling (increase of the target property bias) due to the inclusion of zero SHG response instances. It might thus be a reasonable perspective to turn to more structure-oriented ML models such as MEGNet for example [108].

When considering the MAE or RMSE shown in Figure 3.11, it can be thought that the error on the target is relatively small. However, as depicted in Figure 3.12a and Figure 3.12b, this only mirrors the predictions of the low target values rather than the reality of the overall predictions. Indeed, from Figure 3.12c, it can be seen that the accuracy of the best model, i.e., GA0 with n_s and E_g and without any centrosymmetric instances, is lower for larger values

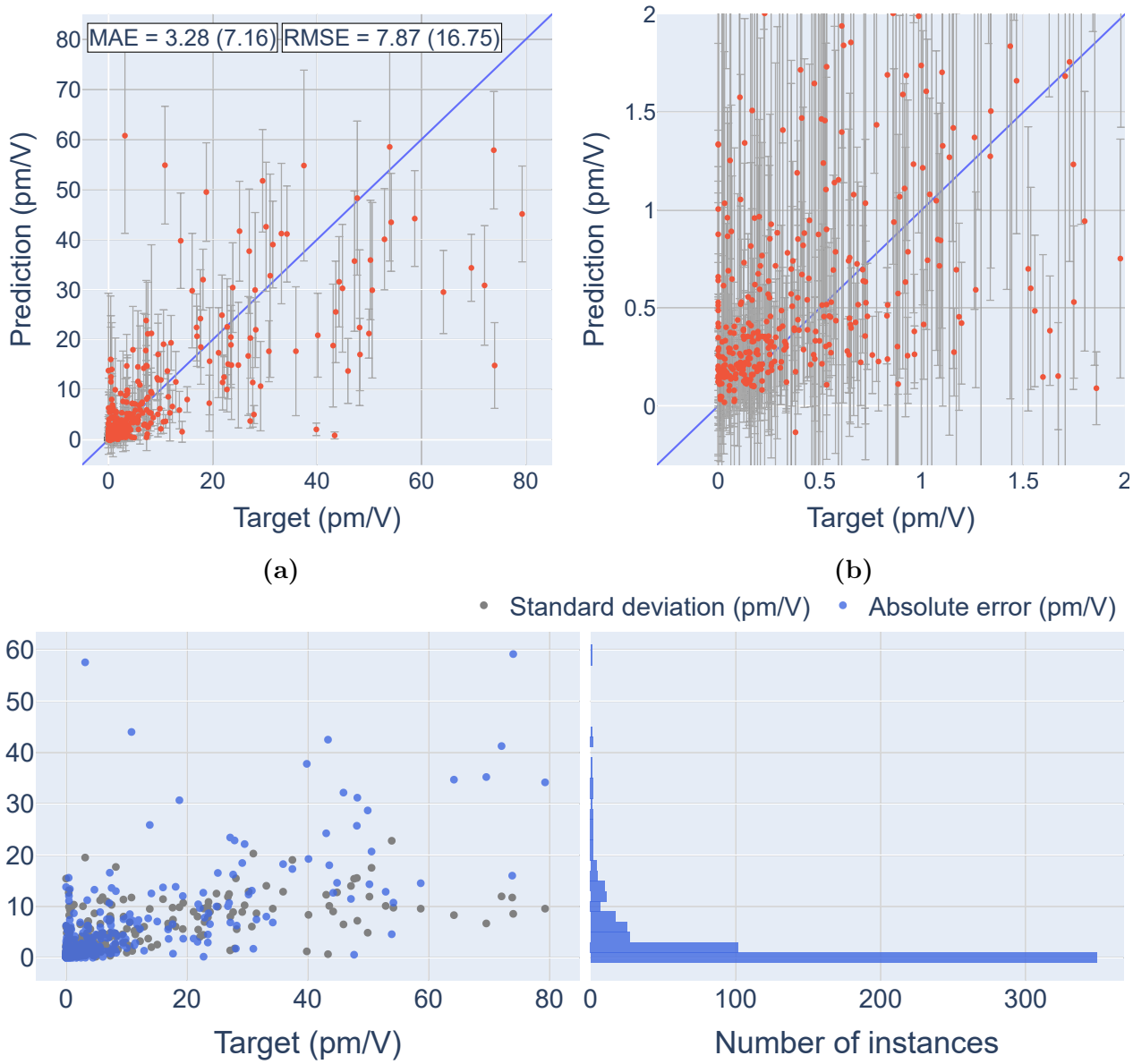
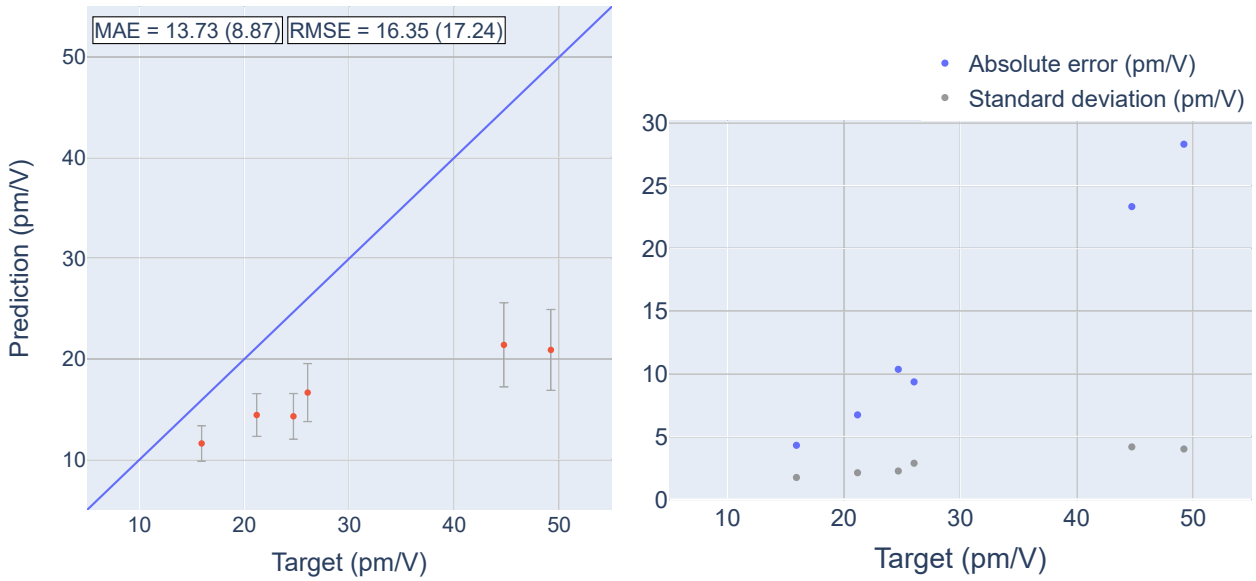


Figure 3.12: (a) Predictions of the test set RMS coefficient (red dots) with their related error bars (in grey) in the case of the GA0 model without any centrosymmetric compounds and including n_s and E_g. The error bars correspond to the standard deviation of the individual predictions of the GA0 (Ensemble MODNet model). The solid blue line indicates the ideal scenario. (b) Zoom-in of (a) to better visualize the small values of d_{RMS}. (c) Scatter plot of the error bars (in grey) and absolute errors with their histogram (in blue) from (a) with respect to the RMS coefficient.

of the target property than for smaller ones. This is reflected both by the absolute error and the uncertainty of the prediction (error bar). This characteristic was to be expected due to the biased nature of the nonlinear dataset which favors low SHG response. The histogram on the right of Figure 3.12c clearly shows a similarly skewed distribution as the one of the DB (see Figure 3.8a).

3.7 Predicting the Mg-IV-V_2 compounds

The initial goal of the ML aspect of the present work was to feed the MP to the final model, thereby identifying candidates with high SHG response, i.e., with high RMS coefficient. However, the poor performance of all the models investigated prevents their intended use. This is illustrated by Figure 3.13 for the case of the Mg-IV-V_2 compounds that were introduced in the previous sections. It clearly shows the non-negligible errors that render this model not yet ready for use. Indeed, although the global trend is respected, the relative ranking of these six materials (which is the same as the one based on their d_{36} component or on their $d_{\mathrm{eff,KP}}$) is not respected by the predictions. Indeed, the latter are 11.61, 14.44, 14.31, 16.66, 21.40, and 20.90 pm/V for $MgSiP_2$, $MgSnP_2$, $MgGeP_2$, $MgSiAs_2$, $MgSnAs_2$, and $MgGeAs_2$ instead of 15.92, 21.17, 24.67, 26.02, 44.73, and 49.20 pm/V, respectively. As expected, $MgGeAs_2$ and $MgSnAs_2$ suffer particularly from the bias of the nonlinear DB since their RMS coefficient is larger than the other ones. Surprisingly, the uncertainty on the predictions are rather small compared to the ones observed previously for the test sets.



(a) **(b)**

Figure 3.13: (a) Predictions of the Mg-IV-V_2 RMS coefficients (red dots) with their related error bars (in grey) in the case of the GA0 model without any centrosymmetric compounds. The error bars correspond to the standard deviation of the individual predictions of the GA0 (Ensemble MODNet model). The solid blue line indicates the ideal scenario. (b) Scatter plot of the error bars (in grey) and absolute errors (in blue) from (a) with respect to the RMS coefficient.

3.8 Conclusion

This third chapter dealt with a ML approach to this investigation of the SHG phenomenon. The pre-requisites (ANN, features selection,...) to understand such method were introduced in Section 3.1.1 while MODNet, the model used in the present work, was presented in Section 3.1.2. The choice of an adequate target property was developed in Section 3.2. Indeed, the KP effective coefficient was found to be dependent on the orientation of the system. Thus, a new effective coefficient, i.e., the RMS of the SHG components, was rather chosen as target. A short analysis of the relevant DB was done in Section 3.3 which showed the latter to be strongly biased towards low values of the target. In Section 3.4, four different copies of the DB were created in order to investigate the inclusion of the static refractive index and the bandgap. These datasets were then duplicated to include several proportions of centrosymmetric materials. They were subsequently fed to MODNet which performed five different hyperparameters optimization. The generalization score was estimated based on a 5×5 nested cross-validation as explained in Section 3.5. The results were presented in Section 3.6. The best model was found to be the one which optimized its hyperparameters using the genetic algorithm. The addition of n_s and E_g also improved the performance although the inclusion of centrosymmetric instances worsened them. In Section 3.7, the best model predicted the RMS coefficients of the six Mg-IV-V_2 of the previous sections. Although the global trend was respected, the exact ordering was not.

Conclusion

Functional materials are nowadays at the core of most technologies. Designing new compounds can thus bring considerable improvements on many aspects and can even lead to the development of new applications. To that end, first-principles quantum simulations allow to single out and theoretically study relevant materials, thereby helping the subsequent experimental investigations. Moreover, this approach has led to the emergence of large materials databases. Therefore, statistical and machine learning algorithms have also become interesting tools in materials science due to their low computational cost. The present work made use of both paradigms (FP and ML) in an attempt to discover new nonlinear optical crystals. Although lacking in certain regions of the spectrum, this type of compound is critical in the field of optoelectronics since it allows the exploitation of a broader range of frequencies.

In Chapter 1, several key concepts of nonlinear optics were introduced. The existence of NLO phenomena (and more precisely of SHG) were explained in a classical framework thanks to the anharmonic oscillator model. Miller's rule, a relationship between linear and nonlinear optics, was presented as it holds great potential in this search for new NLO crystals. The latter are however only effective if they allow phase matching as demonstrated. This necessary condition can be attained through angular phase matching which combines the property of birefringence and the normal dispersion of the refractive index. Historically, the need for large single crystals constituted an experimental challenge to assess the potential of crystals to realize angular phase matching. This led to the Kurtz-Perry powder technique and the introduction of an effective SHG coefficient to characterize the strength of the SHG response of any materials.

This theoretical background was necessary for the subsequent investigation of six NLO crystals candidates in Chapter 2. FP computations with LDA functionals were applied to evaluate the NLO potential of MgSi(-Ge,-Sn)P_2(-As_2) in the mid-IR region which particularly suffers from the lack of suitable compounds. These six candidates adopt a chalcopyrite structure. They are thus described by a ratio c/a different from 2 and a distortion parameter, u, which both indicate the deviation from a zincblende structure. The LDA ratio c/a agrees with experiments while u was found to be in accordance with GGA results from the literature. $MgSnAs_2$ is the only compound with a distortion parameter lower than 0.25. This implies that it displays a deviation from the zincblende structure which is inverted compared to the other five materials. Regarding the electronic structure, MgSi(-Ge)P_2 and $MgSiAs_2$ display a pseudo-direct bandgap resulting from an indirect one in their binary zincblende analogs. $MgGeAs_2$ and $MgSnAs_2$ were found to display an adequate birefringence for angular phase matching as well as a high SHG response. The latter was analyzed with respect to their structural, electronic, and linear optics properties. An initial assumption was the dependence of the SHG on a "degree of centrosymmetry" since such phenomenon is absent from centrosymmetric compounds. The results seem to show that another factor (yet unidentified) possesses more influence since no clear trend could be

observed between the different parameters. However, Miller's coefficient might well be more impacted by this quantification of centrosymmetry. Indeed, it is observed that δ decreases when the atomic structure is modified into a pseudo-zincblende one in order to approach a "more centrosymmetric" compound. A similar outcome was found when investigating zincblende binary analogs of the six initial materials. This must however be nuanced by the investigation of "inverse u" structures which was not able to confirm this claim based on the present understanding of the quantification of the centrosymmetry.

Chapter 3 approached the thematic from a machine learning point of view thanks to the recent emergence of a database with the appropriate NLO properties. The basics of this domain were first presented as well as MODNet, the chosen model for this task. Due to the dependence of the Kurtz-Perry effective coefficient on the orientation of the system, a new quantity was designed based on eigenvalues of d_{ijk} to serve as target variable. However, the dataset turned out to be highly skewed towards low SHG response. This proved to be the main reason behind the poor performance of the best model despite the inclusion of the bandgap and/or the refractive index. Various proportions of centrosymmetric instances were also added in an attempt to improve the sampling of the feature space. However, this led to a worsening of the performances.

Following in the footsteps of previous research, the present document opens the way to many possibilities for future works. Indeed, the six investigated compounds constitute a great system to study the relations between structural, electronic, and NLO properties. It would thus be interesting to observe the variation of the SHG and δ when varying the distortion parameter. The zincblende binary analogs also promise a lot of information. Most of them are well known materials and their relation to their ternary chalcopyrite analogs has already been researched with respect to their structure and bandgap. A deeper bibliographic search should therefore be conducted in order to determine if the known relations between the two systems could have an impact on the SHG phenomenon as well.

Regarding the ML approach, many perspectives exist. The Miller's coefficient and other invariants as the RMS one should be tested as target. The analytic formula of the KP effective coefficient should also be verified. Moreover, MODNet is especially designed for joint-learning purposes, i.e., to predict multiple targets at once and to benefit from them. This aspect could be explored with many of the existing invariants of the SHG tensor. In parallel, other ML models should be tested since each one has its own specialty. A more "structure oriented" model could be interesting to further investigate this concept of "degree of centrosymmetry." An alternative would be to explore different features such as derived from the bandstructure or DOS for example. A deeper analysis of the DB is also necessary.

Nonlinear optics is a fascinating field of research. Although not obvious at first glance, it is present in many aspects of everyday life. This greatly motivates the need for materials research in this area. The questions raised by the present work should thus serve as catalysts.

Bibliography

[1] C. L. Magee, "Towards quantification of the role of materials innovation in overall technological development," *Complexity*, vol. 18, no. 1, pp. 10–25, 2012. [Cited on page 1.]

[2] A. Jain, K. A. Persson, and G. Ceder, "Research update: The materials genome initiative: Data sharing and the impact of collaborative ab initio databases," *APL Materials*, vol. 4, no. 5, p. 053102, 2016. [Cited on page 1.]

[3] J. Hafner, C. Wolverton, and G. Ceder, "Toward computational materials design: the impact of density functional theory on materials research," *MRS bulletin*, vol. 31, no. 9, pp. 659–668, 2006. [Cited on page 1.]

[4] G. R. Schleder, A. C. Padilha, C. M. Acosta, M. Costa, and A. Fazzio, "From DFT to machine learning: recent approaches to materials science–a review," *Journal of Physics: Materials*, vol. 2, no. 3, p. 032001, 2019. [Cited on pages 1 and 2.]

[5] R. A. Ganeev, *Nonlinear optical properties of materials*, vol. 174. Springer, 2013. [Cited on page 2.]

[6] F. Naccarato, *First-principles high-throughput study of linear and nonlinear optical materials*. PhD thesis, University of Luxembourg, Luxembourg, Luxembourg, 2019. [Cited on pages 3, 8, 31, 35, 44, 48, and 51.]

[7] C. Garrett, "Nonlinear optics, anharmonic oscillators, and pyroelectricity," *IEEE Journal of Quantum Electronics*, vol. 4, no. 3, pp. 70–84, 1968. [Cited on page 4.]

[8] X. Gonze, "Optical properties of materials," 2020. (syllabus, LMAPR2014 - Physics of functional materials, UCLouvain). [Cited on pages 5 and 12.]

[9] A. Levi, "The lorentz oscillator model," *Essential classical mechanics for device physics*, pp. 5–1, 2016. [Cited on page 5.]

[10] N. Bloembergen, *Nonlinear Optics*. W.A. Benjamin, Inc., 1965. [Cited on page 5.]

[11] F. Zernike and J. E. Midwinter, *Applied nonlinear optics*. Courier Corporation, 2006. [Cited on pages 6, 8, 10, 12, 13, 15, 17, 18, and 19.]

[12] R. W. Boyd, *Nonlinear optics*. Academic press, 2020. [Cited on pages 6, 7, 13, 14, 15, and 18.]

[13] R. C. Miller, "Optical second harmonic generation in piezoelectric crystals," *Applied Physics Letters*, vol. 5, no. 1, pp. 17–19, 1964. [Cited on page 8.]

[14] G.-M. Rignanese, "Physics of functional materials," 2020. (syllabus, LMAPR2014 - Physics of functional materials, UCLouvain). [Cited on page 8.]

[15] F. Naccarato, F. Ricci, J. Suntivich, G. Hautier, L. Wirtz, and G.-M. Rignanese, "Searching for materials with high refractive index and wide band gap: A first-principles high-throughput study," *Physical Review Materials*, vol. 3, no. 4, p. 044602, 2019. [Cited on pages 9, 31, 32, and 33.]

[16] A. Jackson, M. Ohmer, and S. LeClair, "Relationship of the second order nonlinear optical coefficient to energy gap in inorganic non-centrosymmetric crystals," *Infrared Physics & Technology*, vol. 38, no. 4, pp. 233–244, 1997. [Cited on page 9.]

[17] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013. [Cited on pages 10 and 18.]

[18] M. Simões, W. Bertolino, D. Tovar, W. Braga, O. Santos, D. Luders, A. Sampaio, N. Kimura, and A. Palangana, "Optical signal and optical axes in uniaxial and biaxial nematic phases," *Phase Transitions*, vol. 92, pp. 1–9, 12 2018. [Cited on page 11.]

[19] W. Zhang, H. Yu, H. Wu, and P. S. Halasyamani, "Phase-matching in nonlinear optical compounds: a materials perspective," *Chemistry of Materials*, vol. 29, no. 7, pp. 2655–2668, 2017. [Cited on pages 12, 17, and 18.]

[20] F. D. Bloss, *An introduction to the methods of optical crystallography*. No. 548 BLO, 1961. [Cited on page 12.]

[21] J. Armstrong, N. Bloembergen, J. Ducuing, and P. S. Pershan, "Interactions between light waves in a nonlinear dielectric," *Physical review*, vol. 127, no. 6, p. 1918, 1962. [Cited on page 13.]

[22] D. Kleinman, "Nonlinear dielectric polarization in optical media," *Physical Review*, vol. 126, no. 6, p. 1977, 1962. [Cited on page 13.]

[23] J. Midwinter and J. Warner, "The effects of phase matching method and of uniaxial crystal symmetry on the polar distribution of second-order non-linear optical polarization," *British Journal of Applied Physics*, vol. 16, no. 8, p. 1135, 1965. [Cited on pages 14, 16, and 17.]

[24] V. Dmitriev and D. Nikogosyan, "Effective nonlinearity coefficients for three-wave interactions in biaxial crystal of mm2 point group symmetry," *Optics communications*, vol. 95, no. 1-3, pp. 173–182, 1993. [Cited on pages 14, 17, and 20.]

[25] M. Hobden and J. Warner, "The temperature dependence of the refractive indices of pure lithium niobate," *Physics Letters*, vol. 22, no. 3, pp. 243–244, 1966. [Cited on page 16.]

[26] M. M. Fejer, G. Magel, D. H. Jundt, and R. L. Byer, "Quasi-phase-matched second harmonic generation: tuning and tolerances," *IEEE Journal of quantum electronics*, vol. 28, no. 11, pp. 2631–2654, 1992. [Cited on page 16.]

[27] P. Maker, R. Terhune, M. Nisenoff, and C. Savage, "Effects of dispersion and focusing on the production of optical harmonics," *Physical review letters*, vol. 8, no. 1, p. 21, 1962. [Cited on page 16.]

[28] J. Giordmaine, "Mixing of light beams in crystals," *Physical Review Letters*, vol. 8, no. 1, p. 19, 1962. [Cited on page 16.]

[29] M. Hobden, "Phase-matched second-harmonic generation in biaxial crystals," *Journal of Applied Physics*, vol. 38, no. 11, pp. 4365–4372, 1967. [Cited on page 16.]

[30] V. G. Dmitriev, G. G. Gurzadyan, and D. N. Nikogosyan, *Handbook of nonlinear optical crystals*, vol. 64. Springer, 2013. [Cited on page 18.]

[31] S. Kurtz and T. Perry, "A powder technique for the evaluation of nonlinear optical materials," *Journal of applied physics*, vol. 39, no. 8, pp. 3798–3813, 1968. [Cited on pages 18 and 19.]

[32] I. Aramburu, J. Ortega, C. Folcia, and J. Etxebarria, "Second harmonic generation by micropowders: A revision of the kurtz–perry method and its practical application," *Applied Physics B*, vol. 116, no. 1, pp. 211–233, 2014. [Cited on page 19.]

[33] G. Bourhill, K. Mansour, K. J. Perry, L. Khundkar, E. T. Sleva, R. Kern, J. W. Perry, I. D. Williams, and S. K. Kurtz, "Powder second harmonic generation efficiencies of saccharide materials," *Chemistry of materials*, vol. 5, no. 6, pp. 802–808, 1993. [Cited on page 19.]

[34] T. K. Begzjav and R. Nessler, "On three-dimensional rotational averages of odd-rank tensors," *Physica Scripta*, vol. 94, no. 10, p. 105504, 2019. [Cited on pages 19 and 59.]

[35] S. Cyvin, J. Rauch, and J. Decius, "Theory of hyper-raman effects (nonlinear inelastic light scattering): Selection rules and depolarization ratios for the second-order polarizability," *The Journal of Chemical Physics*, vol. 43, no. 11, pp. 4083–4095, 1965. [Cited on page 19.]

[36] J. Yao and T. S. Fahlen, "Calculations of optimum phase match parameters for the biaxial crystal ktiopo4," *Journal of applied physics*, vol. 55, no. 1, pp. 65–68, 1984. [Cited on page 20.]

[37] B. Boulanger, J. P. Fève, G. Marnier, B. Ménaert, X. Cabirol, P. Villeval, and C. Bonnin, "Relative sign and absolute magnitude of $d^{(2)}$ nonlinear coefficients of ktp from second-harmonic-generation measurements," *JOSA B*, vol. 11, no. 5, pp. 750–757, 1994. [Cited on page 20.]

[38] J. Xiao, S. Zhu, B. Zhao, B. Chen, H. Liu, and Z. He, "Computational assessment of promising mid-infrared nonlinear optical materials Mg–IV–V2 (IV= Si, Ge, Sn; V=

P, As): a first-principles study," *Materials Research Express*, vol. 5, no. 3, p. 035907, 2018. [Cited on pages 22, 25, 26, 27, 28, 34, 35, and 37.]

[39] X. Gonze, B. Amadon, G. Antonius, F. Arnardi, L. Baguet, J.-M. Beuken, J. Bieder, F. Bottin, J. Bouchet, E. Bousquet, N. Brouwer, F. Bruneval, G. Brunin, T. Cavignac, J.-B. Charraud, W. Chen, M. Côté, S. Cottenier, J. Denier, G. Geneste, P. Ghosez, M. Giantomassi, Y. Gillet, O. Gingras, D. R. Hamann, G. Hautier, X. He, N. Helbig, N. Holzwarth, Y. Jia, F. Jollet, W. Lafargue-Dit-Hauret, K. Lejaeghere, M. A. L. Marques, A. Martin, C. Martins, H. P. C. Miranda, F. Naccarato, K. Persson, G. Petretto, V. Planes, Y. Pouillon, S. Prokhorenko, F. Ricci, G.-M. Rignanese, A. H. Romero, M. M. Schmitt, M. Torrent, M. J. van Setten, B. V. Troeye, M. J. Verstraete, G. Zérah, and J. W. Zwanziger, "The abinit project: Impact, environment and recent developments," *Comput. Phys. Commun.*, vol. 248, p. 107042, 2020. [Cited on page 22.]

[40] A. H. Romero, D. C. Allan, B. Amadon, G. Antonius, T. Applencourt, L. Baguet, J. Bieder, F. Bottin, J. Bouchet, E. Bousquet, F. Bruneval, G. Brunin, D. Caliste, M. Côté, J. Denier, C. Dreyer, P. Ghosez, M. Giantomassi, Y. Gillet, O. Gingras, D. R. Hamann, G. Hautier, F. Jollet, G. Jomard, A. Martin, H. P. C. Miranda, F. Naccarato, G. Petretto, N. A. Pike, V. Planes, S. Prokhorenko, T. Rangel, F. Ricci, G.-M. Rignanese, M. Royo, M. Stengel, M. Torrent, M. J. van Setten, B. V. Troeye, M. J. Verstraete, J. Wiktor, J. W. Zwanziger, and X. Gonze, "Abinit: Overview, and focus on selected capabilities," *J. Chem. Phys.*, vol. 152, p. 124102, 2020. [Cited on page 22.]

[41] A. Springthorpe and J. Harrison, "Mg Si P2: a new member of the II IV V2 family of semiconducting compounds," *Nature (London)*, vol. 222, pp. 977–977, 1969. [Cited on page 22.]

[42] R. Trykozko and N. Goryunova, "Preparation and some properties of Mg Si P2," *Izvestiya Akademii Nauk SSSR, Neorganicheskie Materialy*, vol. 4, pp. 2101–2105, 1968. [Cited on page 22.]

[43] A. Petukhov, W. Lambrecht, and B. Segall, "Electronic structure of wide-band-gap ternary pnictides with the calcopyrite structure," *Physical Review, Serie 3. B - Condensed Matter (18,1978-)*, vol. 49, pp. 4549–4558, 1994. [Cited on page 22.]

[44] A. Vaipolin, "Specific defects of the structure of compounds A(2+) B(4+) C2(3-)," *Fizika Tverdogo Tela (Leningrad) (= Solid State Physics)*, vol. 15, pp. 1430–1435, 1973. [Cited on page 22.]

[45] T. Ouahrani, "Chemical and physical insight on the local properties of the phosphides XSip2 (X = Be, Mg, Cd, Zn and Hg) under pressure: from first principles calculations," *The European Physical Journal B*, vol. 86, pp. 369–1–369–11, 2013. [Cited on page 22.]

[46] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. a. Persson, "Commentary: The Materials Project: A materials genome approach to accelerating materials innovation," *APL Materials*, vol. 1, no. 1, p. 011002, 2013. [Cited on page 22.]

[47] G. Kresse and J. Furthmüller, "Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set," *Physical review B*, vol. 54, no. 16, p. 11169, 1996. [Cited on page 22.]

[48] X. Gonze, "First-principles responses of solids to atomic displacements and homogeneous electric fields: Implementation of a conjugate-gradient algorithm," *Phys. Rev. B*, vol. 55, pp. 10337–10354, April 1997. [Cited on page 23.]

[49] X. Gonze and C. Lee, "Dynamical matrices, born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory," *Phys. Rev. B*, vol. 55, pp. 10355–10368, April 1997. [Cited on page 23.]

[50] M. Veithen, X. Gonze, and P. Ghosez, "Nonlinear optical susceptibilities, raman efficiencies, and electro-optic tensors from first-principles density functional perturbation theory," *Phys. Rev. B*, vol. 71, p. 125107, March 2005. [Cited on page 23.]

[51] X. Gonze and J.-P. Vigneron, "Density-functional approach to nonlinear-response coefficients of solids," *Phys. Rev. B*, vol. 39, pp. 13120–13128, 1989. [Cited on page 23.]

[52] X. Gonze, "Adiabatic density-functional perturbation theory," *Phys. Rev. A*, vol. 52, pp. 1096–1114, August 1995. [Cited on page 23.]

[53] D. Hamann, "Optimized norm-conserving vanderbilt pseudopotentials," *Physical Review B*, vol. 88, no. 8, p. 085117, 2013. [Cited on page 23.]

[54] M. J. van Setten, M. Giantomassi, E. Bousquet, M. J. Verstraete, D. R. Hamann, X. Gonze, and G.-M. Rignanese, "The pseudodojo: Training and grading a 85 element optimized norm-conserving pseudopotential table," *Computer Physics Communications*, vol. 226, pp. 39–54, 2018. [Cited on page 23.]

[55] J. Jaffe and A. Zunger, "Theory of the band-gap anomaly in ABC2 chalcopyrite semiconductors," *Physical Review B*, vol. 29, no. 4, p. 1882, 1984. [Cited on pages 24 and 47.]

[56] V. Shaposhnikov, A. Krivosheeva, V. Borisenko, J.-L. Lazzari, and F. A. d'Avitaya, "Ab initio modeling of the structural, electronic, and optical properties of $A^{II}B^{IV}C_2^V$ semiconductors," *Physical Review B*, vol. 85, no. 20, p. 205201, 2012. [Cited on pages 24, 25, and 26.]

[57] M. I. Aroyo, J. M. Perez-Mato, C. Capillas, E. Kroumova, S. Ivantchev, G. Madariaga, A. Kirov, and H. Wondratschek, "Bilbao crystallographic server: I. databases and crystallographic computing programs," *Zeitschrift für Kristallographie-Crystalline Materials*, vol. 221, no. 1, pp. 15–27, 2006. [Cited on page 24.]

[58] M. I. Aroyo, A. Kirov, C. Capillas, J. Perez-Mato, and H. Wondratschek, "Bilbao crystallographic server. II. representations of crystallographic point groups and space groups," *Acta Crystallographica Section A: Foundations of Crystallography*, vol. 62, no. 2, pp. 115–128, 2006. [Cited on page 24.]

[59] M. I. Aroyo, J. M. Perez-Mato, D. Orobengoa, E. Tasci, G. de la Flor, and A. Kirov, "Crystallography online: Bilbao crystallographic server," *Bulg. Chem. Commun*, vol. 43, no. 2, pp. 183–197, 2011. [Cited on page 24.]

[60] O. Madelung, U. Rössler, and M. Schulz, "Ternary compounds, organic semiconductors," *Landolt-Bornstein- Group III Condensed Matter*, 2000. [Cited on page 25.]

[61] M. Arrigoni and G. K. Madsen, "Comparing the performance of LDA and GGA functionals in predicting the lattice thermal conductivity of III–V semiconductor materials in the zincblende structure: The cases of AlAs and BAs," *Computational Materials Science*, vol. 156, pp. 354–360, 2019. [Cited on page 25.]

[62] J. Jaffe and A. Zunger, "Electronic structure of the ternary pnictide semiconductors $ZnSiP_2$, $ZnGeP_2$, $ZnSnP_2$, $ZnSiAs_2$, and $MgSiP_2$," *Physical Review B*, vol. 30, no. 2, p. 741, 1984. [Cited on page 26.]

[63] F. Chiker, B. Abbar, B. Bouhafs, and P. Ruterana, "Interband transitions of wide-band-gap ternary pnictide $BeCN_2$ in the chalcopyrite structure," *physica status solidi (b)*, vol. 241, no. 2, pp. 305–316, 2004. [Cited on page 28.]

[64] J. Shay, E. Buehler, and J. Wernick, "Optical properties and electronic structure of $ZnSiAs_2$," *Physical Review B*, vol. 3, no. 6, p. 2004, 1971. [Cited on page 28.]

[65] J. Shay, B. Tell, E. Buehler, and J. Wernick, "Band structure of $ZnGeP_2$ and $ZnSiP_2$—ternary compounds with pseudodirect energy gaps," *Physical Review Letters*, vol. 30, no. 20, p. 983, 1973. [Cited on page 28.]

[66] Z. Ma, J. Hu, R. Sa, Q. Li, Y. Zhang, and K. Wu, "Screening novel candidates for mid-IR nonlinear optical materials from I 3–V–VI 4 compounds," *Journal of Materials Chemistry C*, vol. 5, no. 8, pp. 1963–1972, 2017. [Cited on page 30.]

[67] S. V. Gallego, J. Etxebarria, L. Elcoro, E. S. Tasci, and J. M. Perez-Mato, "Automatic calculation of symmetry-adapted tensors in magnetic and non-magnetic materials: a

new tool of the Bilbao Crystallographic Server," *Acta Crystallographica Section A: Foundations and Advances*, vol. 75, no. 3, pp. 438–447, 2019. [Cited on page 33.]

[68] D. N. Nikogosyan, *Nonlinear optical crystals: a complete survey*. Springer Science & Business Media, 2006. [Cited on page 34.]

[69] M.-H. Lee, C.-H. Yang, and J.-H. Jan, "Band-resolved analysis of nonlinear optical properties of crystalline and molecular materials," *Physical Review B*, vol. 70, no. 23, p. 235110, 2004. [Cited on page 37.]

[70] M. Mutailipu, Z. Li, M. Zhang, D. Hou, Z. Yang, B. Zhang, H. Wu, and S. Pan, "The mechanism of large second harmonic generation enhancement activated by Zn 2+ substitution," *Physical Chemistry Chemical Physics*, vol. 18, no. 48, pp. 32931–32936, 2016. [Cited on page 37.]

[71] E. W. Weisstein, "Correlation coefficient," *https://mathworld. wolfram. com/*, 2006. [Cited on page 37.]

[72] A. G. Asuero, A. Sayago, and A. Gonzalez, "The correlation coefficient: An overview," *Critical reviews in analytical chemistry*, vol. 36, no. 1, pp. 41–59, 2006. [Cited on page 37.]

[73] Y. Dodge, *The concise encyclopedia of statistics*. Springer Science & Business Media, 2008. [Cited on page 39.]

[74] D. Wackerly, W. Mendenhall, and R. L. Scheaffer, *Mathematical statistics with applications*. Cengage Learning, 2014. [Cited on page 39.]

[75] J. H. Zar, "Significance testing of the Spearman rank correlation coefficient," *Journal of the American Statistical Association*, vol. 67, no. 339, pp. 578–580, 1972. [Cited on page 40.]

[76] S. Chatterjee, "A new coefficient of correlation," *Journal of the American Statistical Association*, vol. 116, no. 536, pp. 2009–2022, 2021. [Cited on pages 40 and 41.]

[77] Z. H. Levine and D. C. Allan, "Calculation of the nonlinear susceptibility for optical second-harmonic generation in III-V semiconductors," *Physical review letters*, vol. 66, no. 1, p. 41, 1991. [Cited on page 48.]

[78] J. Heyd, J. E. Peralta, G. E. Scuseria, and R. L. Martin, "Energy band gaps and lattice parameters evaluated with the Heyd-Scuseria-Ernzerhof screened hybrid functional," *The Journal of chemical physics*, vol. 123, no. 17, p. 174101, 2005. [Cited on page 48.]

[79] B. Levine and C. Bethea, "Nonlinear susceptibility of gap; relative measurement and use of measured values to determine a better absolute value," *Applied Physics Letters*, vol. 20, no. 8, pp. 272–275, 1972. [Cited on page 48.]

[80] J. W. Nicklas and J. W. Wilkins, "Accurate ab initio predictions of III–V direct-indirect band gap crossovers," *Applied Physics Letters*, vol. 97, no. 9, p. 091902, 2010. [Cited on page 48.]

[81] G. Ceder and K. Persson, "The materials project: A materials genome approach," 2010. [Cited on page 51.]

[82] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, "Machine learning for molecular and materials science," *Nature*, vol. 559, no. 7715, pp. 547–555, 2018. [Cited on page 52.]

[83] R. Reed and R. J. MarksII, *Neural smithing: supervised learning in feedforward artificial neural networks*. Mit Press, 1999. [Cited on page 53.]

[84] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proceedings of COMPSTAT'2010*, pp. 177–186, Springer, 2010. [Cited on page 53.]

[85] L. Ward, A. Dunn, A. Faghaninia, N. E. Zimmermann, S. Bajaj, Q. Wang, J. Montoya, J. Chen, K. Bystrom, M. Dylla, *et al.*, "Matminer: An open source toolkit for materials data mining," *Computational Materials Science*, vol. 152, pp. 60–69, 2018. [Cited on page 53.]

[86] J. A. Lee and M. Verleysen, *Nonlinear dimensionality reduction*, vol. 1. Springer, 2007. [Cited on page 54.]

[87] Y. Yang, "Consistency of cross validation for comparing regression procedures," *The Annals of Statistics*, vol. 35, no. 6, pp. 2450–2473, 2007. [Cited on page 54.]

[88] J. Shao, "Linear model selection by cross-validation," *Journal of the American statistical Association*, vol. 88, no. 422, pp. 486–494, 1993. [Cited on page 54.]

[89] S. Raschka and V. Mirjalili, "Learning best practices for model evaluation and hyperparameter tuning," *Python Mach. Learn., Third, Packt Publishing*, pp. 280–321, 2019. [Cited on page 55.]

[90] S. Raschka, "Model evaluation, model selection, and algorithm selection in machine learning," *arXiv preprint arXiv:1811.12808*, 2018. [Cited on page 55.]

[91] S. Varma and R. Simon, "Bias in error estimation when using cross-validation for model selection," *BMC bioinformatics*, vol. 7, no. 1, pp. 1–8, 2006. [Cited on page 55.]

[92] P.-P. De Breuck, G. Hautier, and G.-M. Rignanese, "Materials property prediction for limited datasets enabled by feature selection and joint learning with modnet," *npj Computational Materials*, vol. 7, no. 1, pp. 1–8, 2021. [Cited on pages 55 and 56.]

[93] A. Dunn, Q. Wang, A. Ganose, D. Dopp, and A. Jain, "Benchmarking materials property prediction methods: the Matbench test set and Automatminer reference algorithm," *npj Computational Materials*, vol. 6, no. 1, pp. 1–10, 2020. [Cited on page 55.]

[94] A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," *Physical review E*, vol. 69, no. 6, p. 066138, 2004. [Cited on page 56.]

[95] L. Breiman, "Bagging predictors," *Machine learning*, vol. 24, no. 2, pp. 123–140, 1996. [Cited on page 57.]

[96] P.-P. De Breuck, M. L. Evans, and G.-M. Rignanese, "Robust model benchmarking and bias-imbalance in data-driven materials science: a case study on MODNet," *Journal of Physics: Condensed Matter*, vol. 33, no. 40, p. 404002, 2021. [Cited on page 57.]

[97] B. Lakshminarayanan, A. Pritzel, and C. Blundell, "Simple and scalable predictive uncertainty estimation using deep ensembles," *Advances in neural information processing systems*, vol. 30, 2017. [Cited on page 57.]

[98] R. L. Haupt and S. E. Haupt, *Practical genetic algorithms*. John Wiley & Sons, 2004. [Cited on page 57.]

[99] M. Verleysen, "Machine learning: regression and dimensionality reduction. feature selection," 2021. (slides, LELEC2870 - Machine learning : regression, deep networks and dimensionality reduction, UCLouvain). [Cited on page 57.]

[100] E. W. Swokowski, *Calculus with analytic geometry*. Taylor & Francis, 1979. [Cited on page 58.]

[101] P. Kelly, "Mechanics lecture notes: An introduction to solid mechanics," *Retrieved January*, vol. 25, p. 2019, 2019. [Cited on page 58.]

[102] G. Kindlmann, "Tensor invariants and their gradients," in *Visualization and Processing of Tensor Fields*, pp. 215–224, Springer, 2006. [Cited on page 60.]

[103] L. Qi, "Eigenvalues and invariants of tensors," *Journal of Mathematical Analysis and Applications*, vol. 325, no. 2, pp. 1363–1377, 2007. [Cited on page 60.]

[104] W. H. Yang, "A Generalized von Mises Criterion for Yield and Fracture," *Journal of Applied Mechanics*, vol. 47, pp. 297–300, 06 1980. [Cited on page 60.]

[105] L. Qi, "Transposes, L-Eigenvalues and Invariants of Third Order Tensors," *arXiv preprint arXiv:1704.01327*, 2017. [Cited on page 60.]

[106] L. Qi, H. Chen, and Y. Chen, "Third order tensors in physics and mechanics," in *Tensor eigenvalues and their applications*, pp. 207–248, Springer, 2018. [Cited on page 60.]

[107] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011. [Cited on page 67.]

[108] C. Chen, W. Ye, Y. Zuo, C. Zheng, and S. P. Ong, "Graph networks as a universal machine learning framework for molecules and crystals," *Chemistry of Materials*, vol. 31, no. 9, pp. 3564–3572, 2019. [Cited on page 68.]

[109] C. Spataru, S. Ismail-Beigi, R. Capaz, and S. Louie, "Quasiparticle and Excitonic Effects in the Optical Response of Nanotubes and Nanoribbons," *Topics in Applied Physics*, vol. 111, pp. 195–227, 2008. [Cited on pages 86, 87, 91, 95, 96, and 97.]

[110] F. Giustino, *Materials Modelling Using Density Functional Theory: Properties and Predictions*. Oxford University Press, 2014. [Cited on pages 86, 87, 88, 89, 90, 91, and 96.]

[111] P. Hohenberg and W. Kohn, "Inhomogeneous Electron Gas," *Physical Review*, vol. 136, no. 3B, pp. B864–B871, 1964. [Cited on pages 86 and 87.]

[112] W. Kohn and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects," *Physical Review*, vol. 140, no. 4A, pp. A1133–A1138, 1965. [Cited on page 87.]

[113] P. A. M. Dirac, "Note on Exchange Phenomena in the Thomas Atom," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 26, no. 3, pp. 376–385, 1930. [Cited on page 89.]

[114] D. M. Ceperley and B. J. Alder, "Ground State of the Electron Gas by a Stochastic Method," *Physical Review Letters*, vol. 45, no. 7, pp. 566–569, 1980. [Cited on page 89.]

[115] S. Louie, "Chapter 2 Predicting Materials and Properties: Theory of the Ground and Excited State," *Contemporary Concepts of Condensed Matter Science*, vol. 2, 2006. [Cited on pages 89, 90, 92, 94, 95, 96, and 97.]

[116] J. C. Phillips and L. Kleinman, "New method for calculating wave functions in crystals and molecules," *Physical Review*, vol. 116, no. 2, p. 287, 1959. [Cited on page 89.]

[117] W. E. Pickett, "Pseudopotential methods in condensed matter applications," *Computer Physics Reports*, vol. 9, no. 3, pp. 115–197, 1989. [Cited on page 90.]

[118] S. Cottenier *et al.*, "Density Functional Theory and the family of (L) APW-methods: a step-by-step introduction," *Instituut voor Kern-en Stralingsfysica, KU Leuven, Belgium*, vol. 4, no. 0, p. 41, 2002. [Cited on page 90.]

[119] E. Luppi, H.-C. Weissker, S. Bottaro, F. Sottile, V. Veniard, L. Reining, and G. Onida, "Accuracy of the pseudopotential approximation in ab initio theoretical spectroscopies," *Physical Review B*, vol. 78, no. 24, p. 245124, 2008. [Cited on page 90.]

[120] L. Reining, "The GW approximation: content, successes and limitations," *WIREs Computational Molecular Science*, vol. 8, no. 3, p. e1344, 2018. [Cited on pages 91, 94, 95, and 96.]

[121] L. Hedin and S. Lundqvist, "Effects of Electron-Electron and Electron-Phonon Interactions on the One-Electron States of Solids," in *Solid State Physics* (F. Seitz, D. Turnbull, and H. Ehrenreich, eds.), vol. 23, pp. 1–181, Academic Press, 1970. [Cited on pages 92, 93, 94, and 95.]

[122] P. A. M. Dirac, "On the theory of quantum mechanics," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 112, no. 762, pp. 661–677, 1926. [Cited on page 92.]

[123] L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics: Green's Function Methods in Equilibrium and Nonequilibrium Problems*. Boca Raton: CRC Press, 2019. [Cited on pages 93 and 94.]

[124] M. S. Hybertsen and S. G. Louie, "Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies," *Phys. Rev. B*, vol. 34, pp. 5390–5413, 1986. [Cited on pages 95 and 96.]

[125] L. Hedin, "New method for calculating the one-particle green's function with application to the electron-gas problem," *Phys. Rev.*, vol. 139, pp. A796–A823, 1965. [Cited on pages 95 and 96.]

[126] F. Tran and P. Blaha, "Accurate band gaps of semiconductors and insulators with a semilocal exchange-correlation potential," *Phys. Rev. Lett.*, vol. 102, p. 226401, 2009. [Cited on page 96.]

[127] G. Strinati, "Application of the Green's functions method to the study of the optical properties of semiconductors," *La Rivista del Nuovo Cimento (1978-1999)*, vol. 11, no. 12, pp. 1–86, 1988. [Cited on page 97.]

Appendices

Appendix A

First-principles methods

First-principles methods refer to the use of quantum mechanics associated to some approximations in order to quantitatively compute physical properties of the system of interest. Different approaches allow to play between the accuracy of the results and the computational resources required to obtain them. This section aims at describing the methods of quantum simulations that are used to develop DBs of materials or to investigate the NLO properties of specific compounds. The structure of the section is greatly inspired by the one of Ref. [109].

A.1 Density Functional Theory (DFT)

The first step in this kind of approach is to obtain the optimised geometry of the system, i.e. the geometry minimising its total energy. To do so, the common approach is to use the density functional theory (DFT), a pillar among the frameworks of ab-initio calculations in solid-state physics. This theory is mainly based on three components: the Hohenberg-Kohn theorem, the Hohenberg-Kohn variational principle and the Kohn-Sham reformulation [110]. The Hohenberg-Kohn theorem proves that the total energy of a many-electron system, E, can be expressed as a functional, F, of the electron density, ρ, instead of the many-body wavefunction, Ψ, [111]:

$$E = F[\rho] \tag{A.1}$$

where the use of the squared brackets underlines the fact that it is a functional. By considering fixed nuclei, the many-electron Hamiltonian can be written as [110]:

$$\hat{H} = \sum_i V_\rho(\mathbf{r}_i) \underbrace{- \sum_i \frac{1}{2}\nabla_i^2}_{\hat{T}} + \underbrace{\frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}}_{\hat{W}} \tag{A.2}$$

where i is the index of each electron, which feels a Coulomb potential V_ρ due to all the nuclei of the system, $\hat{T}$ is the kinetic energy, and $\hat{W}$ is the Coulomb energy due to the interaction between each electron. This implies that the functional F can be expressed as [110]:

$$F[\rho] = \int d\mathbf{r}\rho(\mathbf{r})V_\rho(\mathbf{r}) + \langle\Psi[\rho]|\hat{T} + \hat{W}|\Psi[\rho]\rangle \tag{A.3}$$

$$\overset{K-S}{=} \overbrace{\underbrace{\int d\mathbf{r}\rho(\mathbf{r})V_\rho(\mathbf{r})}_{\text{External potential}} - \underbrace{\sum_i \int d\mathbf{r}\phi_i^*(\mathbf{r})\frac{\nabla^2}{2}\phi_i(\mathbf{r})}_{\text{Kinetic energy}} + \underbrace{\frac{1}{2}\iint d\mathbf{r}d\mathbf{r}'\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}}_{\text{Hartree energy}} + \underbrace{E_{xc}[\rho]}_{\text{XC energy}}}^{\text{Total energy in the independent electrons approximation}} \tag{A.4}$$

where ϕ_i is the wavefunction of the i^{th} electron and "K-S" refers to the idea of Kohn and Sham. They actually re-expressed the terms with Ψ in the previous equation as the kinetic and Coulomb energy (Hartree energy) of independent electrons and put the "error" (quantum electron-electron interaction) due to this reformulation in E_{xc}, the exchange and correlation energy [112]. The latter term is therefore the only unknown part. Then, the Hohenberg-Kohn variational principle states that the ground-state density, ρ_0, minimises the total energy [111]:

$$\left.\frac{\delta F[\rho]}{\delta\rho}\right|_{\rho_0} = 0 \tag{A.5}$$

By combining Eq. (A.3) and Eq. (A.5) as well as requiring the orthonormality of the independent electron wavefunction, one gets an equation for ϕ_i after a few steps [110]:

$$\left[-\frac{1}{2}\nabla^2 + V_\rho(\mathbf{r}) + V_{\text{H}}(\mathbf{r}) + V_{xc}(\mathbf{r})\right]\phi_i(\mathbf{r}) = \varepsilon_i\phi_i(\mathbf{r}) \tag{A.6}$$

where V_H is the Hartree potential (mean-field approximation) and V_{xc} is the exchange and correlation potential:

$$V_{xc} = \left.\frac{\delta E_{xc}[\rho]}{\delta\rho}\right|_{\rho(\mathbf{r})} \tag{A.7}$$

Eq. (A.6) are known as the Kohn-Sham equations and are the basis of DFT. By solving them self-consistently (iteratively) for ϕ_i, it becomes possible to obtain the electron density:

$$\rho(\mathbf{r}) = \sum_i^{\text{occ}} |\phi_i(\mathbf{r})|^2 \tag{A.8}$$

The problem of the interacting many-electron system has thus been shifted to a set of effective, self-consistent, single-body equations [109]. In order to determine each K-S wavefunction individually, it is necessary to know V_{xc} and V_H. However, those two terms depend on the charge density which depends itself on all ϕ_i as seen from Eq. (A.8). As mentioned above, the proper way to solve this set of equations is thus to solve them self-consistently, i.e. iteratively until a certain convergence threshold, chosen beforehand, is reached for ρ (often for the total energy in practice) [110]. Figure A.1 represents the usual procedure.

The K-S wavefunctions, ϕ_i are only a tool used to build the density and do not have any direct physical meaning. The same statement applies to the K-S eigenvalues, ε_i. However, they might carry some hidden or approximate physical meaning [110]. They form what can be called the "Kohn-Sham band structure." The latter approaches relatively well the real band structure of the material. However, its biggest issue is its underestimation of the band gap of about 40-50% in general [110]. This topic can become quite complex and a

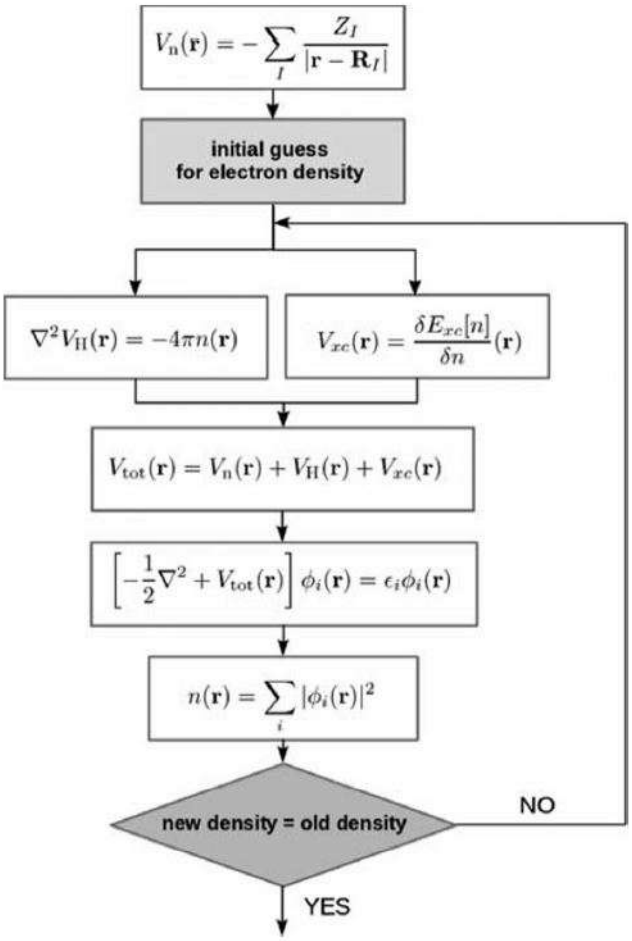


Figure A.1: Self-consistent procedure to solve the K-S equations of Eq. (A.6). Thanks to a set of atomic coordinates given by the user, V_ρ can be constructed and is used to make a first guess of ρ. From this first density, V_H and V_{xc} are calculated and the set of K-S equations are solved via finite differences for example. From the new ϕ_i, a new density is computed which gives a new and better V_{tot}. The procedure is repeated until the new density is equal to the old one up to a convergence threshold defined beforehand. Taken from Ref. [110].

simple approach will be developed later to explain some characteristic of the DFT. Stated otherwise, it is important to stress that the use of the Kohn-Sham variational principle implies that DFT is a ground-state theory. Therefore, it might not be well suited when studying properties dealing with excited states. Unfortunately, it is quite intuitive to consider the optical absorption as one of those quantities and also a quantity of interest in the framework of this paper. Further improvements or alternatives have thus been developed in order to obtain accurate quantitative results in this field as will be seen later. In any case, DFT is a good starting point to get the optimised geometry of the system (via the optimisation of the total energy with respect to the atomic coordinates) and can even be used as a first approximation of the studied properties.

The only practical issue remaining in order to use DFT is the knowledge of E_{xc}. The exact form of this functional would give the exact ground-state density and energy. The goal

is thus to develop and use the best approximations possible of this energy. Many solutions exist and one of the most popular ones is presented hereafter.

A.1.1 Local Density Approximation (LDA)

The LDA corresponds to the simplest approximation of the functional of the exchange and correlation energy. It consists in associating each infinitesimal volume of the real material to a homogeneous electron gas (HEG), i.e. a free electron gas with interaction between the electrons. Each infinitesimal region is described by a local density $\rho(\mathbf{r})$ and this allows to "map" the real system into a HEG [110]:

$$E_{xc} = \int_V \frac{E_{xc}^{\mathrm{HEG}}[\rho(\mathbf{r})]}{V} d\mathbf{r} \tag{A.9}$$

where E_{xc}^{HEG} is the sum of the exchange energy and the correlation energy for a HEG with the right charge density. An analytic expression can be derived for the former [113]:

$$E_{\mathrm{x}} = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}} \rho^{\frac{4}{3}} V \tag{A.10}$$

while the latter requires a numerical parameterisation [114]. The exchange and correlation potential is then obtained via the functional derivative of Eq. (A.7). LDA is attractive because the unknown term in the K-S equations is obtained by knowing only the density at the considered point in space.

A.1.2 Pseudopotentials and planewaves representation

This whole section is based on the appendix of Ref. [110] unless specified otherwise.

As can be intuitively understood, only the outer electrons, i.e. the valence electrons, are of interest when studying materials properties [115]. The density of the valence electrons reaches further away from the nucleus than the core ones. This is equivalent to saying that the valence electrons are mainly the ones participating in chemical bonding and that the core electrons can be considered as inert. This is commonly called the "frozen core approximation" and it allows to discard the core electrons from the computations. The distinction between core and valence electrons is not universally defined and when in doubt, the spatial reach of each atomic wavefunction can help classify them. Often, a good separation is obtained via the usual definition of valence electrons as being the ones of the outer most shell of the periodic table. The valence states wavefunctions display an oscillatory behavior near the nuclei while they are smooth away from them [116]. This results in difficulties to model them in a simple basis set, which would be useful for computational reasons. The pseudopotential approach solves precisely this issue. It consists in substituting the K-S potential by a modified one such that it coincides with the original one far from the nucleus (for $r > r_{\mathrm{cutoff}}$) and such that it results in a precise pseudo-wavefunction close to it (for $r < r_{\mathrm{cutoff}}$). The latter must

be smooth but must still yield the same charge density as the all-electron, i.e. not just valence, wavefunction inside r_{cutoff} as well as give the same value and slope at $r = r_{\text{cutoff}}$. The radial cutoff, r_{cutoff}, is chosen depending on the case and is the limit of the so-called "pseudization region." An example of this approach is depicted in Figure A.2. As mentioned,

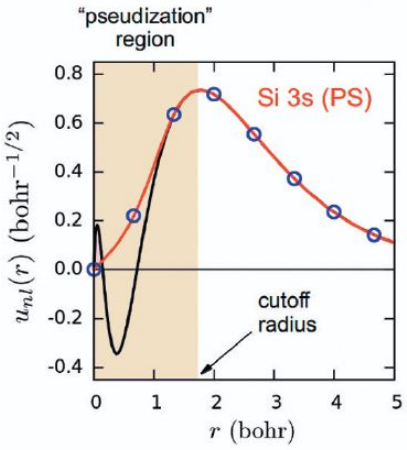


Figure A.2: Example of pseudization of the 3s all-electron wavefunction of Si according to the conditions described in Section A.1.2. Taken from Ref. [110].

the pseudopotential allows the use of a planewaves basis set in order to model the pseudo-wavefunctions. This basis set is the most common strategy. The K-S wavefunctions are expressed as a linear combination of planewaves whose wavevector is a reciprocal lattice vector $\mathbf{G}$:

$$\phi_i(\mathbf{r}) = \sum_{\mathrm{G}} c_i(\mathbf{G}) \exp(i\mathbf{G} \cdot \mathbf{r}) \tag{A.11}$$

However, this Fourier serie must be cutoff at a value $\mathbf{G}_{\text{max}}$, chosen in order to reach convergence of the property of interest. The K-S equations then become:

$$\frac{|\mathbf{G}|^2}{2} c_i(\mathbf{G}) + \sum_{\mathrm{G}'} v_{\text{tot}}\left(\mathbf{G} - \mathbf{G}'\right) c_i\left(\mathbf{G}'\right) = \varepsilon_i c_i(\mathbf{G}) \tag{A.12}$$

This set of equations can be solved self-consistently for a chosen sampling of the first Brillouin zone. The choice of the sampling is also done in order to reach convergence of the studied property.

An infinite number of valid pseudopotentials can be developed, each of them presenting different accuracy and efficiency [115]. In order to optimise both those quantities, many methods of development came to light [117]. As briefly explained in Ref. [118], a pseudopotential can be judged on two criteria: its softness and its transferability. The softer a pseudopotential is, the fewer planewaves will be needed, and the more situation-specific the pseudopotential is. On the contrary, a transferable pseudopotential is one that performs relatively well in any situation. The goal is to optimise both aspects at once.

A detailed analysis of the accuracy of various pseudopotentials compared to the all-electron approach can be found in Ref. [119]. This analysis investigates more precisely the optical

spectra and even non-linear responses. This can be comforting in the use of pseudopotentials for the situation of interest for the current paper.

Up to now, common and necessary requirements to the practical use of the DFT have been presented. However, as explained, DFT is not suited to study properties related to excited states such as the optical properties. It is thus mandatory to go beyond the single-body theory that is DFT [109].

A.2 Beyond DFT

This section explains in more details the need to go beyond the formalism of DFT and presents the usual alternative to it that one adopts.

A few observables are usually the main goal of a study and are expressed as:

$$O = \langle \Psi | \hat{O} | \Psi \rangle \tag{A.13}$$

where Ψ is the ground-state many-body wavefunction for example. They cannot be obtained via the charge density of DFT. The ground-state many-body wavefunction can instead be approximated by the K-S ground-state Slater determinant, itself resulting from the K-S wavefunctions. This implies that the K-S eigenvalues form a band structure, namely the K-S one. This approximation can be quite non-accurate since the K-S eigenvalues are just a mathematical tool. As mentioned, this might not be fully the case but this subject can be complex and is left aside [120]. The next paragraph specifies the meaning of what has been stated.

DFT is a ground-state theory, i.e. it is not suited for investigations of excited states properties. This is translated in what is now known as the "band gap problem" among other things [110]. This corresponds to the underestimation of the band gap by about 40-50% in general when dealing with semiconductors or insulators. A heuristic explanation of this issue is developed in Ref. [110] and is summarised in what follows. The band gap of the K-S band structure can be expressed as:

$$E_g^{\mathrm{KS}} = \varepsilon_{\mathrm{CBM}} - \varepsilon_{\mathrm{VBM}} \tag{A.14}$$

where 'CBM' and 'VBM' denotes respectively the minimum of the conduction band, i.e. lowest unoccupied K-S state, and the maximum of the valence band, i.e. highest occupied K-S state. These two specific K-S eigenvalues can be approximately expressed as the ionisation potential and the electron affinity:

$$E_{N+1} \simeq E_N + \varepsilon_{\mathrm{CBM}} \tag{A.15}$$

$$E_{N-1} \simeq E_N - \varepsilon_{\mathrm{VBM}} \tag{A.16}$$

The equality is not strictly correct due to the interaction between the added/removed electron

with the rest of them, which is neglected in KS-DFT. This implies:

$$E_g^{\mathrm{qp}} = (E_{N-1} - E_N) - (E_N - E_{N+1}) \simeq E_g^{\mathrm{KS}} \tag{A.17}$$

where 'qp' stands for 'quasi-particle' and indicates that this is the quantity actually measured in experiment. In the case of small molecules, E_N, etc can be directly measured, but in the case of solid materials, the density of electrons is so large that adding or removing one is insignificant. A direct measure cannot be made and KS-DFT is required. It can then be shown that:

$$\lim_{\Delta n \to 0} E_{\mathrm{g}}^{\mathrm{qp}} = E_{\mathrm{g}}^{\mathrm{KS}} + \Delta_{xc} \tag{A.18}$$

where

$$\Delta_{xc} = \lim_{\Delta n \to 0} V_{xc}[n + \Delta n] - V_{xc}[n - \Delta n] \tag{A.19}$$

Since, by LDA and other usual approximations of the exchange and correlation potential, V_{xc} is continuous, the constant Δ_{xc} is always zero and thus DFT-LDA cannot give the correct band gap. It is an inherent problem due to the K-S formulation.

The previous 'quasi-particle' denotation is physically meaningful. It stems from the fact that a new particle can represent the excited electron interacting with all others. This many-body interaction actually affects its dispersion relation and its lifetime, and a quasi-particle formalism is thus necessary to describe its new properties. This is another reason to use a different theory than DFT in order to take into account those interactions and to shift from an independent-particle to a many-body one [115].

Another issue arises when considering that the excited electron and hole can interact via Coulomb interaction to form a new quasi-particle called an exciton, i.e. an electron-hole pair excitation. Thus, optical excitation becomes a two-body problem, at least. By reducing the dimensions of a system, the constituting particles are "forced" to interact in a stronger manner.

The following sections present the main battle horse method that yields accurate quantitative results dealing with excited states properties.

A.2.1 Prerequisite: Green's functions

As a preliminary, it must be noted that quantum field theory is needed to be perfectly rigorous in what follows but it has been deemed unnecessary to go into the details. For practical reasons, the method of second quantization is adopted when dealing quantum mechanically with many-body situations [121]. This formalism describes the quantum many-body states by Fock states, i.e. a quantum state defined by a precise number of identical particles. This number is modified by the field operators Ψ and $\Psi^\dagger$, respectively the annihilation and creation operator, which removes or adds a quantum of the field, i.e. the wave function of matter [122]. Instead of the Schrödinger's picture, it is also appropriate to

work in the Heisenberg representation where the operators are time dependent [121]:

$$\Psi_H(\mathbf{x},t) = e^{iHt}\Psi_S(\mathbf{x})e^{-iHt} \tag{A.20}$$

This allows the introduction of the Green's functions, a must when dealing with many-particle systems. As said in Ref. [123], these functions are *thermodynamic averages of products of the field operators* taken at two different time:

$$\mathrm{G}\left(1,1'\right) = \left(1/\mathrm{i}\right)\left\langle \mathrm{T}\left(\Psi(1)\Psi^{\dagger}\left(1'\right)\right)\right\rangle \tag{A.21}$$

$$\mathrm{G}_2\left(12,1'2'\right) = \left(1/\mathrm{i}^2\right)\left\langle \mathrm{T}\left(\Psi(1)\Psi(2)\Psi^{\dagger}\left(2'\right)\Psi^{\dagger}\left(1'\right)\right)\right\rangle \tag{A.22}$$

where 1 and $1'$ substitute $\mathbf{r}_1 t_1$ and $\mathbf{r}_{1'} t_{1'}$ which indicate the same quantum of the field, i.e. particle, but at a different place and time. G is the one-particle Green's function while G_2 is the two-particle one. The Dyson time-ordering operation, T, has the following effect:

$$\mathrm{T}\left(\Psi(1)\Psi^{\dagger}\left(1'\right)\right) = \Psi(1)\Psi^{\dagger}\left(1'\right) \qquad \text{for } t_1 > t_{1'} \tag{A.23}$$

$$= \pm\Psi^{\dagger}\left(1'\right)\Psi(1) \qquad \text{for } t_1 < t_{1'} \tag{A.24}$$

where the positive sign in the second equation is in the case of bosons while the negative sign is in the case of fermions. The physical meaning of G is the following [123]: if $t_1 > t_{1'}$, the creation operator adds a single particle which then travels and is removed later at another point in space and time by the annihilation operator. The single-particle Green's function thus describes the propagation of a disturbance. The same train of thought can be followed if $t_1 < t_{1'}$ but this time the disturbance corresponds to the absence of a particle. In a similar manner, G_2 has the same physical interpretation but allows to address more complex combinations of addition and removal of particles since two of them are now involved.

Another necessary addition is the spectral function A [123]:

$$\mathrm{A}(\mathbf{p},\omega) = \mathrm{G}^{>}(\mathbf{p},\omega) \mp \mathrm{G}^{<}(\mathbf{p},\omega) \tag{A.25}$$

where the superscript of G indicates whether $t_1 \gtrless t_{1'}$ and where the Fourier transforms were taken. The positive sign is for fermions and the negative one for bosons. This convention (upper for bosons, lower for fermions) is always followed here. It can be shown that G has boundary conditions which can then be re-written as:

$$G^{>}(\mathbf{p},\omega) = [1 \pm \mathrm{f}(\omega)]\mathrm{A}(\mathbf{p},\omega) \tag{A.26}$$

$$G^{<}(\mathbf{p},\omega) = \mathrm{f}(\omega)\mathrm{A}(\mathbf{p},\omega) \tag{A.27}$$

where

$$f(\omega) = 1/\left[e^{\beta(\omega-\mu)} \mp 1\right] \tag{A.28}$$

which corresponds to the average occupation number of a mode with energy ω in the grand-

canonical ensemble. Thus, the spectral function corresponds to spectral weights associated to each energy [123].

In an orbital basis representation, the amplitude of occupation of an orbital by a particle at a time t if one was created into this same orbital at a prior time is given by the diagonal elements of G. This prevents the need to deal with the full many-electron excited-state wavefunction [115].

A.2.2 The spectral function

From Ref. [121], it can be proven that, in the orbital basis representation, the poles of G are directly linked to the quasi-particle energies and lifetime by their position on the complex frequency plane.

It can be shown that the matrix elements of the spectral function can be expressed as:

$$A_{ii}(\omega) \;=\; \frac{1}{2\pi}\left|\mathrm{Im}\int \mathrm{d}x\;\mathrm{d}x'\varphi_i^*(x)\varphi_i\left(x'\right)G\left(x,x';\omega\right)\right| \tag{A.29}$$

where φ_i is the single-particle orbital [120]. By linking this equation to the statement of the previous paragraph, it comes that the spectral function can be used to derive the quasi-particles properties: the position and width of its peak are respectively the energy and the inverse of the lifetime of the quasi-particle [115].

If a non-interacting system is considered, then $A(\omega)$ does not display all the possible excitations since the matrix elements are taken between Slater determinants and not between superposition of them. On the contrary, an interacting system with the same number of electrons displays more peaks and each matrix element has several of them. The number of peaks increases with the number of electrons in the system. This leads to a continuous energy spectrum of each matrix element when real, i.e. extended, systems are considered. By comparing the interacting and the non-interacting spectrum, one can still identify a main peak. The latter dominates the others and is precisely the peak of the quasi-particle, although it has undergone a shift and a broadening due to the interaction. As mentioned before, this last effect relates to the finite lifetime of the quasi-particle and its possible decay into other excited states. Those are present in the spectrum as satellite peaks. They emphasise the coupling of the excitations in an interacting system. Figure A.3 illustrates a typical matrix element of a spectral function as described in this paragraph. On another note, the band structure of an extended system can be obtained by collecting the position of the qp peaks as a function of the momentum quantum number [120]. An example of such a band structure will be shown later.

The main issue that must now be addressed is how to calculate the Green's function. Since the many-body wavefunction is unknown, the definition of G (Eq. (A.22)) is not helpful for this matter.

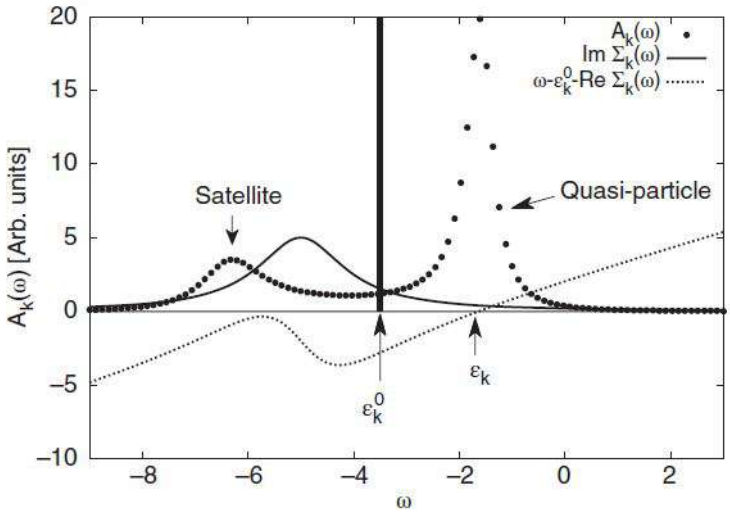


Figure A.3: Typical matrix element of a spectral function in an interacting system (fat dots), quasi-particle peak of a non-interacting system (vertical line). Taken from Ref. [120].

A.2.3 The GW approximation (GWA)

As shown in Refs. [121] and [124], the equation of motion of the field operators in the Heisenberg picture can be re-expressed to an equation of motion of G by using its definition. This equation can then be solved for a particular form of G as a function of $\psi_{n\mathbf{k}}(\mathbf{r})$ and $E_{n\mathbf{k}}$ which are the solutions to an equation called the quasiparticle Dyson equation [115]:

$$\left[E_{n\mathbf{k}} - H_0(\mathbf{r}) - V_{\mathrm{H}}(\mathbf{r})\right]\psi_{n\mathbf{k}}(r) - \int \Sigma\left(\mathbf{r},\mathbf{r}';E_{n\mathbf{k}}\right)\psi_{n\mathbf{k}}\left(\mathbf{r}'\right)\mathrm{d}\mathbf{r}' \;=\; 0 \tag{A.30}$$

where n and $\mathbf{k}$ are the band and wavevector quantum number of the electronic states. Eq. (A.30) closely resembles the K-S equations (Eq. (A.6)) but with a self-energy operator, Σ, which is non-local, non-Hermitian and energy-dependent. It contains the exchange-correlation potential that modifies the excited electron, i.e. the qp, energy due to the electronic environment [109]. This term is thus due to the electron-electron interaction and $E_{n\mathbf{k}}$ is complex as a consequence. The latter can be shown to be linked to the spectral function such that its real part is the qp energy and its imaginary one relates to the lifetime of the qp [115].

The problem of finding the poles of G has thus been shifted to the resolution of Eq. (A.30). For this purpose, it remains to evaluate or approximate the unknown self-energy operator.

Without going into the details, an expansion of Σ in terms of G and an operator, called the screened Coulomb operator W, can be realised [125, 124]. By only considering the first term of this expansion, one obtains the so-called "GW approximation" (GWA) [121]:

$$\Sigma\left(\mathbf{r},\mathbf{r}';E\right) \;=\; i\int \frac{\mathrm{d}\omega}{2\pi}\mathrm{e}^{-\mathrm{i}\delta\omega}G\left(\mathbf{r},\mathbf{r}';E-\omega\right)W\left(\mathbf{r},\mathbf{r}';\omega\right) \tag{A.31}$$

where

$$W(\mathbf{r},\mathbf{r}';\omega) = \int \varepsilon^{-1}(\mathbf{r},\mathbf{r}'';\omega)\, V_c(\mathbf{r}'',\mathbf{r}')\, d\mathbf{r}'' \tag{A.32}$$

with ε the time-ordered dielectric response function which dictates the charge response of the interacting system to a perturbation [120]. This function "screens" the bare Coulomb potential. This means that the GWA takes into account the dynamical and the position-dependent screening due to the other charges in the material [125].

Two quantities are thus important for the practical use of the GWA: ε and G. The latter can be approximated via calculation yielding the static limit ($\omega \to 0$) and extending it to finite frequencies while the former is first obtained via the K-S wavefunctions and then updated when Eq. (A.30) is solved [124].

It is also interesting to notice that DFT is an approximation of GWA in the sense that the non-local, energy-dependent self-energy is replaced with a local, energy-independent potential [109]. This justifies the considerable improvement regarding the band gap of insulators and semiconductors as illustrated in Figure A.4. The proper development and limitations of

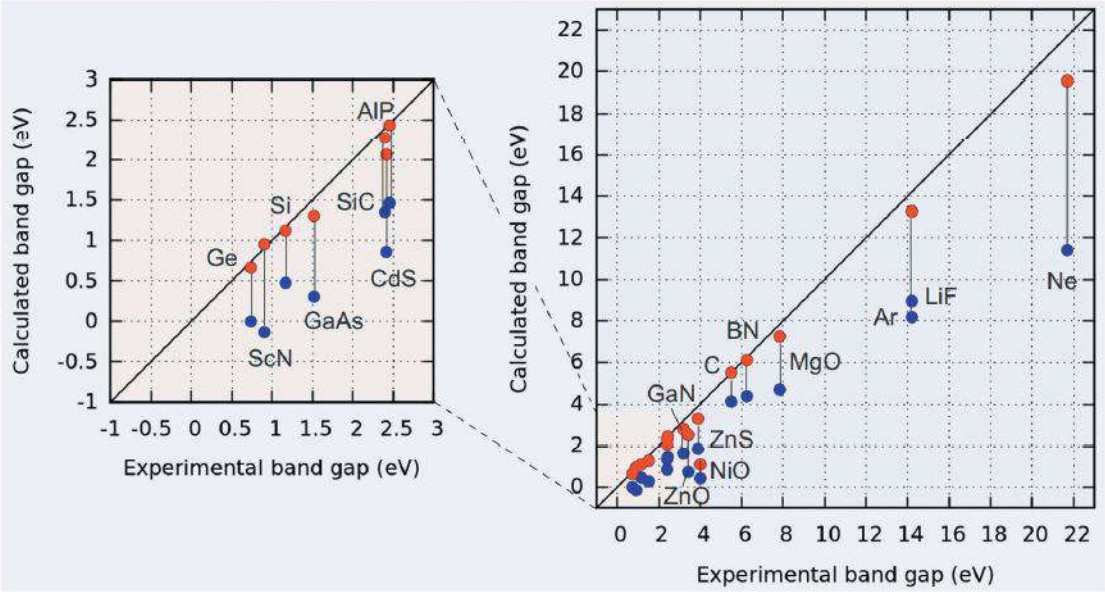


Figure A.4: Band gaps of several semiconductors and insulators from DFT-LDA and GWA. Figure Ref. from [110] and data from Ref. [126].

GWA is reviewed in Ref. [120] and the reader is invited to turn to the source materials for exhaustive information of what has been presented.

Despite the success of the GWA, the theory up to now still lacks the inclusion of the electron-hole interactions, i.e. the excitonic effects. Those are of great importance when dealing with nanotubes since the confinement due to the reduced dimensions strengthen these effects by naturally bringing the electron and the hole spatially closer [115].

A.2.4 Inclusion of the excitonic effects

The GWA can still be considered as a single-particle theory in the sense that its energies involves the addition or removal of a single electron. However, optical transitions seen as an interband transition implies adding both an electron and a hole to the ground-state. By ignoring the interaction between those two particles, one assumes that the excitation energy, i.e. $E_g = \varepsilon_{\mathrm{CBM}} - \varepsilon_{\mathrm{VBM}}$, is the sum of the excited electron and hole quasi-particle energies. This assumption does not hold anymore once the exciton, i.e. the quasi-particle composed of the electron and the hole via the Coulomb interaction, binding energy cannot be neglected against E_g. As mentioned, nanostructures favour low exciton binding energy due to the confinement helping the Coulomb interaction [109].

The way to include the excitonic effects is a natural continuation of the GWA for the single-particle excited states. In addition to the usual GW theory approximations and in the case of optical transitions, the interaction between excitons, i.e. between different electron-hole configurations, can be neglected [115]. From Ref. [127], the exciton wavefunction, which describes spatially the electron-hole configuration, can be written as:

$$\chi_S(\mathbf{r}, \mathbf{r}') = \sum_{cv} A^S_{cv} \psi_c(\mathbf{r}) \psi^*_v(\mathbf{r}') \tag{A.33}$$

where ψ corresponds to the wavefunction of the (single-particle) quasi-particle and S to the label of the exciton, i.e. the excited state. The coefficients A^S_{cv} as well as the energy Ω_S of the exciton can be found via the "Bethe-Salpeter equation" (BSE), itself derived from the equation of motion of G_2:

$$(E_c - E_v) A^S_{cv} + \sum_{cv,c'v'} K_{cv,c'v'}(\Omega_S) A^S_{c'v'} = \Omega_S A^S_{cv} \tag{A.34}$$

where E refer to qp energies from GWA and K relates to the interaction between the two components of the exciton [127]. Physically, the latter coefficient contains a long-range screened Coulomb attraction and short-range exchange repulsion [109, 127]. Once the exciton wavefunction and energy are obtained by resolving Eq. (A.34), the imaginary part of the dielectric function can be computed and thus various optical properties and linear responses to EM fields can be investigated [115, 109].

Appendix B

Parameters

B.1 Relaxation

The relaxation of the atomic structure was realized based on the Abipy Flow available[1].

- ecut=42
- kptopt=1
- ngkpt=[10 10 10]
- shiftk=[0 0 0]
- tolvrs=1e-10
- diemac=12.0
- optcell=2
- ecutsm=0.5
- dilatmx=1.15
- ionmov=22
- ntime=100
- nstep=100

B.2 Bandgap VASP

B.2.1 Ground-state density

- ENCUT = 600
- EDIFF = 1E-10
- ISTART = 0
- ICHARG = 2
- NELMIN = 6
- NELM = 100
- PREC = Normal
- LWAVE = .True.
- ALGO = Fast
- ISMEAR = -5
- AMIX = 0.2
- BMIX = 0.00001
- AMIX_MAG = 0.8
- BMIX_MAG = 0.00001
- Automatic mesh 0
- Gamma 8 8 8 0 0 0

B.2.2 HSE bandgap

- ENCUT = 600
- EDIFF = 1E-6
- ISTART = 1
- ICHARG = 1
- NELMIN = 6
- NELM = 100
- PREC = Normal
- LWAVE = .True.
- ISMEAR = -5
- AMIX = 0.2
- BMIX = 0.00001
- AMIX_MAG = 0.8
- BMIX_MAG = 0.00001
- LHFCALC = .TRUE.

[1] https://abinit.github.io/abipy/flow_gallery/index.html ; Accessed on the 17th April 2022.

- HFSCREEN = 0.2
- ALGO = D
- LDIAG = .True.
- AEXX = 0.25
- TIME = 0.4

B.3 Electronic structure

B.3.1 Ground-state density

- nband=60
- ngkpt=[10, 10, 10]
- diemac=5.0
- nbdbuf=4
- shiftk=[0, 0, 0]
- prtden=1
- ecut=42
- tolvrs=1.0e-12
- nstep=500

B.3.2 Bandstructure

- tolwfr=1e-12
- ndivsm=10
- nb_extra=60

B.3.3 JDOS

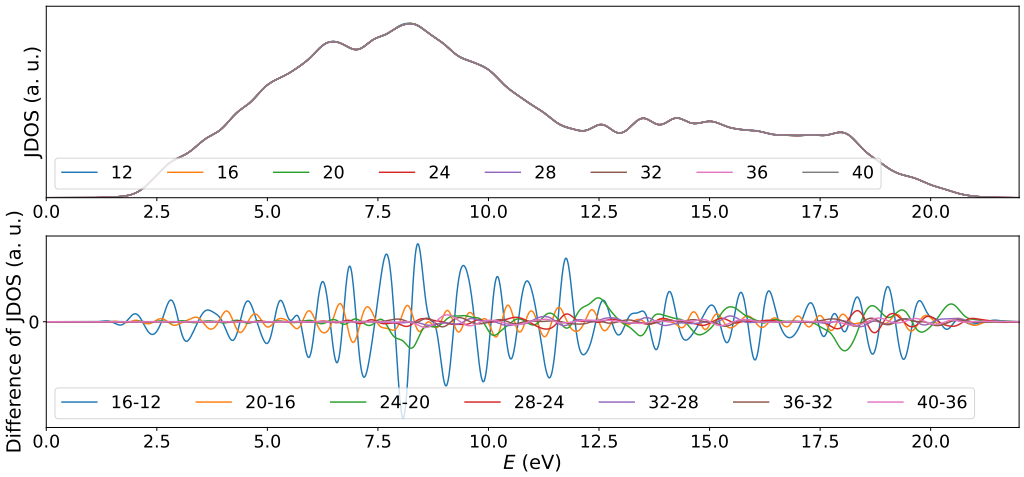
- tolwfr=1e-12
- ngkpt=[16 16 16]
- nscf_nband=60

B.3.4 PDOS

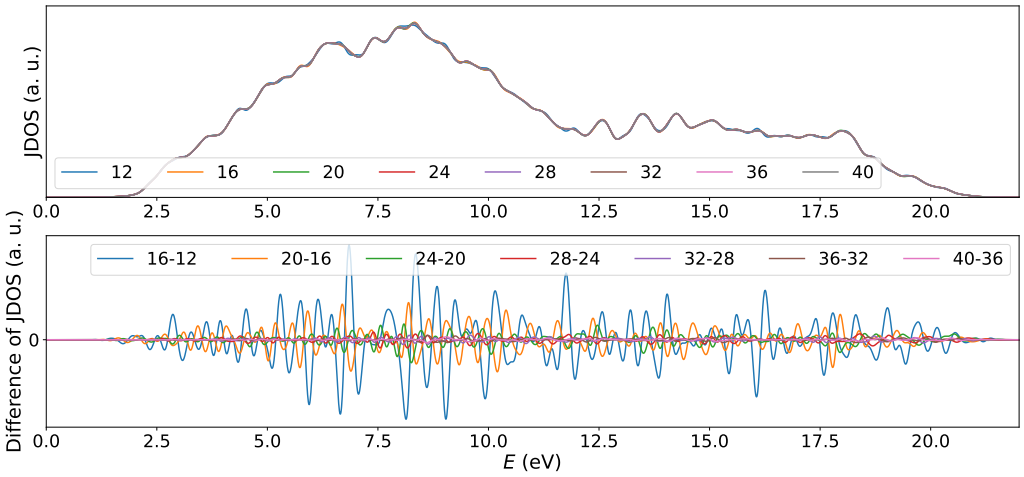
- prtdos 3
- natsph 8
- ratsph 2 2 2 2 2 2 2 2
- dosdeltae 0.00005
- iatsph 1 2 3 4 5 6 7 8

B.3.5 Convergence study JDOS

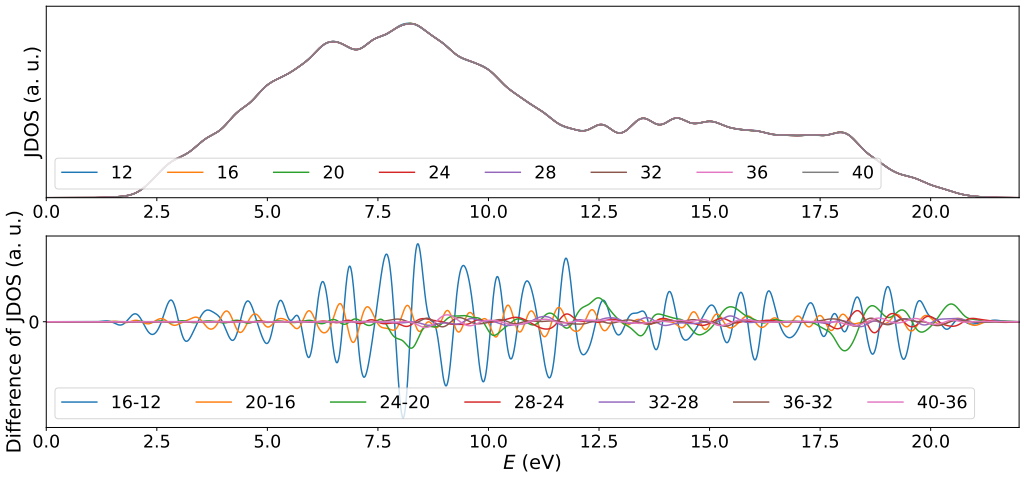
Figure B.1 shows the convergence study that determined the parameters used for the integration needed to get the JDOS. "Step" determines the energy mesh while "Width" corresponds to the width of the Gaussian of integration. From the looks of the different plots, a step of 0.025 and a width of 0.2 as well as an uniform k-point grid of 16×16×16 were arbitrarily decided as good compromise regarding both accuracy and numerical instabilities.



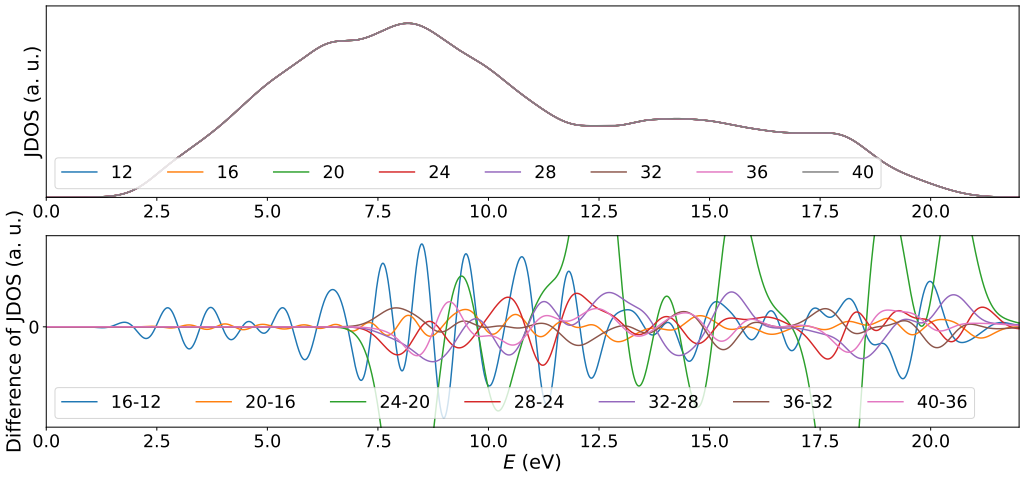
(a) Step = 0.025 ; Width = 0.2



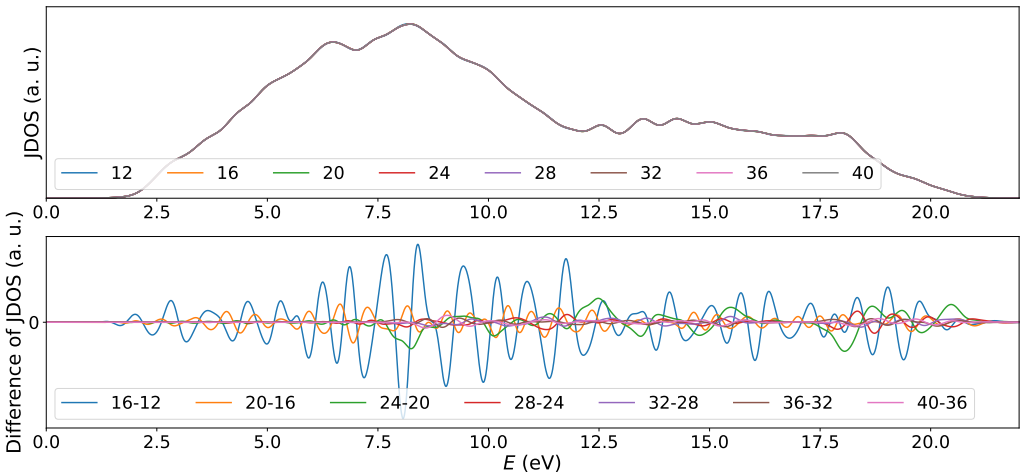
(b) Step = 0.025 ; Width = 0.1



(c) Step = 0.05 ; Width = 0.2



(d) Step = 0.025 ; Width = 0.4



(e) Step = 0.0125 ; Width = 0.2

Figure B.1: Convergence study of the JDOS with respect to an uniform k-point grid and by varying the energy step and the width (of the Gaussians) of the integration.

B.4 Optical properties

B.4.1 Ground-state density

- nband=54
- ecut=42
- ngkpt=[16 16 16]
- nshiftk=1
- shiftk=[0, 0, 0]
- nstep=500
- iscf=7
- diemac=5.0
- toldfe=1.0e-22
- nbdbuf=0
- kptopt=1
- paral_kgb=0

B.4.2 Bandstructure

- tolwfr=1e-12
- ndivsm=10
- nb_extra=64
- nstep=300
- nbdbuf=4

B.4.3 WFK in the irreducible Brillouin zone

- iscf=-2
- kptopt=1
- nband=64
- nbdbuf=4
- tolwfr=1e-20
- paral_kgb=0

B.4.4 WFK in the full Brillouin zone

- iscf=-2
- kptopt=3
- nband=64
- nbdbuf=4
- tolwfr=1e-20
- paral_kgb=0

B.4.5 DDK inputs

From Section B.4.4.

- kptopt=3
- nstep=1
- nline=0
- prtwf=1

B.4.6 Optic

- broadening=0.001
- domega=0.0003
- maxomega=1.0
- scissor=0.78*abu.eV_to_Ha
- tolerance=0.002
- num_lin_comp=3

- lin_comp=(11, 22, 33)
- num_nonlin_comp=1
- nonlin_comp=(123)
- nscf_node=nscf_work[1]
- ddk_nodes=ddks_work.ddk_tasks
- use_ddknc=False

B.5 Static NLO properties

B.5.1 DFPT

- nband=44
- ecut=ecut
- ngkpt=ngkpt
- nshiftk=1
- shiftk=[0, 0, 0]
- nstep=500
- iscf=7
- diemac=5.0
- toldfe=1.0e-22
- nbdbuf=0
- kptopt=1
- dfpt_sciss=0.7*abu.eV_to_Ha

B.5.2 Convergence study lattice parameters

k-point grid	10	12	14
a (Å)	5.868955023010328	5.86895116152383	5.8689511016134865
c (Å)	10.545267668788407	10.54527838199723	10.545277889266643

Table B.1: Convergence study for the uniform k-point grid to be used when relaxing the structure.

B.5.3 Convergence study NLO properties

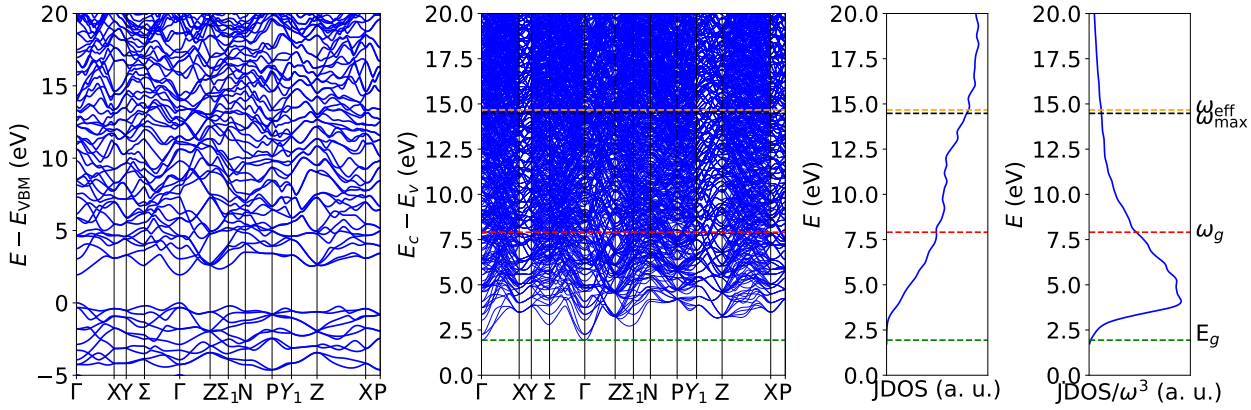
k-point grid	10	12	14	16
d_{36} (pm/V)	32.9141383	33.3372587	33.5983533	33.6971293

Table B.2: Convergence study for the uniform k-point grid to be used when computing the SHG response or related quantities. A threshold of 0.02% was arbitrarily chosen.

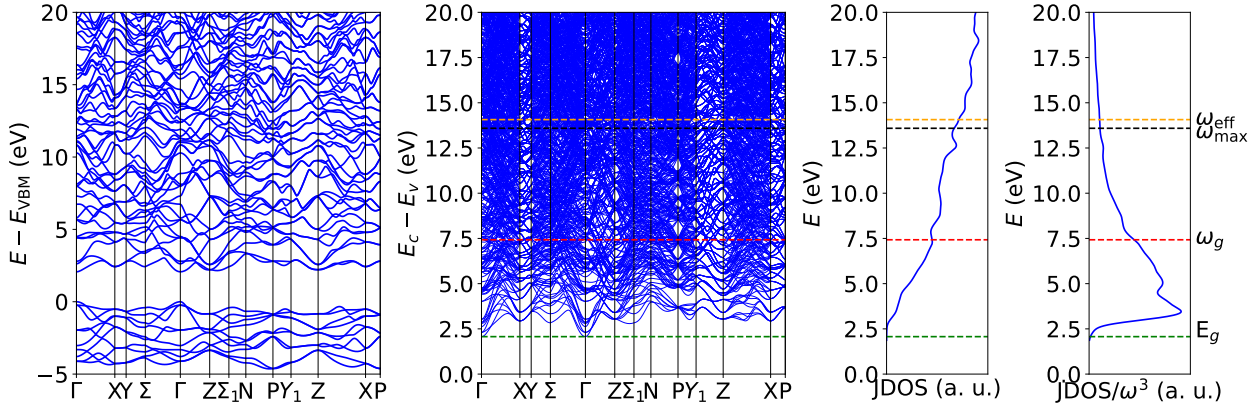
Appendix C

Mg-IV-V_2 compounds

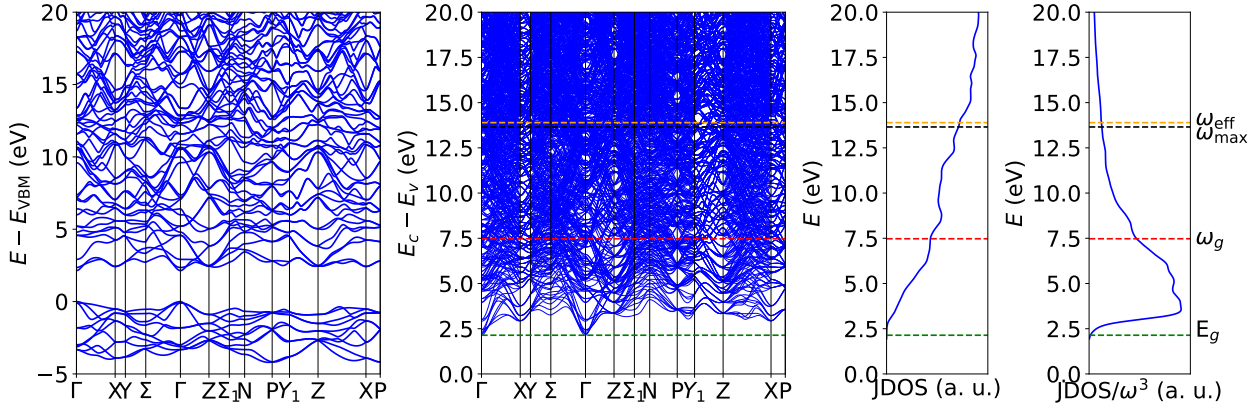
C.1 JDOS and optical coefficients



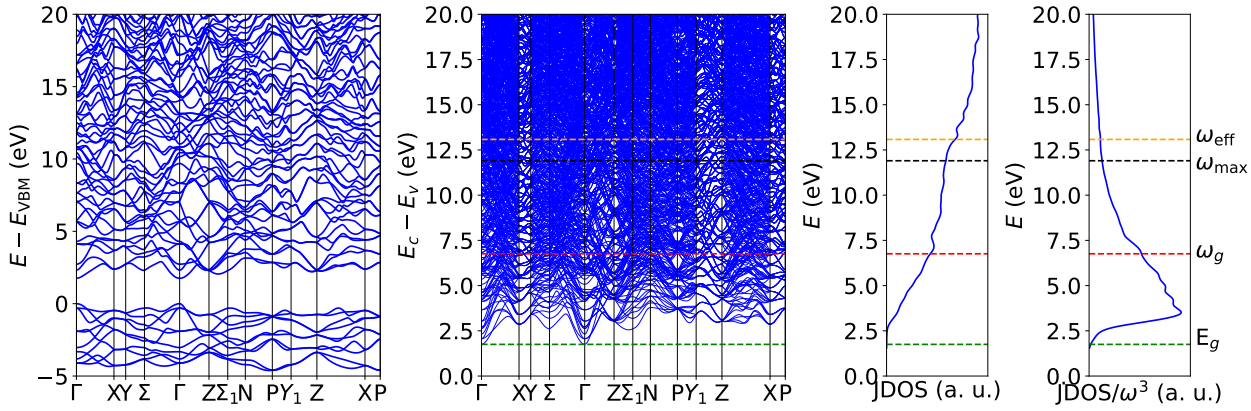
(a) $MgSiP_2$



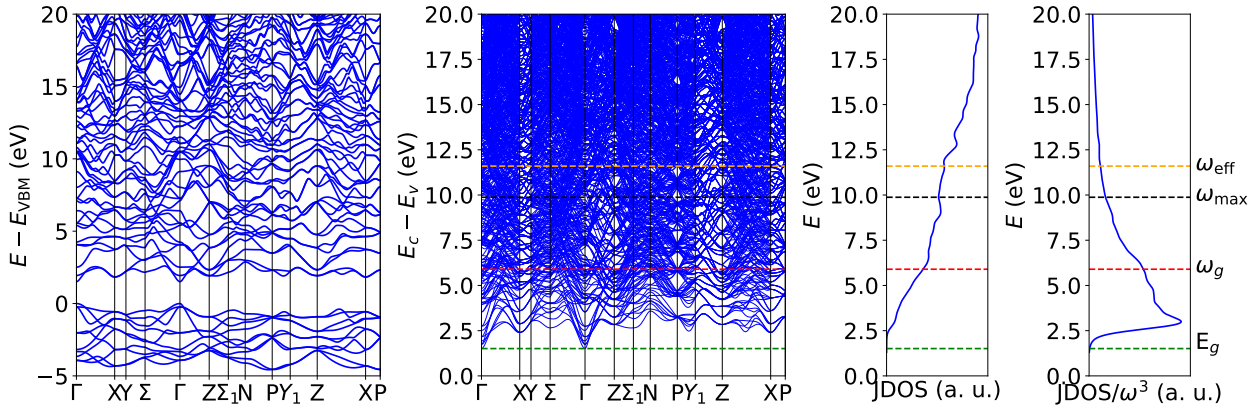
(b) $MgGeP_2$



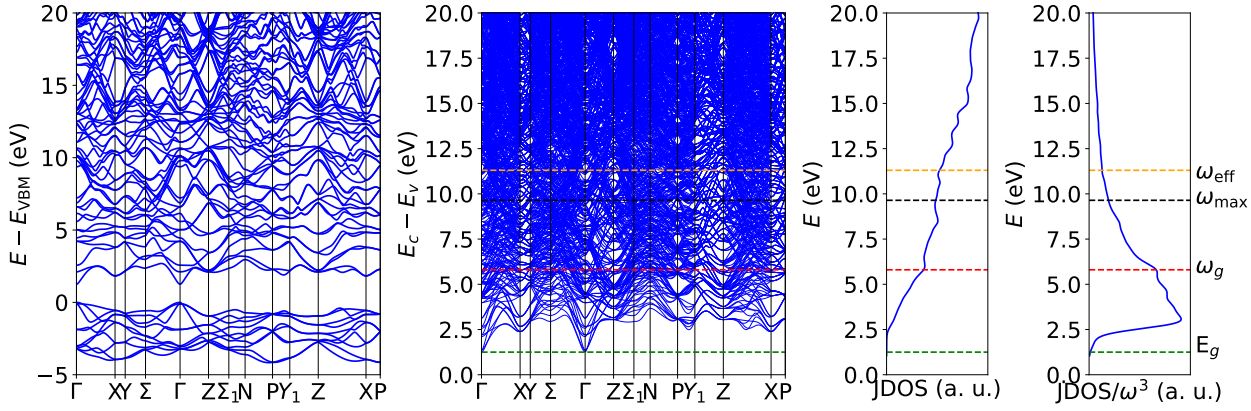
(c) $MgSnP_2$



(d) $MgSiAs_2$



(e) $MgGeAs_2$



(f) $MgSnAs_2$

Figure C.1: Bandstructure, joint density of states and optical coefficients of the six compounds of interest.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
École polytechnique de Louvain
Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/epl**