



REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE  
وزارة التعليم العالي و البحث العلمي  
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA  
RECHERCHESCIENTIFIQUE  
جامعة عمار تليجي بالاغواط  
UNIVERSITE AMAR TELIDJI LAGHOUAT  
كلية العلوم  
FACULTE DES SCIENCES  
قسم علوم المادة  
DEPARTEMENT SCIENCES DE LA MATIERE



## Mémoire master

Filière : Physique des Matériaux

Option : Physique des Matériaux

Présenté par M<sup>elle</sup> GUETTAF Aroua Rokaya

### THEME

## Etude datamining des matériaux pérovskites

Devant le jury :

<b>LEFKAIER Iben Khaldoun</b>	<b>Professeur</b>	Univ. de Laghouat	Président
<b>BENTRIA Bachir</b>	<b>Professeur</b>	Univ. de Laghouat	Examineur
<b>HEBBOUL Zoulikha</b>	<b>MCA</b>	Univ. de Laghouat	Examinatrice
<b>BENGHIA Ali</b>	<b>MCB</b>	Univ. de Laghouat	Encadreur
<b>FADLA Mohamed Abdelilah</b>	<b>Doctorant</b>	Univ. de Laghouat	Co-encadreur

Année universitaire 2020/2021

## **Remerciements**

*Je remercie ALLAH le tout-puissant de m'avoir donné le courage, la volonté et la patience de terminer ce travail.*

*Ce travail a été réalisé dans le Laboratoire de Physique des Matériaux de l'université de Laghouat. Aussi, j'aimerais témoigner ma gratitude et mes remerciements sincères au mon encadreur Mr.BENGHIA Ali et mon co-encadreur, Mr. FADLA Mohamed Abdelilah. J'exprime ma gratitude pour leur patience, leurs encouragements, leur aide, leur dévouement et leur disponibilité lors de la préparation de ce mémoire.*

*J'exprime mes sincères remerciements à monsieur Mr.LEFKAIER Iben Khaldoun pour avoir accepté de présider le jury de ce mémoire. Je remercie également tous les membres de jury, en l'occurrence Mr.BENTRIA Bachir et Mme. HEBBOUL Zoulikha d'avoir accepté d'examiner ce travail et de l'enrichir par leurs remarques et suggestions. À tous les enseignants du département des sciences de la matière, qui ont contribué à mon formation. Enfin, je remercie tous ceux qui m'ont aidé et soutenu dans mon travail.*

*Je dédie ce travail*

*A ma chère mère et mon cher père et pour leur soutien et leurs sacrifices durant mes études et dans toute ma vie.*

*A toutes mes sœurs, A toute ma famille.*

*A tous mes amis, particulièrement : S.Namarek*

*B.Rokaya, S.Fatima Zahra et H.Mamma*

# Sommaire

Liste des figures .....	3
Liste des tableaux.....	4
Introduction générale .....	5
Chapitre I Généralités sur les pérovskites.....	7
I.1 Introduction .....	8
I.2 Les matériaux pérovskites .....	8
I.2.1 Les pérovskites idéales.....	8
I.2.2 Conditions de stabilité structurale de pérovskite.....	9
I.2.2.1 Facteur de tolérance .....	9
I.2.2.2 Facteur d’octaèdre.....	10
I.2.3 Les propriétés et applications des pérovskites .....	10
I.2.4 Les limites des pérovskites dans le photovoltaïque .....	11
Chapitre II Méthodes de calcul.....	13
II.1 La Théorie de la Fonctionnelle de la Densité .....	14
II.1.1 Introduction .....	14
II.1.2 L’équation de Schrödinger .....	14
II.1.3 Les approximations .....	15
II.1.4 Théorie de la fonctionnelle de la densité (DFT) .....	15
II.1.4.1 Les théorèmes de Hohenberg et Kohn .....	15
II.1.4.2 L’approche de Kohn-Sham.....	16
II.1.4.3 Approximations d’échange-corrélation .....	17
II.1.4.3.1 L’approximation du gradient généralisé (GGA) .....	17
II.1.4.4 Approximation du pseudo potentiel.....	18
II.1.4.4.1 Pseudo-potentiel de type OTFG .....	19
II.1.4.5 Base des ondes planes.....	19
II.1.4.6 Procédure de calcul dans la DFT .....	20
II.1.5 Code de calcul : CASTEP .....	22
II.2 Datamining.....	23
II.2.1 Définition .....	23
II.2.2 Principe.....	23
II.2.3 Les méthodes du datamining.....	24

II.2.3.1	Les méthodes non- supervisées.....	24
II.2.3.2	Les méthodes supervisées.....	24
II.2.3.3	Les méthodes de réduction de dimensions : .....	24
II.2.4	Les techniques du datamining .....	24
II.2.4.1	Les techniques descriptives.....	25
II.2.4.2	Les techniques prédictives .....	25
II.2.5	Analyse en composant principales .....	25
II.2.5.1	Introduction.....	25
II.2.5.2	Définition .....	26
II.2.5.3	Principe de l’analyse en composantes principales (ACP).....	28
II.2.5.4	Formulation mathématique de l’ACP .....	28
II.2.5.5	Propriétés des composantes principales.....	29
Chapitre III Résultat et discussion .....		30
III.1	Introduction.....	31
III.2	Etude ab-initio.....	31
III.2.1	Paramètres de calcul .....	31
III.2.2	Propriétés structurales.....	31
III.3	Analyse par l'approche de datamining .....	33
III.3.1	Base de données.....	33
III.3.2	Outil d’analyse.....	37
III.3.3	Choix des axes .....	37
III.3.4	Matrice de corrélation.....	39
III.3.5	Les coordonnées des individus .....	40
III.3.6	Score plot.....	43
III.3.7	Les coordonnées des variables .....	44
III.3.8	Loading plot.....	44
III.3.9	Biplot des variables et des individus : .....	46
III.4	Conclusion .....	47
Conclusion et perspective .....		48
<i>References</i> .....		49

## Liste des figures

<b>Figure I.1</b> <i>Représentation de structure pérovskite idéale.</i> .....	9
<b>Figure II.1</b> <i>Illustration schématisant le potentiel de tout-électron (lignes discontinues) et pseudo-électron (lignes continues) et leurs fonctions d'ondes correspondantes</i> .....	19
<b>Figure II.2</b> <i>Schéma de cycle auto-cohérent (SCF) de la théorie de la fonctionnelle de la densité</i> .....	21
<b>Figure II.3</b> <i>Le processus de datamining</i> .....	23
<b>Figure II.4</b> <i>Tableau de donné pour l'ACP</i> .....	27
<b>Figure III.1</b> <i>Représentation graphique des axes principaux</i> .....	39
<b>Figure III.2</b> <i>Score plot des individus (<math>ABX_3</math>)</i> .....	43
<b>Figure III.3</b> <i>Loading plot</i> .....	45
<b>Figure III.4</b> <i>Graphe de « Biplot » superposant les individus et leurs propriétés.</i> .....	46

## Liste des tableaux

<b>Tableau I.1</b> <i>L'évolution des structures cristallines en fonction de la valeur du facteur de tolérance</i> .....	10
<b>Tableau III.1</b> <i>Les valeurs de paramètres optimisés des composés BaTiO<sub>3</sub>, ZnCNi<sub>3</sub>, AlCNi<sub>3</sub>, SrTiO<sub>3</sub></i> .....	32
<b>Tableau III.2</b> <i>Jeu des données des individus et leurs propriétés</i> .....	34
<b>Tableau III.3</b> <i>Les valeurs propres des axes principaux</i> .....	38
<b>Tableau III.4</b> <i>Matrice de corrélation</i> .....	39
<b>Tableau III.5</b> <i>Les coordonnées des individus</i> .....	40
<b>Tableau III.6</b> <i>Les coordonnées des variables</i> .....	44

## Introduction générale

Les matériaux ont de tout temps défini le niveau de développement de notre civilisation. Aujourd'hui, ils jouent un rôle déterminant dans toutes les mutations technologiques. Depuis de nombreuses années, les matériaux à structure pérovskite suscitent l'intérêt de la communauté scientifique. En effet, en raison de la très grande variété de compositions chimiques qu'elles proposent, les structures pérovskite présentent une large gamme de propriétés physico-chimiques. Suivant la nature des cations présents au sein de la structure, celle-ci s'écarte de la symétrie cubique idéale par distorsion de la maille. Toutes ces variétés de compositions chimiques et de symétries structurales, font des structures pérovskites d'excellents matériaux multifonctionnels, applicables dans des domaines aussi variés que les cellules solaires, catalyseurs, ...[1]

En raison de l'intérêt croissant pour ces composés, et de leur utilisation extensive dans la technologie existante et les applications potentielles possibles dans les dispositifs futuristes de haute technologie, l'étude des propriétés structurales, électroniques et optiques de ces pérovskites est fortement utilisée. La théorie de la fonctionnelle de la densité est actuellement la théorie la plus utilisée en physique des matériaux pour étudier et prédire les propriétés des matériaux. Elle permet d'obtenir une large gamme d'informations sur l'origine microscopique des propriétés de ces matériaux. On peut utiliser toutes ces informations pour faire des analyses, des explorations et voir des comportements corrélés de ces grands nombres d'informations. C'est l'idée de datamining.

Forage de données, explorations de données ou fouilles de données, ce sont les traductions possibles du datamining en Français. En règle générale, le terme datamining désigne l'analyse de données depuis différentes perspectives et le fait de transformer ces données en informations utiles, en établissant des relations entre les données ou en repérant des modèles [2]. Dans ce mémoire nous avons choisi l'une des méthodes de datamining qui est l'analyse en composante principale ACP.

L'objectif de ce travail est de générer une base de données, de l'analyser et de découvrir des corrélations entre elles afin de pouvoir développer un modèle prédictif dans le futur. Le travail développé dans ce mémoire a été réparti sur trois chapitres. Le premier chapitre est dédié à une brève revue bibliographique dans laquelle sont exposés tour à tour les composés

pérovskites, leurs principaux critères de stabilité, leurs propriétés et applications et aussi les principaux obstacles qui limitent leur commercialisation. Le deuxième chapitre présente une idée générale concernant la méthode ab-initio, les méthodes datamining et l'analyse en composantes principale ACP. Dans le troisième chapitre, nous exposons les principaux résultats des calculs ab-initio et les analyses élaborées par datamining. Enfin, une conclusion générale récapitule les principaux résultats obtenus et propose des perspectives de travaux futurs.

# *Chapitre I Généralités sur les pérovskites*

---

## **I.1 Introduction**

Le terme pérovskite est utilisé pour désigner une classe de solides cristallins inorganiques de formule  $ABX_3$ , où les éléments A et B sont des cations et X est un anion. La structure pérovskite occupe la position de premier choix pour tous les systèmes ternaires en termes de composition chimique de type  $ABX_3$ , et ce n'est pas à cause de leur grande abondance mais aussi à cause des propriétés utiles et intéressantes qui leur sont associées concernant cette structure. Le comportement cristallographique des pérovskites est d'un grand intérêt car la plupart des structures sont cubiques. Cependant, ils sont souvent légèrement déformés, ce qui donne des structures à plus faible symétrie telles que la structure orthorhombique, tétragonal, rhomboédrique et trigonal. [3]

Les pérovskites sont largement étudiés, non seulement en raison de leur structure, mais aussi parce qu'ils présentent de nombreuses propriétés intéressantes : électroniques, magnétique, électro-optique, électromécaniques et conductrices, qui sont la base de plusieurs applications existantes et potentielles.[3]

Dans ce premier chapitre nous présentons une généralité sur les pérovskites, leurs principaux critères de stabilité, leurs propriétés et applications et aussi les principaux obstacles qui limitent leur commercialisation.

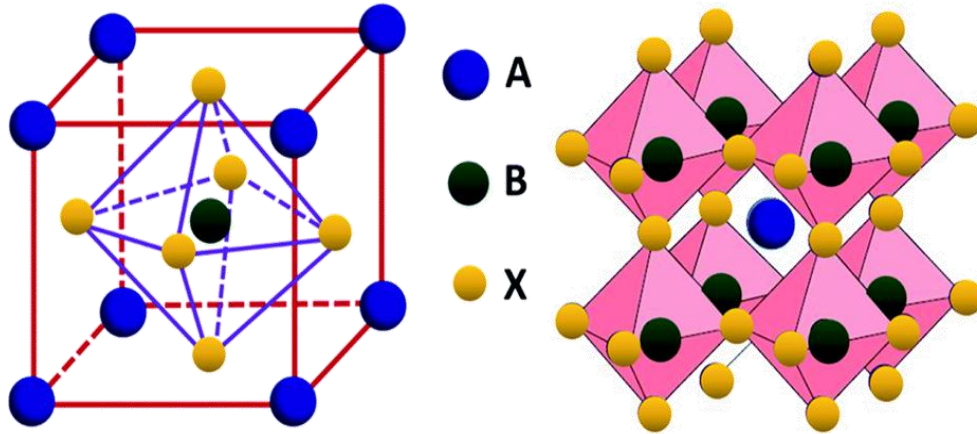
## **I.2 Les matériaux pérovskites**

La pérovskite désigne à l'origine le minéral  $CaTiO_3$ , décrit par Gustav Rose en 1839 et nommé ainsi en hommage au minéralogiste russe Lev Alexeïevitch Perovski. Mais aujourd'hui le terme pérovskite désigne un ensemble de composés possédant tous un même arrangement atomique  $ABX_3$ , où A étant le cation le plus gros, B le plus petit et X l'anion. Cet anion peut être oxyde, fluorure et, en quelques cas, chlorure, bromure, iodure, sulfure ou hydrure. La structure type pérovskite, occupe une place très majoritaire dans les systèmes ternaires connus sous la composition  $ABX_3$ , cela est dû non seulement à son occurrence large, mais également à une série de propriétés intéressantes et utiles liées à ce type structural.[4]

### **I.2.1 Les pérovskites idéales**

Les pérovskites idéales ont la formule générale  $ABX_3$ , où A est un alcalino-terreux ou une terre rare, B est un métal de transition et X est généralement l'oxygène, le fluor ou un halogène. [3]

La pérovskite idéale présente une maille cubique simple dans le groupe d'espace Pm-3m (groupe spatial n° 221 dans les tables cristallographiques internationale). Le motif cristallographique de la pérovskite est ABX<sub>3</sub> contient 5 atomes (Figure 1).



1. **Figure I.1** Représentation de structure pérovskite idéale.

En fonction du choix de l'origine, il y a deux façons de décrire la structure :

- **Dans la première :** A se trouve à l'origine, dans la position (0, 0, 0), B se trouve au centre du cube, dans la position (1/2, 1/2, 1/2) et les atomes X se trouvent au milieu de chaque face, dans la position (0, 1/2, 1/2)
- **Dans la deuxième façon :** l'origine est déplacée d'un vecteur (1/2, 1/2, 1/2), ce qui amène A à occuper la position (1/2, 1/2, 1/2), B la position (0, 0, 0) et les atomes X se trouvent au milieu de chaque arête, dans la position (0, 0, 1/2).

## I.2.2 Conditions de stabilité structurale de pérovskite

La stabilité de la structure pérovskite dépend essentiellement de deux facteurs : facteur de tolérance et facteur d'octaèdre.

### I.2.2.1 Facteur de tolérance

Goldschmidt en 1927[5] a défini un critère géométrique, appelé facteur de tolérance  $t$  ou de Goldschmidt, qui tient compte de la distance entre les ions pour définir les différentes structures dérivées de la structure pérovskite, suivant l'équation :

$$t = \frac{r_A + r_X}{\sqrt{2}(r_B + r_X)} \quad \text{ou} \quad t = \frac{1}{\sqrt{2}} \frac{d_{A-X}}{d_{B-X}}$$

où,  $r_A$ ,  $r_B$  et  $r_X$  représentent respectivement les rayons des ions A, B et X,  $d_{A-O}$  et  $d_{B-O}$  sont les distances cation- anion.

D'après ce critère, la structure cubique est observée pour  $t$  très proche de 1, les limites de stabilité de la phase pérovskite étant définies pour  $t$  compris entre 0.75 et 1.06.

J. B. Phillip et al. [6] ont résumés l'évolution des structures cristallines en fonction de la valeur du facteur de tolérance  $t$  comme suit :

**Tableau I.1** L'évolution des structures cristallines en fonction de la valeur du facteur de tolérance

$t < 0.75$ Ilménite (Trigonale)	$0.75 < t < 1.06$ Pérovskite			$t > 1.06$ Hexagonal
	$0.75 < t < 0.95$ Distorsion Orthorhombique	$0.96 < t < 0.99$ Distorsion Rhomboédrique	$0.99 < t < 1.06$ Cubique	

### I.2.2.2 Facteur d'octaèdre

Le facteur d'octaèdre est un autre paramètre pour prédire la stabilité de la structure. Il combine le rayon ionique de site B et le rayon ionique de site X au sein d'un octaèdre dans l'équation suivante :

$$\mu = \frac{r_B}{r_X}$$

Selon les observations empiriques, un composé de formule chimique  $ABX_3$  peut former une structure de pérovskite si le facteur de tolérance se situe entre 0,75 et 1,06 et si le facteur d'octaèdre est supérieur à 0,4. [7]

### I.2.3 Les propriétés et applications des pérovskites

Les pérovskites, un véritable coffre au trésor pour la science des matériaux. Ces matériaux céramiques avec leur structure cristalline particulière présentent une variété étonnante de propriétés électroniques et magnétiques dont on cite quelques une :

- ✓ **Supraconductivité** : Certains matériaux, une fois refroidis sous une température élevée spécifique, présentent une résistance électrique nulle et une expulsion de champs de flux magnétique, ces phénomènes appelés supraconductivité. Le type de structure oxyde pérovskites fournit un excellent cadre structurel en raison de l'existence de la supraconductivité. Les pérovskites qui contiennent du Cu agissent comme des supraconducteurs à haute température. Elles sont utilisées dans les condensateurs, les appareils à micro-ondes et l'électronique ultrarapide.[8]
- ✓ **Piézoélectricité** : Ces pérovskites transforment la pression mécanique ou la chaleur en électricité comme le titanate de baryum ( $\text{BaTiO}_3$ ), elles sont utilisées dans les microphones, circuit d'allumage et capteurs sous-marins.
- ✓ **Catalyseurs** : Elles accélèrent les réactions chimiques utilisées comme cathode dans les piles à combustible. On peut citer quelques matériaux comme  $\text{LaCoO}_3$ ,  $\text{LaMnO}_3$ ,  $\text{BaCuO}_3$ . [9]
- ✓ **Magnétorésistance** : Les pérovskites changent soudainement leur résistance électrique lorsqu'elles sont placées dans un champ magnétique, par exemple  $\text{LaMnO}_3$  et  $\text{LaFeO}_3$ . Elles sont utilisées dans les bandes et les disques magnétiques.
- ✓ **Ferroélectricité** : La propriété selon laquelle un matériau possède une polarisation électrique à l'état spontané, polarisation qui peut être renversée par l'application d'un champ électrique extérieur comme le niobate de potassium  $\text{KNbO}_3$ . [1]
- ✓ **Photovoltaïque** : La pérovskite est l'un des matériaux prometteurs du moment et pourrait remplacer le silicium des panneaux solaires grâce à leur coefficient d'absorption élevé, grande mobilité de porteurs de charges, une longueur de diffusion de porteurs élevée et rendement important [10]. Les cellules à base de Pérovskite ont montré une très grande évolution de leur rendement qui est passé de 2,2% à 26% depuis 2009, montrant une perspective intéressante pour leur commercialisation [11]

#### **I.2.4 Les limites des pérovskites dans le photovoltaïque**

Le plus gros problème avec les pérovskites aujourd'hui est l'instabilité à long terme. Il a été démontré que cette instabilité est due à des voies de dégradation impliquant des facteurs externes (tels que l'eau, la lumière et l'oxygène), mais également à une instabilité inhérente due à la nature du matériau, telle que la dégradation thermique. De plus, le matériau de cellule solaire pérovskite le plus efficace contient du plomb, qui est une matière toxique qui provoque

des effets néfastes. Plusieurs stratégies pour améliorer la stabilité ont été proposées, et la plus réussie est de changer le choix des composants. L'utilisation de systèmes de cations mixtes (par exemple, comprenant des cations inorganiques tels que le rubidium ou le césium) peut améliorer la stabilité et l'efficacité.[12]

# *Chapitre II Méthodes de calcul*

---

## II.1 La Théorie de la Fonctionnelle de la Densité

### II.1.1 Introduction

Les calculs ab-initio sont basés sur la résolution de l'équation de Schrödinger. Ces méthodes permettent de déterminer les grandeurs physiques et chimiques d'un système telles que sa structure électronique. La résolution de cette équation pour un système multi particule étant impossible numériquement, sa simplification en un système d'équations mono-particule est relativement plus aisée à résoudre, notamment grâce à quelques approximations. Dans cette partie, les principales méthodes de résolution de ces équations seront brièvement exposées.

### II.1.2 L'équation de Schrödinger

Considérons un système cristallin constitué par  $n$  électrons et  $N$  noyaux atomiques. L'état stationnaire de ces particules est décrit par l'équation de Schrödinger [13] :

$$H\psi(\{\vec{R}_I\}, \{\vec{r}_i\}) = E\psi(\{\vec{R}_I\}, \{\vec{r}_i\}) \quad (1)$$

$\Psi$  représente la fonction d'onde de toutes les coordonnées et  $E$  est la valeur numérique de l'énergie totale du système.

$\{\vec{R}_I\} = \vec{R}_1, \vec{R}_2, \dots \dots \dots \vec{R}_N$  représente l'ensemble des coordonnées des noyaux et  $\{\vec{r}_i\} = \vec{r}_1, \vec{r}_2, \dots \dots \dots \vec{r}_N$  représente l'ensemble des coordonnées des électrons.

$H$  représente l'opérateur Hamiltonien total du système à plusieurs corps donne par la relation suivante :

$$H = T_e + T_N + V_{eN} + V_{ee} + V_{NN} \quad (2)$$

Les termes apparaissant dans l'Hamiltonien prennent la forme explicite suivante :

- $T_e = -\sum_i^{N_e} \frac{\hbar^2}{2m_e} \nabla_i^2$  est l'énergie cinétique des électrons.
- $T_N = -\sum_I^N \frac{\hbar^2}{2M_I} \nabla_I^2$  est l'énergie cinétique des noyaux.
- $V_{eN} = -\sum_{i=1}^{N_e} \sum_{I=1}^N \frac{e^2}{4\pi\epsilon_0} \frac{Z_I}{|r_i - R_I|}$  est l'énergie potentielle d'attraction noyaux-électrons.
- $V_{ee} = \sum_{i=1}^{N_e} \sum_{j>i}^{N_e} \frac{e^2}{4\pi\epsilon_0} \frac{1}{|r_i - r_j|}$  est l'énergie potentielle de répulsion entre les électrons.

- $V_{NN} = \sum_{I=1}^N \sum_{J>I}^N \frac{e^2}{4\pi\epsilon_0} \frac{Z_I Z_J}{|R_I - R_J|}$  est l'énergie potentielle d'interaction entre les noyaux.

Tell que  $\hbar$  est la constante de Planck ( $h$  divisée par  $2\pi$ ),  $m_e$  est la masse de l'électron,  $e$  est la charge de l'électron,  $M_I$  est la masse du noyau I,  $Z_I$  et  $Z_J$  sont les charges des noyaux I et J,  $|\vec{r}_i - \vec{r}_j|$  définissent les positions des électrons ( $i$ ) et ( $j$ ), respectivement et  $|\vec{R}_I - \vec{R}_J|$  définissent les positions des noyaux ( $I$ ) et ( $J$ ), respectivement et  $|\vec{r}_i - \vec{R}_I|$  définissent les positions électrons ( $i$ ) et noyaux ( $I$ ) :

### II.1.3 Les approximations

La première approximation faite pour résoudre l'équation (1) est l'approximation de Born et Oppenheimer [14], et cela en séparant la partie électronique de la partie nucléaire dans la fonction d'onde  $\psi$ . Hartree [15] vient compléter ensuite cette approximation. Elle repose sur l'hypothèse d'électron libre, où on ne tient pas compte des interactions entre électrons et les états de spin. Pour résoudre le précédent problème, Hartree et Fock [16] ont généralisé ce concept en montrant que le principe de Pauli est respecté si l'on écrit la fonction d'onde sous la forme d'un «déterminant de Slater ».

### II.1.4 Théorie de la fonctionnelle de la densité (DFT)

Toutes les approches précédentes sont basées sur la fonction d'onde électronique ce qui rend les calculs très lourd comme un axe de la résolution d'équation de Schrödinger. En 1964, Kohn et Hohenberg ont fournis une autre méthode qui s'appuie sur la densité électronique comme une variable fondamentale. Ils ont réactivé l'idée de Thomas [17] et Fermi [18] qui l'ont proposés à la fin des années 20. Ce concept est appelé la Théorie de la fonctionnelle de la densité (DFT : density functional theory).

#### II.1.4.1 Les théorèmes de Hohenberg et Kohn

La théorie de la fonctionnelle de densité se base sur deux théorèmes énoncés par Hohenberg et Kohn en 1964 [19] :

Théorème 1 : L'énergie totale de l'état fondamental E est une fonctionnelle unique universelle de la densité électronique  $\rho(\vec{r})$  pour un potentiel externe  $V_{ext}(\vec{r})$  donné.

$$E = E[\rho(\vec{r})] \quad (5)$$

L'énergie de l'état fondamental :

$$E[\rho(r)] = T_e[\rho(\vec{r})] + V_{ee}[\rho(\vec{r})] + V_{eN}[\rho(\vec{r})] \quad (6)$$

Théorème 2: Pour un potentiel  $V_{ext}$  particulier, l'énergie de l'état fondamental du système est la valeur qui minimise cette fonctionnelle; la densité  $\rho$  qui lui est associée correspond à la densité exacte  $\rho_0$  de l'état fondamental.

$$E[\rho_0] = \min E\rho(\vec{r}) \quad (7)$$

#### II.1.4.2 L'approche de Kohn-Sham

Les travaux de Kohn et Sham [20], publiés en 1965, viennent compléter les travaux ultérieurs sur la DFT. Kohn et Sham proposent de remplacer le système multi particules en interaction qui obéit à l'Hamiltonien de l'équation de Schrödinger par un système sans interactions plus facile à résoudre. L'approche de Kohn et Sham suppose que la densité électronique à l'état fondamental du système réel est égale à celui d'un autre système fictif de particules sans interactions. Ceci conduit à résoudre un ensemble d'équations pour des particules indépendantes comme celles de Hartree ou Hartree-Fock.

L'équation de Schrödinger à résoudre dans le cadre de l'approche de Kohn et Sham est de la forme :

$$\left[ -\frac{\nabla_i^2}{2e} + V_{eff}(\vec{r}) \right] |\varphi_i(\vec{r})\rangle = \varepsilon_i |\varphi_i(\vec{r})\rangle \quad i = 1, \dots, N_e \quad (8)$$

où le potentiel effectif est de la forme

$$V_{eff} = V_{ext} + \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' + V_{xc}$$

où  $V_{ext}$  est le potentiel d'échange et corrélation est donné par la fonctionnelle dérivée :

$$V_{xc}(\vec{r}) = \frac{\delta E_{xc}[\rho(\vec{r})]}{\delta \rho(\vec{r})} \quad (9)$$

où  $E_{xc}$  est l'énergie d'échange et corrélation.

La densité  $\rho(\vec{r})$  est donnée par une somme sur l'ensemble des orbitales occupées :

$$\rho(\vec{r}) = \sum_{i=1}^{N_e} |\varphi_i(\vec{r})|^2 \quad (10)$$

Les équations (8) correspondent aux équations de Kohn et Sham et doivent être résolues de façon auto-cohérente.

### II.1.4.3 Approximations d'échange-corrélation

La fonctionnelle d'échange et de corrélation dans les équations de Kohn et Sham n'est possible qu'en donnant une forme analytique à l'énergie d'échange et de corrélation. Les approximations les plus couramment utilisées: l'approximation de la densité locale (LDA) [21] et l'approximation du gradient généralisé (GGA) [22].

#### II.1.4.3.1 L'approximation du gradient généralisé (GGA)

La méthode GGA [22] consiste à introduire des termes en gradient dans l'expression de l'énergie d'échange et de corrélation, Pour améliorer certains nombres de problème de la LDA, on parle du gradient de la densité qui est permet de prendre en compte la variance de la densité au voisinage de chaque point. On peut écrit donc:

$$E_{xc}^{GGA} = [\rho(\vec{r}), \nabla\rho(\vec{r})] = \int \varepsilon_{xc}^{GGA} [\rho(\vec{r}), \nabla\rho(\vec{r})] \rho(\vec{r}) d\vec{r} \quad (11)$$

La GGA est donnée par différentes paramétrisations, parmi elles celles (GGA PBE) et (GGA PBEsol).

#### ❖ La fonctionnel PBE

La fonctionnelle de Perdew, Burke et Ernzerhof (PBE) [23] utilise l'approximation GGA. Elle est basée sur une approche analytique. Les paramètres de la fonctionnelle sont des constantes fondamentales.

La forme GGA pour l'énergie d'échange est simplement

$$E_{xc}^{GGA}[\rho] = \int d^3r e_x^{unif}(\rho(\vec{r})) F_{xc}(s(\vec{r})) \quad (12)$$

$e_x^{unif}$  est la densité d'énergie d'échange d'un gaz d'électrons uniforme.

$s(\vec{r})$  est le dimensionnel gradient de densité.

$F_{xc}(s(\vec{r}))$  est le facteur d'amélioration pour le GGA donné.

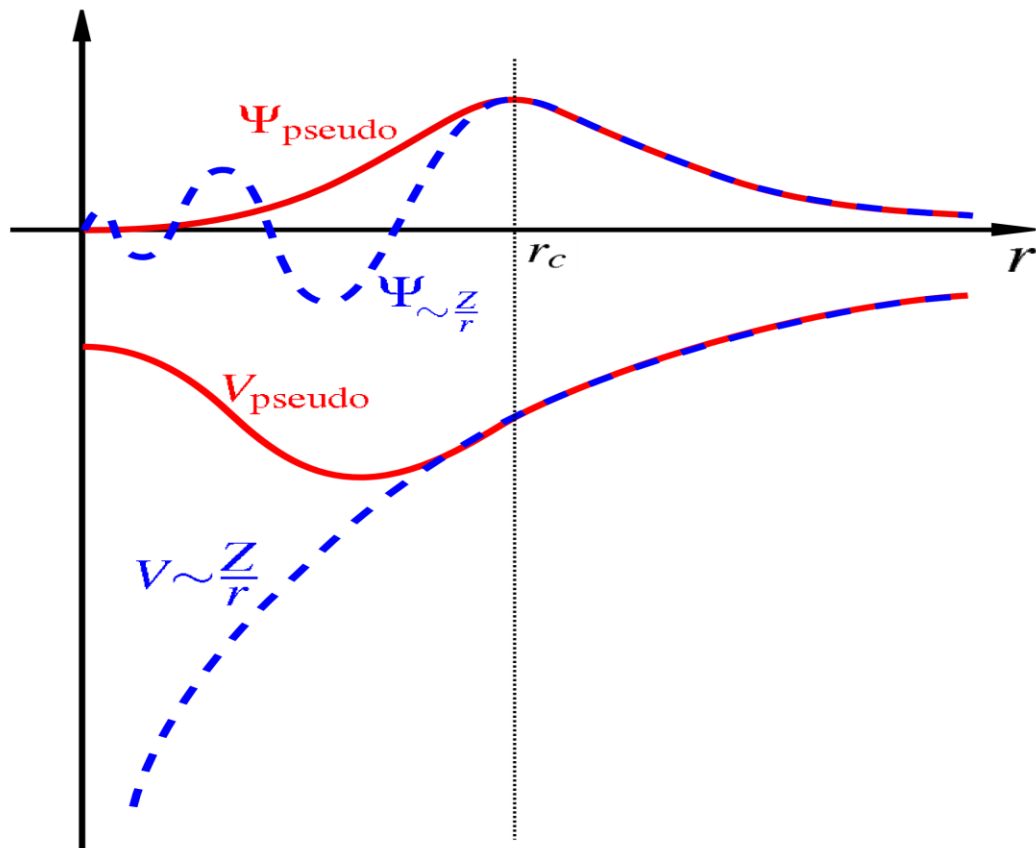
### ❖ La fonctionnel PBEsol

La fonctionnel PBEsol [24] construite à partir de PBE pour calculer l'énergie et les différentes propriétés des solides (paramètre de maille, Bulk modulus...). Elle est explicitement conçue pour améliorer l'expansion de gradient de premier principe pour l'échange sur une large gamme de gradients de densité, et donc pour améliorer la description des propriétés d'équilibre des solides densément tassés et de leurs surfaces.

$$E_{xc}^{GGA}[\rho] = \int d^3r e_x^{unif}(\rho(\vec{r})) F_{xc}(r_s(\vec{r}), s(\vec{r})) \quad (13)$$

#### II.1.4.4 Approximation du pseudo potentiel

Dans les molécules et les composées solides, les électrons de valence sont les seuls à intervenir dans les liaisons chimiques. Les électrons de cœurs, qui sont sur les couches les plus internes, proches du noyau, sont très peu sensibles à l'environnement. Ils sont difficiles à représenter sur une base d'ondes planes car ils possèdent généralement de fortes oscillations autour du noyau. On peut donc regrouper ces électrons du cœur avec les noyaux pour former des ions rigides dont les états électroniques resteront inchangés quel que soit l'environnement dans lequel l'atome sera placé [25]. On peut aller plus loin en remplaçant l'interaction des électrons de valence avec l'ensemble (noyaux et électrons de cœur) par un potentiel effectif, beaucoup moins attractif que le potentiel créé par le noyau avec tous les électrons. C'est ce potentiel effectif est appelé pseudo-potentiel. (figure 2)



**Figure II.1** Illustration schématisant le potentiel de tout-électron (lignes discontinues) et pseudo-électron (lignes continues) et leurs fonctions d'ondes correspondantes

#### II.1.4.4.1 Pseudo-potential de type OTFG

Le code serveur CASTEP de matériaux studio vous permet de générer des pseudo-potentiels à la volée, c'est-à-dire que vous pouvez fournir les paramètres qui régissent la génération plutôt qu'un fichier tabulé de la base de données. Cette approche présente un certain nombre d'avantages; par exemple, la même fonctionnelle de corrélation d'échange est utilisée dans les calculs à l'état atomique et solide. Il est possible de générer des potentiels "plus doux" ou "plus durs" en changeant le rayon du noyau. Il est possible d'étudier des configurations excitées avec un trou central, etc....

#### II.1.4.5 Base des ondes planes

L'ensemble des ondes planes est une base de projection des fonctions d'ondes. On les retrouve dans de nombreux codes: CASTEP, Abinit, VASP,... Dans le cadre de la théorie de la DFT, la base des ondes planes sont associées à des conditions aux limites périodiques.

La décomposition en ondes planes des fonctions d'onde consiste à exprimer ces dernières en séries de Fourier :

$$\varphi_k^n(r) = \frac{1}{\sqrt{\Omega}} \sum_k C_{\vec{k}}^{n,\vec{k}} e^{i(\vec{k}+\vec{K})\vec{r}} \quad (14)$$

où  $\vec{k}$  est le vecteur du réseau réciproque.

$\vec{K}$  est un vecteur d'onde de la première zone de Brillouin.

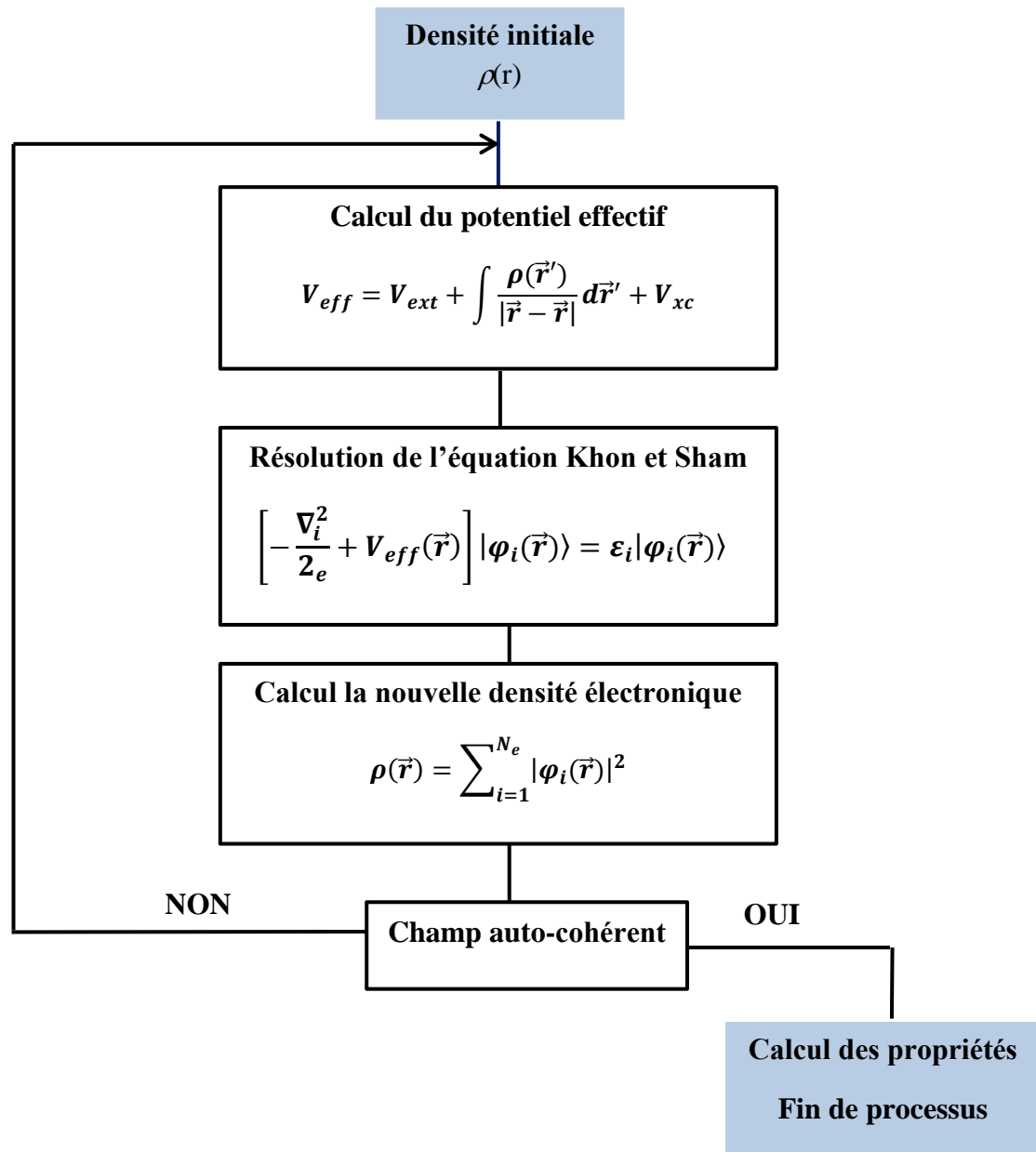
$C_{\vec{k}}^{n,\vec{k}}$  représente le coefficient du développement en ondes planes.

$\Omega$  est le volume du système.

#### II.1.4.6 Procédure de calcul dans la DFT

Les équations de Kohn et Sham sont résolues d'une manière itérative en utilisant un cycle auto cohérent qui est illustré dans la figure 3. Les étapes suivantes résument ce cycle :

1. Commencez par une densité électronique de test pour la première itération ;
2. Calculer la densité et le potentiel d'échange corrélation pour un point ;
3. Résoudre l'équation de Kohn-Sham ;
4. Calculer la nouvelle densité ;
5. Vérifier le critère de convergence (en comparant l'ancienne et la nouvelle densité) ;
6. En cas de convergence calculer les différents grandeurs physiques (Energie, ...) ;
7. Fin du calcul.



**Figure II.2.** Schéma de cycle auto-cohérent (SCF) de la théorie de la fonctionnelle de la densité

### **II.1.5 Code de calcul : CASTEP**

Le code CASTEP (Cambridge Serial Total Ennery Package Software) [26] est un programme de simulation très utilisé en physique du solide. Il s'agit d'un code de calcul ab-initio et il fait partie d'un ensemble de logiciels de simulation numériques nommé Materials Studio (MS) et commercialisés par Dassault systèmes Biovia ©. CASTEP est développé dans le groupe de la théorie de la matière condensée à l'Université de Cambridge, Royaume-Uni, c'est un programme qui emploie la théorie fonctionnelle de la densité (DFT) pour simuler les propriétés des solides, et peut prédire les propriétés, y compris les constantes élastiques, les propriétés structurales, les diagrammes de bande d'énergie, les densités d'états électroniques, les densités de charge et les propriétés optiques ainsi que les propriétés vibrationnelles et thermodynamiques.

## II.2 Datamining

### II.2.1 Définition

Le datamining que l'on peut traduire par "fouille de données" ou l'exploitation de données ou encore extraction de connaissances à partir des données (ECD) apparaît au milieu des années 1990 aux États-Unis comme une nouvelle discipline à l'interface de la statistique et des technologies de l'information : bases de données, intelligence artificielle, apprentissage automatique « machine Learning ». [27]

### II.2.2 Principe

Le datamining présente l'avantage de trouver des structures originales et des corrélations informelles entre les données. Il permet de mieux comprendre les liens entre des phénomènes en apparence distincts et d'anticiper des tendances encore peu discernables.

Le datamining est un processus d'extraction de connaissances en suivant les étapes principales suivantes :

- ❖ Formaliser un problème que l'organisation cherche à résoudre en terme de données.
- ❖ Préparer les données en vue des traitements et utilisations futures.
- ❖ Modéliser les données en leur appliquant des algorithmes d'analyses.
- ❖ Evaluer et valider les connaissances ainsi extraites des analyses.

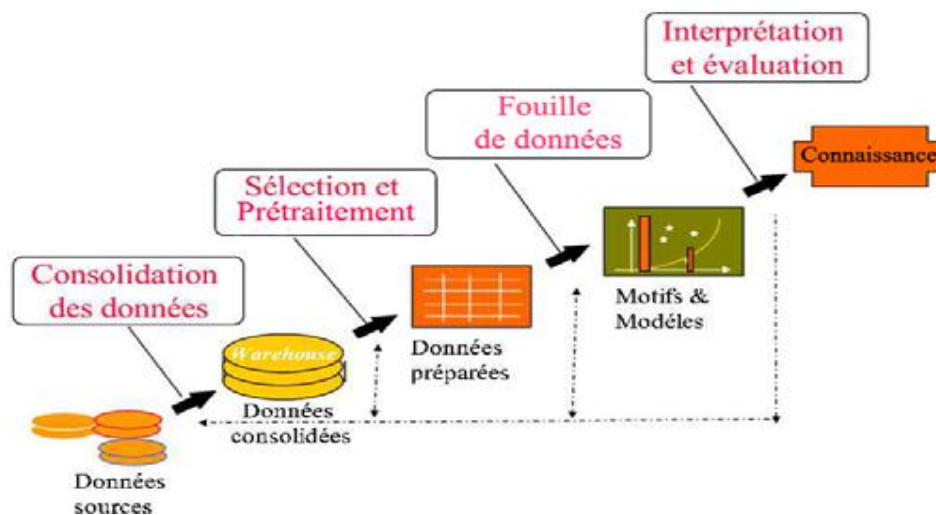


Figure II.3 Le processus de datamining [28]

### **II.2.3 Les méthodes du datamining**

Résoudre une problématique avec un processus de datamining impose généralement l'utilisation d'un grand nombre de méthodes différents. On peut distinguer 3 grandes familles :

- ❖ Les méthodes non- supervisées.
- ❖ Les méthodes supervisées.
- ❖ Les méthodes de réduction de données.

#### **II.2.3.1 Les méthodes non- supervisées**

Elles permettent de traiter un ensemble de données dans lequel aucune des données ou des variables à disposition n'a d'importance particulière par rapport aux autres, c'est-à-dire un ensemble de données dans lequel aucune variable n'est considérée individuellement comme la cible, l'objectif de l'analyse. On les utilise par exemple pour dégager d'un ensemble d'individus des groupes homogènes (typologie), pour obtenir une compression de l'information [29]

#### **II.2.3.2 Les méthodes supervisées**

Leur raison d'être est d'expliquer et/ou de prévoir un ou plusieurs phénomènes observables et effectivement mesurés. Concrètement, elles vont s'intéresser à une ou plusieurs variables de la base de données définies comme étant les cibles de l'analyse.[29]

#### **II.2.3.3 Les méthodes de réduction de dimensions :**

Elles permettent de réduire un ensemble de données volumineux à un ensemble de taille plus réduite, épuré ce que l'on considèra comme de l'information non pertinente ou non signifiante, comme du bruit. Elles sont ainsi très souvent, mais pas systématiquement, utilisées en amont des techniques supervisées ou non supervisées. Elles sont notamment très complémentaires des techniques non supervisées classiquement utilisées dans le domaine des statistiques.[29]

### **II.2.4 Les techniques du datamining**

Le Datamining propose différentes techniques à choisir en fonction de la nature des données et du type d'étude que l'on souhaite entreprendre. On peut distinguer deux grandes familles d'outils :

### **II.2.4.1 Les techniques descriptives**

Les techniques descriptives(ou exploratoires) visent à mettre en évidence des informations présentes mais cachées par le volume des données. On peut citer quelques techniques [30]:

- ❖ l'analyse en composantes principales ACP.
- ❖ analyse factorielle des correspondances AFC.
- ❖ analyse des correspondances multiples ACM, etc....

### **II.2.4.2 Les techniques prédictives**

Les techniques prédictives (explicatives) visent à extrapoler de nouvelles informations à partir des informations présentes, permettant ainsi de prédire de nouvelles conceptions : il y a une variable<<cible>> à prédire.

Voici les principales techniques disponibles [30]:

- ❖ Arbres de décision.
- ❖ Réseaux de neurones.
- ❖ Régression de moindre carré partiel PLS.
- ❖ Modèle linéaire général GLM.
- ❖ Analyse discriminante de Fisher.
- ❖ Régression logistique.

## **II.2.5 Analyse en composant principales**

### **II.2.5.1 Introduction**

L'ACP prend sa source dans un article de Karl Pearson publié en 1901. Le père du Test du  $\chi^2$  y prolonge ses travaux dans le domaine de la régression et des corrélations entre plusieurs variables. [31]

Pearson utilise ces corrélations non plus pour expliquer une variable à partir des autres (comme en régression), mais pour décrire et résumer l'information contenue dans ces variables. Encore connue sous le nom de transformée de Karhunen-Loève ou de transformée de Hotelling, l'ACP a été de nouveau développée et formalisée dans les années 1933 par Harold Hotelling [32]. La puissance mathématique de l'économiste et statisticien américain le conduira aussi à

développer l'analyse canonique, et la généralisation des analyses factorielles dont fait partie l'ACP.

Les champs d'application sont aujourd'hui multiples, allant de la biologie à la recherche économique et sociale, et plus récemment le traitement d'images. L'ACP est majoritairement utilisée pour :

- Décrire et visualiser des données ;
- Les décorréler ; la nouvelle base est constituée d'axes qui ne sont pas corrélés entre eux ;
- Les débruiter, en considérant que les axes que l'on décide d'oublier sont des axes bruités.

### **II.2.5.2 Définition**

L'Analyse en Composantes Principales (ACP) est une méthode de la famille de l'analyse des données et plus généralement de la statistique multi variée, qui consiste à changer des variables liées entre elles (dites "corrélées" en statistique) en nouvelles variables décorrélées les unes des autres. Ces nouvelles variables sont nommées "composantes principales", ou axes. Elle permet au praticien de réduire l'information en un nombre de composantes plus limité que le nombre initial de variables. Chaque composante principale est une combinaison linéaire des variables initiales. La mesure de la quantité d'informations représente sa variance. Les variances associées à chaque composante principale sont classées par ordre décroissant. La composante principale la plus informative est donc la première, et la moins informatives est la dernière [32].

Les données pour l'ACP sont présentées sous forme de tableau des données (figure 5). Ces données sont constituées d'individus et de variables qui doivent être quantitatives, continues et corrélées entrent-elles.

		Variables				
		1	.....	p	.....	P
Individus	1					
	.....					
	n			$x_n^p$		
	.....					
	N					

**Figure II.4.** *Tableau de donnés pour l'ACP*

L'ACP est utilisé pour minimiser la dimension (le nombre de variables) d'un problème. Cette diminution du nombre de variables doit s'effectuer en perdant un minimum d'informations. Le but de l'ACP est de condenser les données originelles en de nouveaux groupements, appelées nouvelles composantes. [33].

Pour effectuer une analyse en composantes principales, nous devons passer par les étapes suivantes:

- ❖ Normalisation des données pour être indépendants des paramètres P.
- ❖ Calcul d'une matrice de similarité C (bien souvent la corrélation).
- ❖ Recherche des éléments propres de C, qui donnent les axes principaux.
- ❖ Représentation des individus dans le nouvel espace (en ne considérant que les valeurs propres) [34].

L'ACP permet d'explorer les liaisons entre variables et les ressemblances entre individus.

### II.2.5.3 Principe de l'analyse en composantes principales (ACP)

L'idée de l'ACP est de déterminer un nouveau repère de  $R_p$  associé de manière naturelle à la structure du nuage considéré, de façon à pouvoir l'examiner plus commodément. Pour s'affranchir des effets d'échelle dus à l'hétérogénéité éventuelle des variables, ces dernières sont en général normalisées, c'est à dire que chaque colonne est divisée par son écart-type toutes sont dès lors exprimées dans la même échelle standard.

D'autre part, l'origine est placée au centre de gravité du nuage. C'est le nuage ainsi transformé qui est en fait considéré. L'utilisateur n'a cependant pas à se préoccuper de ces transformations préalables, sauf demande contraire elles sont exécutées automatiquement par les logiciels d'ACP.[35]

### II.2.5.4 Formulation mathématique de l'ACP

L'ACP est un simple changement de repère .sa grande force est de pouvoir ainsi recourir à l'algèbre linéaire comme outil mathématique principal, et de recevoir une interprétation géométrique simple.

Etant donné un ensemble d'observations décrites par des variables exclusivement numériques  $\{x_1, x_2, \dots, x_p\}$ , en plaçant en ligne les individus et en colonnes les variables. Soit X ce tableau de données.

$$X = \begin{matrix} & & & & 1 & \dots & p \\ \text{individu} & 1 & & & \left[ \begin{matrix} x_i^1 & \dots & x_i^p \\ \vdots & & \vdots \\ x_i^n & \dots & x_i^n \end{matrix} \right] \end{matrix}$$

Par convention, on placera en exposant ce qui se rapporte aux individus : le premier individu est donc  $x_i$ , et en indice ce qui se rapporte aux variables :  $x^j$  désigne la première variable.  $x_i^j$  désigne la valeur numérique prise par la première variable  $j$ , pour le premier individu  $i$ . Et notre tableau de données regroupe toutes les valeurs prises par tous les individus (de 1 à n) par p variables, l'ACP a pour objectif de décrire ce même ensemble de données par de nouvelles variables en nombre réduit. Ces nouvelles variables seront des combinaisons linéaires des variables originales, et porteront le nom de composantes principales (CP).

En général, la réduction du nombre de variables utilisées pour décrire un ensemble de données provoque une perte d'information. L'ACP procède de façon à ce que cette perte d'information soit la plus faible possible. L'ACP peut donc être vu comme une technique de réduction de la dimensionnalité. [36][37]

#### **II.2.5.5 Propriétés des composantes principales**

##### **❖ Nombre**

Bien que l'objectif soit en général de n'utiliser qu'un petit nombre de composantes principales, l'ACP en construit initialement  $p$ , autant que de variables originales. Ce n'est que par la suite que l'analyste décidera du nombre de composantes à retenir. "Retenir  $k$  composantes principales" veut dire "remplacer les observations originales par leur projections orthogonales dans le sous-espace à  $k$  dimensions défini par les  $k$  premières composantes principales."

##### **❖ Orthogonalité**

Les composantes principales définissent des directions de l'espace des observations qui sont deux à deux orthogonales. Autrement dit, l'ACP procède à un changement de repère orthogonal, les directions originales étant remplacées par les composantes principales.

##### **❖ Décorrélation**

Les composantes principales sont des variables qui s'avèrent être deux à deux décorrélées. [27]

# *Chapitre III Résultat et discussion*

---

### **III.1 Introduction**

La famille des pérovskites de formule générale  $ABX_3$  a suscité beaucoup d'intérêt. Ils présentent une variété de propriétés. L'Analyse en composante principale (ACP) une des techniques utilisées afin d'identifier les tendances des propriétés d'un grand nombre de matériaux.

Ce chapitre est divisé en deux parties. La première partie, est consacrée à l'étude des propriétés structurales de 120 pérovskites sélectionnées par DFT, et dans la deuxième partie nous avons effectué une analyse par l'ACP sur les propriétés calculés de ces matériaux.

### **III.2 Etude ab-initio**

A partir de l'article Boubchir et al.[38] , nous avons sélectionné 120 composés pour calculer leurs propriétés structurelles afin de créer notre propre base de données. Nous ne nous sommes pas appuyés sur les données Boubchir car elles ont été collectées à partir de la littérature et chaque propriété a été calculée différemment, ce qui affecte évidemment sur les résultats de notre analyse.

#### **III.2.1 Paramètres de calcul**

Dans cette partie, nous allons déterminer les propriétés structurelles des 120 pérovskites cubiques  $ABX_3$  pour avoir les paramètres du réseau, le module de compressibilité B et l'énergie totale, en utilisant le code CASTEP (24) dans le cadre de la théorie de la fonctionnelle de la densité (DFT) et la méthode des ondes planes comme base de projection de la fonction d'onde. Le potentiel d'échange et de corrélation est traité par approximation du gradient généralisé (GGA) de Perdew Burk-Ernzerhof pour les solides(PBEsol) (22). Nous avons choisi le pseudo-potential de type OTFG *norm conserving*.

#### **III.2.2 Propriétés structurales**

Avant de calculer les propriétés que l'on cherche, il faut tout d'abord optimiser la structure pour obtenir une structure plus stable qui correspond à une énergie totale minimale. Cela consiste à minimiser la pression exercée sur chaque atome du solide.

Les structures des pérovskites étudiés sont optimisées en variant les paramètres de maille, et les coordonnées des ions (x, y, z) tout en respectant la symétrie des systèmes. Les structures

optimisées représentent des mailles relaxées (force nulle sur les atomes donc état d'équilibre pour le système). Le code CASTEP exploite l'algorithme proposé par Broyden-Fletcher-Goldfarb-Shanno (BFGS) pour faire varier tous les paramètres structuraux ensemble et trouver la géométrie d'équilibre. Les structures sont considérées optimales lorsque les critères suivants sont remplis :

- Tolérance en énergie :  $0.5 \times 10^{-5}$  eV/atome ;
- Stress maximal : 0.02 GPa.

Les valeurs obtenues dans cette étude par une approximation du type GGA-PBEsol sont en accord avec les valeurs expérimentales disponibles. Le tableau 2 résume les valeurs des paramètres de maille pour les composés BaTiO<sub>3</sub>, ZnCNi<sub>3</sub>, AlCNi<sub>3</sub> et SrTiO<sub>3</sub>.

**Tableau III.1** Les valeurs de paramètres optimisés des composés BaTiO<sub>3</sub>, ZnCNi<sub>3</sub>, AlCNi<sub>3</sub>, SrTiO<sub>3</sub>

		BaTiO <sub>3</sub>	ZnCNi <sub>3</sub>	AlCNi <sub>3</sub>	SrTiO <sub>3</sub>
a (Å)	Nos calculs	3.97	3.72	3.73	3.89
	Exp	4.01 [39]	3.66 [40]	3.59 [41]	3.90 [42]
	Théo	3.95 [43]	3.72 [44]	3.77 [45]	3.86 [43]

Les résultats obtenus de l'optimisation (paramètres structuraux et module de compressibilité B) pour les matériaux restants sont présentés par la suite.

### III.3 Analyse par l'approche de datamining

Dans cette partie nous présentons les résultats liés à l'utilisation de l'analyse en composantes principales (ACP) sur les 120 pérovskites choisis.

Afin d'identifier les tendances des propriétés des matériaux, nous construisons une base de données pour les pérovskites, y compris le paramètre de maille ( $a$ ), le module de compressibilité ( $B$ ), le facteur de tolérance ( $t$ ), ainsi que l'énergie de cohésion ( $E_{coh}$ ), le facteur d'octaèdre ( $\mu$ ), et enfin les rayons ioniques des sites A, B, X ( $r_A, r_B, r_X$ ).

#### III.3.1 Base de données

La base de données contient 120 pérovskites, chacun avec 8 variables. Elle a été construite en se basant sur des valeurs obtenues à partir du calcul ab-initio en utilisant le code CASTEP.

Le tableau 3 présente les éléments sélectionnés et leurs propriétés. Il est à noter que l'énergie de cohésion  $E_{coh}$ , le facteur d'octaèdre  $\mu$  et le facteur de tolérance  $t$  ont été calculés selon les relations suivantes :

$$E_{coh} = E_{tot} - (E_A + E_B + 3E_X)$$

$$\mu = \frac{r_B}{r_X}$$

$$t = \frac{r_A + r_X}{\sqrt{2} (r_B + r_X)}$$

Sans oublier les rayons ioniques  $r_A$ ,  $r_B$  et  $r_X$  que nous avons recueillis à partir de littérature [46].

Nous n'utiliserons pas les valeurs des rayons de Shannon, car on peut obtenir une expression incorrecte pour la prédiction du paramètre de maille et du facteur de tolérance en fonction du rayon ionique [47]. De plus Le facteur de tolérance basé sur ces rayons ne permet pas de prédire avec précision la stabilité de certains composés pérovskites[48] [49].

**Tableau III.2** *Jeu des données des individus et leurs propriétés.*

N°	Individus	a (Å)	B(GPa)	r <sub>A</sub> (Å)	r <sub>B</sub> (Å)	r <sub>X</sub> (Å)	t	E <sub>coh</sub> (eV)	μ
1	AgCNi <sub>3</sub>	3.78	347.9	1.6	0.7	1.35	1.02	-36.809	0.52
2	AgNNi <sub>3</sub>	3.77	348.3	1.6	0.65	1.35	1.04	-36.393	0.48
3	AlCMn <sub>3</sub>	3.73	194.4	1.25	0.7	1.4	0.89	-45.487	0.50
4	AlCNi <sub>3</sub>	3.73	21.6	1.25	0.7	1.35	0.90	-38.939	0.52
5	AlCTi <sub>3</sub>	4.16	223.9	1.25	0.7	1.4	0.89	-38.419	0.50
6	AlNNi <sub>3</sub>	3.72	439.8	1.25	0.65	1.35	0.92	-38.498	0.48
7	AgNCu <sub>3</sub>	3.90	214.2	1.6	0.65	1.35	1.04	-23.825	0.48
8	AlCFe <sub>3</sub>	3.69	181.8	1.25	0.7	1.4	0.89	-44.832	0.50
9	AlCY <sub>3</sub>	4.83	65.1	1.25	0.7	1.85	0.86	-31.375	0.38
10	AlNFe <sub>3</sub>	3.67	435.2	1.25	0.65	1.4	0.91	-44.607	0.46
11	AsNBa <sub>3</sub>	5.38	49.1	1.15	0.65	2.15	0.83	-24.002	0.30
12	AsNCa <sub>3</sub>	4.71	52.3	1.15	0.65	1.8	0.85	-26.049	0.36
13	AsNMg <sub>3</sub>	4.21	97.7	1.15	0.65	1.5	0.87	-22.955	0.43
14	AsNSr <sub>3</sub>	5.03	37.3	1.15	0.65	2	0.84	-24.172	0.33
15	BaAmO <sub>3</sub>	4.28	194.6	2.15	1.75	0.6	0.83	-39.221	2.92
16	CdCNi <sub>3</sub>	3.81	453.3	1.55	0.7	1.35	1.00	-35.491	0.52
17	CdNNi <sub>3</sub>	3.80	374.7	1.55	0.65	1.35	1.03	-34.839	0.48
18	TcCNi <sub>3</sub>	3.75	419.3	1.35	0.7	1.35	0.93	-42.769	0.52
19	BaPaO <sub>3</sub>	4.42	500.0	2.15	1.8	0.6	0.81	-43.098	3.00
20	BaPrO <sub>3</sub>	4.38	416.3	2.15	1.85	0.6	0.79	-40.366	3.08
21	BaPuO <sub>3</sub>	4.30	845.7	2.15	1.75	0.6	0.83	-40.371	2.92
22	BaRuO <sub>3</sub>	4.00	359.2	2.15	1.3	0.6	1.02	-36.511	2.17
23	BaSnO <sub>3</sub>	4.12	167.5	2.15	1.45	0.6	0.95	-33.553	2.42
24	BaTaO <sub>3</sub>	4.08	330.9	2.15	1.45	0.6	0.95	-42.397	2.42
25	BaThO <sub>3</sub>	4.53	124.9	2.15	1.8	0.6	0.81	-42.178	3.00
26	BaTiO <sub>3</sub>	3.97	248.6	2.15	1.4	0.6	0.97	-41.268	2.33
27	BaTbO <sub>3</sub>	4.27	500.0	2.15	1.75	0.6	0.83	-36.617	2.92
28	BaUO <sub>3</sub>	4.36	138.4	2.15	1.75	0.6	0.83	-42.796	2.92
29	BaZrO <sub>3</sub>	4.18	8.8	2.15	1.55	0.6	0.90	-42.836	2.58
30	BiAlO <sub>3</sub>	3.77	330.5	1.6	1.25	0.6	0.84	-34.397	2.08
31	BiGaO <sub>3</sub>	3.85	656.7	1.6	1.3	0.6	0.82	-30.47	2.17
32	BiInO <sub>3</sub>	4.10	40.6	1.6	1.55	0.6	0.72	-27.988	2.58

33	BiNCa <sub>3</sub>	4.85	109.9	1.6	0.65	1.8	0.98	-24.852	0.36
34	BiNBa <sub>3</sub>	5.50	23.4	1.6	0.65	2.15	0.95	-23.339	0.30
35	BiScO <sub>3</sub>	4.02	148.9	1.6	1.6	0.6	0.71	-35.549	2.67
36	BiNSr <sub>3</sub>	5.17	25.7	1.6	0.65	2	0.96	-23.276	0.33
37	CaSnO <sub>3</sub>	4.01	186.9	1.8	1.45	0.6	0.83	-32.926	2.42
38	CaTaO <sub>3</sub>	3.98	350.9	1.8	1.45	0.6	0.83	-41.874	2.42
39	CaCrO <sub>3</sub>	3.72	682.1	1.8	1.4	0.6	0.85	-39.985	2.33
40	CaTiO <sub>3</sub>	3.84	104.6	1.8	1.4	0.6	0.85	-41.185	2.33
41	CaZrO <sub>3</sub>	4.09	185.5	1.8	1.55	0.6	0.79	-41.955	2.58
42	CaCNi <sub>3</sub>	3.90	594.8	1.8	0.7	1.35	1.09	-36.214	0.52
43	CdNCu <sub>3</sub>	3.97	208.1	1.55	0.65	1.35	1.03	-21.831	0.48
44	CoNFe <sub>3</sub>	3.61	932.1	1.35	0.65	1.4	0.95	-47.678	0.46
45	CuNMn <sub>3</sub>	3.68	1522.8	1.35	0.65	1.4	0.95	-44.453	0.46
46	GaCMn <sub>3</sub>	3.73	306.2	1.3	0.7	1.4	0.91	-44.461	0.50
47	CuNNi <sub>3</sub>	3.69	87.4	1.35	0.65	1.35	0.95	-37.456	0.48
48	GaCNi <sub>3</sub>	3.74	713.4	1.3	0.7	1.35	0.91	-37.764	0.52
49	GaN Mn <sub>3</sub>	3.71	799.4	1.3	0.65	1.4	0.93	-44.614	0.46
50	GaNNi <sub>3</sub>	3.73	152.2	1.3	0.65	1.35	0.94	-37.392	0.48
51	InCNi <sub>3</sub>	3.83	67.4	1.55	0.7	1.35	1.00	-37.167	0.52
52	InCTi <sub>3</sub>	4.21	171.7	1.55	0.7	1.4	0.99	-37.367	0.50
53	InCY <sub>3</sub>	4.85	82.6	1.55	0.7	1.85	0.94	-31.003	0.38
54	InNCo <sub>3</sub>	3.78	210.6	1.55	0.65	1.35	1.03	-39.801	0.48
55	InNNi <sub>3</sub>	3.82	125.7	1.55	0.65	1.35	1.03	-36.74	0.48
56	LaCRh <sub>3</sub>	4.20	225.0	1.95	0.7	1.35	1.14	-41.466	0.52
57	LuCRh <sub>3</sub>	4.08	538.6	1.75	0.7	1.35	1.07	-42.601	0.52
58	MgCNi <sub>3</sub>	3.77	148.2	1.5	0.7	1.35	0.98	-36.478	0.52
59	MgNNi <sub>3</sub>	3.77	500.0	1.5	0.65	1.35	1.01	-35.734	0.48
60	MoCNi <sub>3</sub>	3.77	588.2	1.45	0.7	1.35	0.97	-43.519	0.52
61	NbCNi <sub>3</sub>	3.81	513.5	1.45	0.7	1.35	0.97	-43.916	0.52
62	NiCW <sub>3</sub>	4.06	500.0	1.35	0.7	1.35	0.93	-52.906	0.52
63	NiNCu <sub>3</sub>	3.76	500.0	1.35	0.65	1.35	0.95	-28.303	0.48
64	NiNW <sub>3</sub>	4.03	263.8	1.35	0.65	1.35	0.95	-52.989	0.48
65	OsNFe <sub>3</sub>	3.69	1157.8	1.3	0.65	1.4	0.93	-49.469	0.46
66	PbNCa <sub>3</sub>	4.88	46.8	1.8	0.65	1.8	1.04	-22.872	0.36
67	PbTaO <sub>3</sub>	4.03	500.0	1.8	1.45	0.6	0.83	-40.475	2.42

68	PbTiO <sub>3</sub>	3.92	222.2	1.8	1.4	0.6	0.85	-39.022	2.33
69	PdCNi <sub>3</sub>	3.75	500.0	1.4	0.7	1.35	0.95	-37.014	0.52
70	PdNCu <sub>3</sub>	3.82	414.9	1.4	0.65	1.35	0.97	-25.069	0.48
71	PdNNi <sub>3</sub>	3.74	326.3	1.4	0.65	1.35	0.97	-36.542	0.48
72	PNCa <sub>3</sub>	4.66	61.6	1	0.65	1.8	0.81	-26.268	0.36
73	RuNCu <sub>3</sub>	3.82	77.2	1.3	0.65	1.35	0.94	-28.921	0.48
74	RuCNi <sub>3</sub>	3.74	762.3	1.3	0.7	1.35	0.91	-41.929	0.52
75	RuNFe <sub>3</sub>	3.68	500.0	1.3	0.65	1.4	0.93	-48.218	0.46
76	PtNNi <sub>3</sub>	3.75	224.8	1.35	0.65	1.35	0.95	-40.04	0.48
77	RhCNi <sub>3</sub>	3.73	262.5	1.35	0.7	1.35	0.93	-40.929	0.52
78	RhNCu <sub>3</sub>	3.81	380.3	1.35	0.65	1.35	0.95	-28.276	0.48
79	RhNFe <sub>3</sub>	3.68	594.3	1.35	0.65	1.4	0.95	-46.904	0.46
80	SbNBa <sub>3</sub>	5.46	27.0	1.45	0.65	2.15	0.91	-23.756	0.30
81	SbNCa <sub>3</sub>	4.81	50.1	1.45	0.65	1.8	0.94	-25.454	0.36
82	SbNMg <sub>3</sub>	4.34	85.3	1.45	0.65	1.5	0.97	-21.693	0.43
83	SbNNi <sub>3</sub>	3.89	600.2	1.45	0.65	1.35	0.99	-37.382	0.48
84	SbNSr <sub>3</sub>	5.12	23.5	1.45	0.65	2	0.92	-23.785	0.33
85	ScCRh <sub>3</sub>	4.05	500.0	1.6	0.7	1.35	1.02	-42.285	0.52
86	SnNCa <sub>3</sub>	4.86	39.0	1.45	0.65	1.8	0.94	-23.374	0.36
87	SnNMn <sub>3</sub>	3.81	358.6	1.45	0.65	1.4	0.98	-44.719	0.46
88	SnNNi <sub>3</sub>	3.86	500.0	1.45	0.65	1.35	0.99	-37.367	0.48
89	SrCrO <sub>3</sub>	3.77	295.5	2	1.4	0.6	0.92	-39.862	2.33
90	SrFeO <sub>3</sub>	3.74	252.1	2	1.4	0.6	0.92	-36.803	2.33
91	SrNbO <sub>3</sub>	4.04	263.7	2	1.45	0.6	0.90	-42.25	2.42
92	SrPuO <sub>3</sub>	4.25	151.7	2	1.75	0.6	0.78	-39.623	2.92
93	SrTaO <sub>3</sub>	4.02	181.2	2	1.45	0.6	0.90	-42.064	2.42
94	SrThO <sub>3</sub>	4.50	41.8	2	1.8	0.6	0.77	-41.215	3.00
95	SrTiO <sub>3</sub>	3.89	301.1	2	1.4	0.6	0.92	-41.267	2.33
96	SrUO <sub>3</sub>	4.33	161.3	2	1.75	0.6	0.78	-42.109	2.92
97	SrVO <sub>3</sub>	3.82	293.4	2	1.35	0.6	0.94	-40.518	2.25
98	SrZrO <sub>3</sub>	4.12	206.0	2	1.55	0.6	0.86	-42.337	2.58
99	TiCTi <sub>3</sub>	4.22	121.8	1.9	0.7	1.4	1.11	-36.632	0.50
100	WClr <sub>3</sub>	4.06	951.6	1.35	0.7	1.35	0.93	-49.118	0.52
101	WCRh <sub>3</sub>	4.02	554.5	1.35	0.7	1.35	0.93	-44.347	0.52
102	YCNi <sub>3</sub>	3.92	101.0	1.85	0.7	1.35	1.10	-39.929	0.52

103	YCRh <sub>3</sub>	4.12	6.6	1.85	0.7	1.35	1.10	-42.303	0.52
104	ZnNCu <sub>3</sub>	3.89	151.0	1.35	0.65	1.35	0.95	-22.084	0.48
105	ZnNNi <sub>3</sub>	3.71	93.3	1.35	0.65	1.35	0.95	-35.445	0.48
106	ZrCNi <sub>3</sub>	3.85	1359.9	1.55	0.7	1.35	1.00	-42.742	0.52
107	SrCNi <sub>3</sub>	3.99	11.5	2	0.7	1.35	1.16	-34.904	0.52
108	ZnCNi <sub>3</sub>	3.72	146.7	1.35	0.7	1.35	0.93	-36.065	0.52
109	BaCNi <sub>3</sub>	4.09	306.9	2.15	0.7	1.35	1.21	-34.032	0.52
110	BaCrO <sub>3</sub>	3.88	500.0	2.15	1.4	0.6	0.97	-39.565	2.33
111	BaFeO <sub>3</sub>	3.86	187.0	2.15	1.4	0.6	0.97	-36.47	2.33
112	BaMoO <sub>3</sub>	4.04	278.9	2.15	1.45	0.6	0.95	-41.164	2.42
113	BaNbO <sub>3</sub>	4.10	261.2	2.15	1.45	0.6	0.95	-42.638	2.42
114	BaNpO <sub>3</sub>	4.33	136.0	2.15	1.75	0.6	0.83	-42.361	2.92
115	TiCY <sub>3</sub>	4.85	55.7	1.9	0.7	1.85	1.04	-30.485	0.38
116	TiNca <sub>3</sub>	4.88	59.0	1.9	0.65	1.8	1.07	-20.893	0.36
117	GeNca <sub>3</sub>	4.77	58.0	1.25	0.65	1.8	0.88	-23.742	0.36
118	IrNFe <sub>3</sub>	3.69	500.0	1.35	0.65	1.4	0.95	-48.723	0.46
119	SnCMn <sub>3</sub>	3.83	439.1	1.45	0.7	1.4	0.96	-44.743	0.50
120	WCCo <sub>3</sub>	3.73	500.0	1.35	0.7	1.35	0.93	-47.947	0.52

### III.3.2 Outil d'analyse

Il existe plusieurs logiciels dans le domaine de l'informatique qui peuvent utiliser et représenter les données statistiques de l'analyse en composantes principales, telles que: S, SPSS, statistica, XLstat, R, Origin ... etc.

Dans ce travail, nous avons opté pour le logiciel Origin pour analyser à travers l'ACP le jeu de données récolté. Origin est un logiciel de traitement et d'analyse de données scientifiques pour environnement Microsoft Windows développé par OriginLab. L'Origin est un outil de mise en forme graphique fournissant tout un éventail de fonctionnalités (analyse de pics, ajustement de courbes, statistiques...) permettant de satisfaire aux exigences de qualité et aux besoins spécifiques de la communauté scientifique (chercheurs, étudiants, ingénieurs, techniciens...).

### III.3.3 Choix des axes

Pour un plan factoriel donné, on regarde le pourcentage de l'information du nuage initial retranscrite par le plan factoriel. Ainsi, un axe expliquant moins de 10% sera rarement intéressant, les axes sont rangés dans l'ordre décroissant, de telle manière que le premier plan

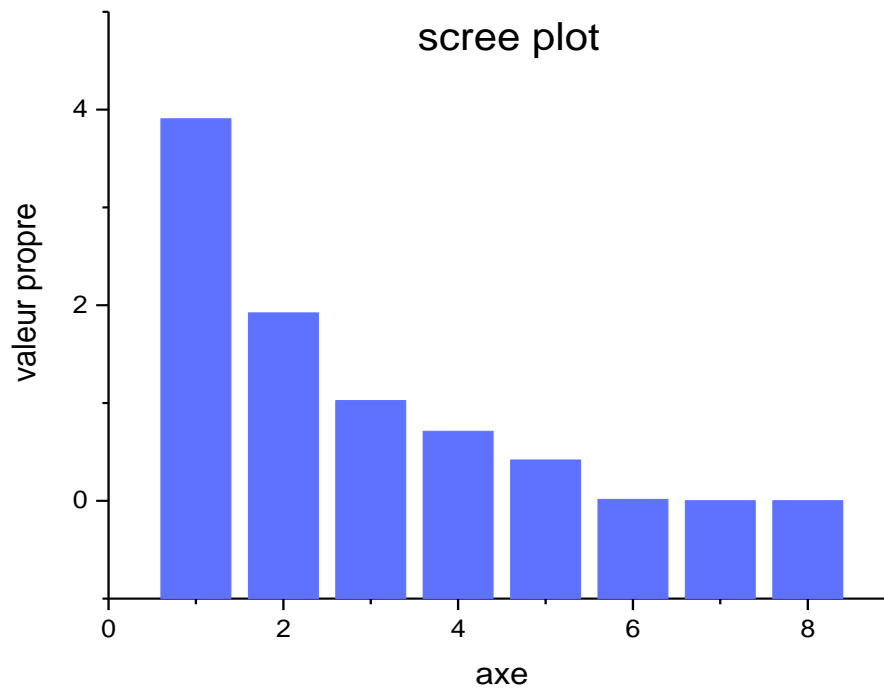
factoriel - constitué par les deux premiers axes factoriels - soit toujours celui qui est le plus riche en renseignements sur les propriétés du nuage étudié. Dans notre étude, les axes PC1 et PC2 sont les plus riches en information. (figure 6)

**Tableau III.3.** *Les valeurs propres des axes principaux*

	Valeur propre	Variabilité (%)	Cumulative
PC1	3,90595	48.82%	48.82%
PC2	1,92169	24.02%	72.85%
PC3	1,02542	12.82%	85.66%
PC4	0,71188	8.90%	94.56%
PC5	0,4181	5.23%	99.79%
PC6	0,01477	0.18%	99.97%
PC7	0,0013	0.02%	99.99%
PC8	8,89E-4	0.01%	100.00%

L'analyse en composantes principales nous déduit les axes principaux, dans notre cas la première composante principale (PC1) est le vecteur propre qui correspond à la plus grande valeur propre (3,90595), et (PC2) est le vecteur propre correspondant à la deuxième plus grande valeur propre (1,92169).

D'après le tableau 4 on voit que PC1 porte 48.82% de l'information pour la base de tableaux 3 et l'axe PC2 porte 24.02% de l'information pour la même base, donc nous récupérons 72.85% de l'information contenue dans tout le tableau 2. Autrement dit, nous avons un bon résumé qui synthétise presque totalement les 8 variables. La perte d'information de 27.15% est répartie sur les 6 autres PC avec des valeurs propres relativement faibles (tableau 4), ce qui signifie qu'il y a très peu d'informations perdues et il n'est pas nécessaire d'aller visualiser ces PC.



**Figure III.1** Représentation graphique des axes principaux

### III.3.4 Matrice de corrélation

Suite à une analyse en composantes principales, le premier résultat intéressant à décrire est la matrice de corrélation qui permet d'analyser les relations binaires qui existent entre les différentes variables sélectionnées. Cette matrice est présentée dans le tableau 5.

**Tableau III.4** Matrice de corrélation

	a	B	$r_A$	$r_B$	$r_X$	T	$E_{coh}$	$\mu$
A	1	-0,33923	0,10086	0,04211	0,39147	-0,1597	0,57701	-0,01284
B	-0,33923	1	-0,05982	-0,0138	-0,1006	0,00781	-0,38301	-0,00562
$r_A$	0,10086	-0,05982	1	0,78201	-0,70977	0,00614	-0,15073	0,78023
$r_B$	0,04211	-0,0138	0,78201	1	-0,88798	-0,60029	-0,23567	0,99616
$r_X$	0,39147	-0,1006	-0,70977	-0,88798	1	0,41018	0,43168	-0,91612
T	-0,1597	0,00781	0,00614	-0,60029	0,41018	1	0,06392	-0,58239
$E_{coh}$	0,57701	-0,38301	-0,15073	-0,23567	0,43168	0,06392	1	-0,2471

$\mu$	-0,01284	-0,00562	0,78023	0,99616	-0,91612	-0,58239	-0,2471	1
-------	----------	----------	---------	---------	----------	----------	---------	---

Sur la base du tableau précédent on remarque que le facteur de tolérance  $t$  est inversement corrélé avec le rayon ionique  $r_B$  et le facteur d'octaèdre  $\mu$  avec une valeur de -0,60029 et -0,58239 respectivement. On peut voir aussi une corrélation entre l'énergie de cohésion  $E_{coh}$  et le paramètre de maille  $a$  (0,57701).

### III.3.5 Les coordonnées des individus

Les coordonnées des individus sont les coordonnées artificielles des composés dans le nouvel espace. Ils sont ensuite représentés graphiquement par un graphe dit « score plot ». Le tableau 6 résume ces coordonnées pour les 120 composés considérés.

**Tableau III.5** *Les coordonnées des individus.*

N°	individus	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
1	AgCNi3	-0,99855	-0,70746	0,77112	-0,21389	0,27796	0,14776	-0,01361	0,01526
2	AgNNi3	-1,16421	-0,76963	1,00831	-0,1866	0,33773	0,05582	0,04283	0,02259
3	AlCMn3	-0,87876	-1,079	-0,89476	-1,18117	-0,61736	0,14606	-0,01462	0,01264
4	AlCNi3	-1,01904	-0,43675	-0,71587	-1,41515	0,02226	0,1007	-0,03857	-0,00559
5	AlCTi3	-1,13735	-0,00557	-0,95372	-0,54784	-0,34223	-0,22374	-0,08503	0,00485
6	AlNNi3	-1,13548	-0,97672	-0,7563	-0,54556	0,27687	-0,0064	0,00651	-0,00932
7	AgNCu3	-1,53106	0,45853	1,11826	-0,08365	1,45062	0,0123	0,0044	0,03699
8	AlCFe3	-0,88889	-1,08552	-0,87684	-1,24256	-0,50102	0,19589	-0,00925	-0,01331
9	AlCY3	-1,90025	1,82277	-1,3099	0,06356	-0,71737	-0,01174	-0,00378	0,03835
10	AlNFe3	-1,01106	-1,46313	-0,82348	-0,72278	-0,33286	0,10971	0,03803	-0,02058
11	AsNBa3	-2,65322	3,22092	-1,81327	0,71661	-0,73217	-0,10813	0,12496	0,07863
12	AsNCa3	-2,13083	1,97872	-1,49113	-0,10178	0,02967	-0,0762	0,02593	0,02913
13	AsNMg3	-1,82422	1,31487	-1,19277	-0,53362	1,12307	-0,11741	-0,04696	0,00757
14	AsNSr3	-2,43865	2,67347	-1,66029	0,29333	-0,27348	-0,0329	0,08536	0,05471
15	BaAmO3	3,47807	1,10321	-0,02186	0,04535	-0,22857	0,21967	-0,01867	0,02201
16	CdCNi3	-1,03491	-0,68673	0,46893	-0,01442	0,44703	0,11158	-0,02522	0,01479
17	CdNNi3	-1,21417	-0,63108	0,74967	-0,13736	0,48846	0,02148	0,02933	0,02067
18	TcCNi3	-0,86665	-1,18227	-0,49148	-0,54111	-0,23001	0,07088	-0,02573	-0,00311
19	BaPaO3	3,73895	0,74558	-0,406	0,72203	-0,65811	0,11457	-0,02724	0,03934
20	BaPrO3	3,82007	1,02927	-0,49203	0,53979	-0,35405	0,25428	-0,01631	0,05963
21	BaPuO3	3,76554	-2,05735	-1,61747	5,39611	0,87084	0,05266	-0,03472	0,02549
22	BaRuO3	1,95467	0,03353	1,75689	0,3461	0,45566	-0,30213	0,0713	0,02805
23	BaSnO3	2,37212	0,88039	1,17144	0,05669	0,53768	-0,09345	0,00483	-0,00733
24	BaTaO3	2,62651	0,02019	1,04298	0,18478	-0,28008	-0,12355	0,02337	-0,0158

25	BaThO3	3,65933	1,3999	-0,19479	0,08999	-0,85566	0,03934	-0,04092	0,04234
26	BaTiO3	2,43378	-0,04086	1,33475	-0,0486	-0,07332	-0,10057	0,04892	-0,00907
27	BaTbO3	3,44091	0,91798	-0,193	0,7041	0,19695	0,22586	-0,02388	0,0246
28	BaUO3	3,55137	1,04023	-0,0168	-0,04792	-0,7283	0,11899	-0,02395	0,02099
29	BaZrO3	2,92898	0,64926	0,80923	-0,42464	-0,58374	-0,01917	-0,00105	-0,01884
30	BiAlO3	1,7302	0,01025	-0,75093	-0,61879	1,02893	-0,26267	0,00615	-0,05905
31	BiGaO3	1,81356	0,10177	-1,14985	0,17957	1,49747	-0,2494	-0,02151	-0,05538
32	BiInO3	2,44415	1,70227	-1,70096	-0,85051	1,18406	-0,04461	-0,04983	8,74155E-5
33	BiNCa3	-2,03602	2,11556	0,39061	0,72875	-0,07068	-0,00343	-0,02003	-0,04272
34	BiNBa3	-2,49942	3,42553	-0,02644	1,32608	-0,91038	0,03487	0,03695	-0,05764
35	BiScO3	2,82082	0,98599	-1,93718	-0,87734	0,5344	0,05731	-0,0174	0,01076
36	BiNSr3	-2,31104	2,87797	0,17513	0,95512	-0,42988	0,08126	0,01378	-0,05598
37	CaSnO3	2,32429	0,84958	-0,50096	-0,45928	0,80576	-0,08183	-0,01821	-0,02686
38	CaTaO3	2,57976	-0,00472	-0,63183	-0,32187	-0,03373	-0,12195	-8,2226E-4	-0,03517
39	CaCrO3	2,44504	-0,71744	-0,58573	0,12673	0,62555	0,04534	0,0341	-0,04305
40	CaTiO3	2,39763	0,05422	-0,26528	-0,93328	0,08688	-0,05135	0,02242	-0,04172
41	CaZrO3	2,8685	0,47399	-0,90992	-0,58539	-0,24084	-0,05496	-0,01277	-0,01538
42	CaCNi3	-0,98103	-0,84349	1,56682	0,67857	0,28145	0,07818	-0,01135	0,02914
43	CdNCu3	-1,60255	0,72093	0,87463	-0,05239	1,58901	-0,0545	-0,01488	0,03642
44	CoNFe3	-0,85689	-2,36826	-0,64434	0,29644	-0,3823	0,14408	0,05822	-0,01904
45	CuNMn3	-0,8946	-2,74495	-0,99516	1,61424	0,16383	0,0721	0,04085	-0,01366
46	GaCMn3	-0,88609	-1,15093	-0,72119	-0,87077	-0,46771	0,16055	-0,01302	-0,01068
47	CuNNi3	-1,17816	-0,57854	-0,04641	-1,14684	0,24671	0,08143	0,02512	-0,00166
48	GaCNi3	-0,97542	-1,17169	-0,88616	0,07478	0,45387	0,07903	-0,0401	-8,91132E-4
49	GaN Mn3	-0,97532	-1,84266	-0,80862	0,12239	-0,22236	0,06185	0,03521	-0,01643
50	GaNNi3	-1,18753	-0,57065	-0,33635	-1,03446	0,24334	0,02323	0,01209	-0,00449
51	InCNi3	-1,03233	-0,32354	0,68999	-0,805	0,06695	0,10293	-0,02287	0,01324
52	InCTi3	-1,12688	0,12136	0,4922	-0,19293	-0,37585	-0,17541	-0,0692	0,01227
53	InCY3	-1,80545	1,81856	-0,06399	0,4923	-0,7628	0,11128	-0,03394	-0,02499
54	InNCo3	-1,09554	-0,82245	0,83171	-0,58787	-0,08346	0,02394	0,04202	0,01538
55	InNNi3	-1,19163	-0,44896	0,88829	-0,6589	0,15263	0,0073	0,03171	0,01905
56	LaCRh3	-0,90979	-0,39552	2,43843	0,33472	-0,81345	-0,18002	-0,02535	0,03913
57	LuCRh3	-0,85861	-0,94955	1,30426	0,56728	-0,61496	-0,14688	-0,03041	0,0257
58	MgCNi3	-1,04059	-0,43115	0,4129	-0,74829	0,24755	0,14015	-0,0225	0,01012
59	MgNNi3	-1,17864	-0,87257	0,43235	-9,87367E-4	0,50035	0,02896	0,02817	0,01505
60	MoCNi3	-0,81931	-1,43024	-0,11909	-0,06096	-0,27547	0,06295	-0,01804	0,002
61	NbCNi3	-0,82243	-1,32034	-0,08168	-0,18338	-0,39199	0,03086	-0,02171	0,00245
62	NiCW3	-0,64807	-1,53804	-0,63283	-0,24611	-1,61696	-0,29741	-0,05201	-0,00376
63	NiNCu3	-1,39144	-0,31203	-0,26852	-0,06305	1,31971	0,04247	-0,00299	0,00905
64	NiNW3	-0,81326	-1,38577	-0,27165	-0,73227	-1,69541	-0,35118	0,00362	-0,00623
65	OsNFe3	-0,80929	-2,62938	-1,04012	0,73553	-0,54575	0,03068	0,04379	-0,02067
66	PbNCa3	-2,03536	2,33918	1,29084	0,91556	0,02963	0,07113	-0,03832	-0,07429
67	PbTaO3	2,54935	-0,01427	-0,72311	0,05389	0,13064	-0,16647	-0,01043	-0,03248

68	PbTiO3	2,33862	0,17924	-0,3396	-0,57508	0,28216	-0,11912	0,00751	-0,0376
69	PdCNi3	-1,0031	-0,88241	-0,27678	-0,20037	0,40027	0,11152	-0,03118	0,00468
70	PdNCu3	-1,49008	0,09163	0,02909	-0,04149	1,53769	0,01674	-0,01031	0,01763
71	PdNNi3	-1,18253	-0,72924	0,04898	-0,52327	0,38119	0,03217	0,02094	0,00514
72	PNCa3	-2,16562	1,92248	-2,13582	-0,31056	0,08747	-0,10764	0,03983	0,05347
73	RuNCu3	-1,43318	0,23846	-0,27135	-0,9346	1,00105	-0,00649	-0,01383	0,00523
74	RuCNi3	-0,86053	-1,52643	-0,9319	0,09119	0,04169	0,05455	-0,03288	-0,0046
75	RuNFe3	-0,90509	-1,78469	-0,63942	-0,57914	-0,71103	0,08287	0,04684	-0,02081
76	PtNNi3	-1,1075	-0,83406	-0,14921	-0,85584	-0,03007	4,77103E-5	0,0204	-0,00216
77	RhCNi3	-0,92818	-0,88739	-0,38804	-0,83569	-0,09536	0,10246	-0,02622	-0,00205
78	RhNCu3	-1,41358	-0,09187	-0,20566	-0,24816	1,20491	-0,00332	-0,00977	0,01032
79	RhNFe3	-0,923	-1,81652	-0,4505	-0,29799	-0,53924	0,09864	0,0492	-0,01701
80	SbNBa3	-2,54667	3,35464	-0,62052	1,1131	-0,87824	-0,01732	0,06614	-0,0123
81	SbNCa3	-2,06739	2,11229	-0,21416	0,3725	-0,08477	-0,03278	-0,00499	-0,01885
82	SbNMg3	-1,82214	1,52799	0,1961	-0,01639	1,04031	-0,13854	-0,04915	0,00814
83	SbNNi3	-1,15351	-0,91349	0,10136	0,23132	0,24397	-0,11721	0,00709	0,01281
84	SbNSr3	-2,34911	2,8059	-0,43432	0,72094	-0,4057	0,03976	0,03767	-0,01921
85	ScCRh3	-0,88686	-0,88916	0,61402	0,26347	-0,52924	-0,14943	-0,04156	0,01758
86	SnNCa3	-2,13215	2,34411	-0,20737	0,44179	0,07169	-0,0685	-0,01537	-0,01569
87	SnNMn3	-0,81351	-3,58641	-1,03781	3,56069	0,36015	-0,072	0,02917	-0,0043
88	SnNNi3	-1,15782	-0,84291	0,16689	-0,00505	0,23634	-0,07965	0,01221	0,0118
89	SrCrO3	2,4183	-0,2359	0,61302	-0,32968	0,35251	0,06051	0,05412	-0,02542
90	SrFeO3	2,3389	-0,01289	0,65634	-0,38986	0,68835	0,10993	0,05327	-0,02355
91	SrNbO3	2,60119	0,09183	0,37038	-0,18756	-0,21738	-0,11859	0,01367	-0,02431
92	SrPuO3	3,4462	1,11872	-0,65488	-0,26062	-0,23051	0,18247	-0,01125	0,03928
93	SrTaO3	2,59132	0,17634	0,42347	-0,36886	-0,21645	-0,09607	0,01625	-0,02473
94	SrThO3	3,58527	1,55216	-0,7892	-0,27317	-0,72894	0,01249	-0,03251	0,06434
95	SrTiO3	2,43557	-0,1779	0,58519	-0,22747	0,07796	-0,05797	0,04091	-0,02368
96	SrUO3	3,49885	1,04306	-0,68334	-0,20935	-0,57512	0,093	-0,0176	0,0391
97	SrVO3	2,25255	-0,28785	0,82201	-0,27099	0,22842	-0,08437	0,06008	-0,01655
98	SrZrO3	2,91518	0,41301	0,00332	-0,26557	-0,34543	-0,02192	-0,00438	-0,01757
99	TiCTi3	-1,09025	0,138	2,18005	0,17391	-0,39537	-0,06623	-0,04197	0,01898
100	WCr3	-0,70234	-1,80848	-0,88706	0,73244	-1,00054	-0,29391	-0,06036	1,51588E-4
101	WCRh3	-0,85982	-1,06422	-0,62333	-0,01934	-0,63934	-0,20614	-0,06077	0,00276
102	YCNi3	-0,92802	-0,51146	2,08428	-0,3031	-0,37434	0,07638	2,82617E-4	0,02809
103	YCRh3	-0,91049	-0,28009	2,09821	-0,33209	-0,89598	-0,12596	-0,02254	0,03109
104	ZnNCu3	-1,61201	0,71796	-0,05522	-0,51757	1,66601	-0,02503	-0,02902	0,01741
105	ZnNNi3	-1,23547	-0,40356	-0,04646	-1,06737	0,42916	0,06586	0,0178	8,80045E-4
106	ZrCNi3	-0,76349	-2,19156	-0,11128	1,72679	0,05477	-0,02558	-0,0246	0,01094
107	SrCNi3	-1,0584	-9,11509E-4	2,8678	-0,12388	0,00662	0,0873	-0,00121	0,0416
108	ZnCNi3	-0,96655	-1,60269	-0,89536	1,02681	0,84797	0,09846	-0,0376	0,00323
109	BaCNi3	-1,04728	-0,18882	3,39873	0,78581	0,09523	0,02279	-0,00215	0,0528
110	BaCrO3	2,43126	-0,35749	1,20734	0,38953	0,33594	-0,00862	0,05807	-0,00977

111	BaFeO3	2,32265	0,19833	1,4107	-0,2043	0,53795	0,04588	0,05737	-0,00801
112	BaMoO3	2,59615	0,10994	1,08582	0,0614	-0,12885	-0,07447	0,0271	-0,01585
113	BaNbO3	2,62293	0,10974	1,08094	0,05713	-0,35775	-0,13789	0,02183	-0,01567
114	BaNpO3	3,54606	1,02226	-0,00772	-0,08137	-0,64302	0,1562	-0,01978	0,02043
115	TiCY3	-1,69514	1,82198	1,42436	0,8603	-0,78731	0,29345	-0,06394	-0,09984
116	TiNCa3	-2,05361	2,44413	1,7217	1,09786	0,22428	0,1289	-0,04699	-0,08958
117	GeNCa3	-2,16928	2,20268	-1,06513	0,13703	0,18604	-0,07542	0,00833	0,01573
118	IrNFe3	-0,88602	-1,82112	-0,40289	-0,51377	-0,78523	0,08418	0,05162	-0,01854
119	SnCMn3	-0,85573	-1,23572	-0,10765	-0,31929	-0,57015	0,10761	-0,01371	-0,00571
120	WCCo3	-0,71915	-1,66843	-0,55777	-0,49623	-0,71596	0,05401	-0,01479	-0,0081

### III.3.6 Score plot

La première analyse consiste à examiner les différences entre les 120 pérovskites. Le résultat des "Score plot" de cette analyse est montré sur la figure 7.

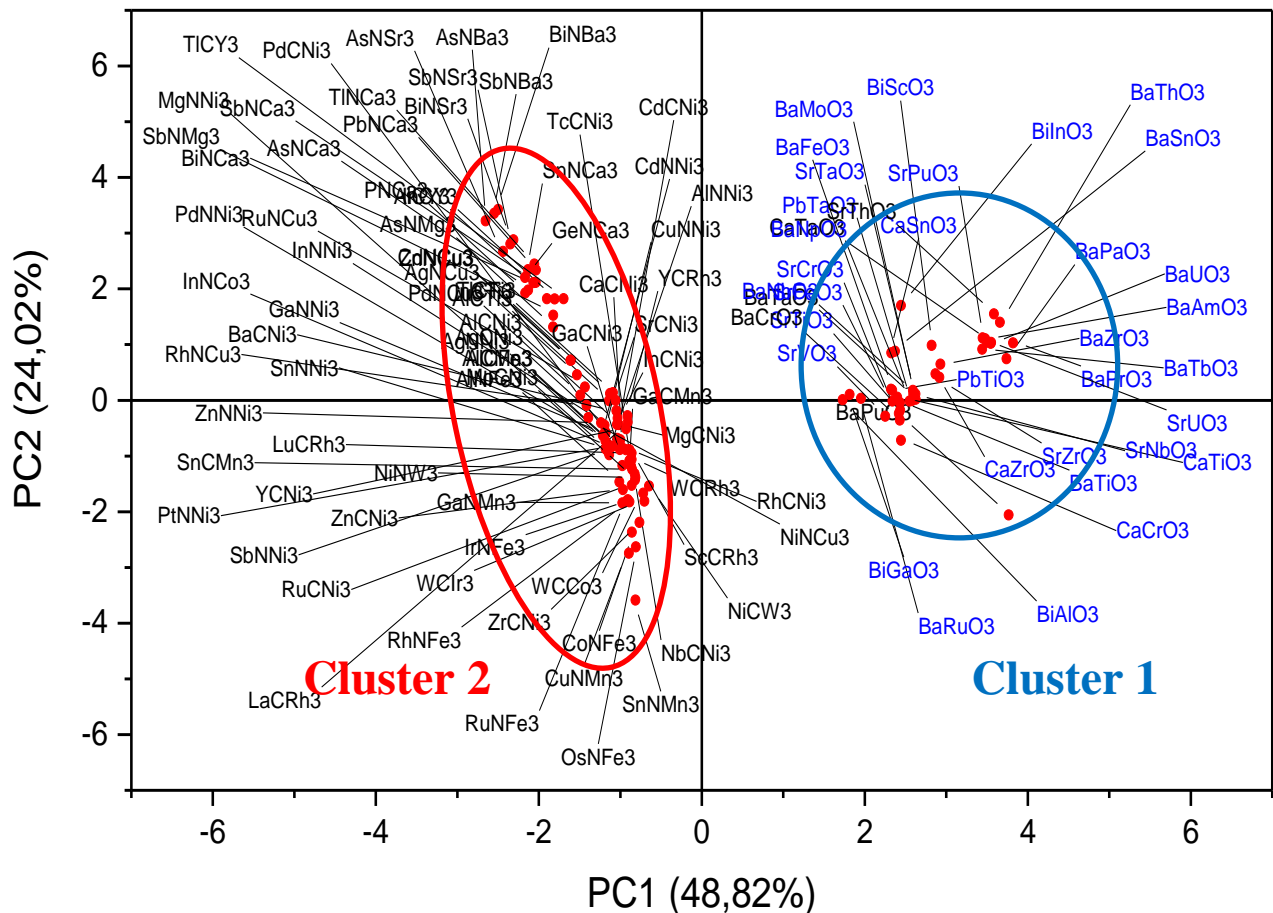


Figure III.2 Score plot des individus ( $ABX_3$ )

Pour cette analyse, le signe de chaque composante principale n'a qu'une signification relationnelle. A partir de figure 7, il ressort deux regroupements importants, l'un (que nous appellerons cluster1) qui inclut les pérovskites (valeurs positives de PC1) et le cluster 2 qui correspond aux antipérovskites (PC1 négatif). Par conséquent, l'ACP fait clairement la distinction entre les pérovskites et antipérovskites.

### III.3.7 Les coordonnées des variables

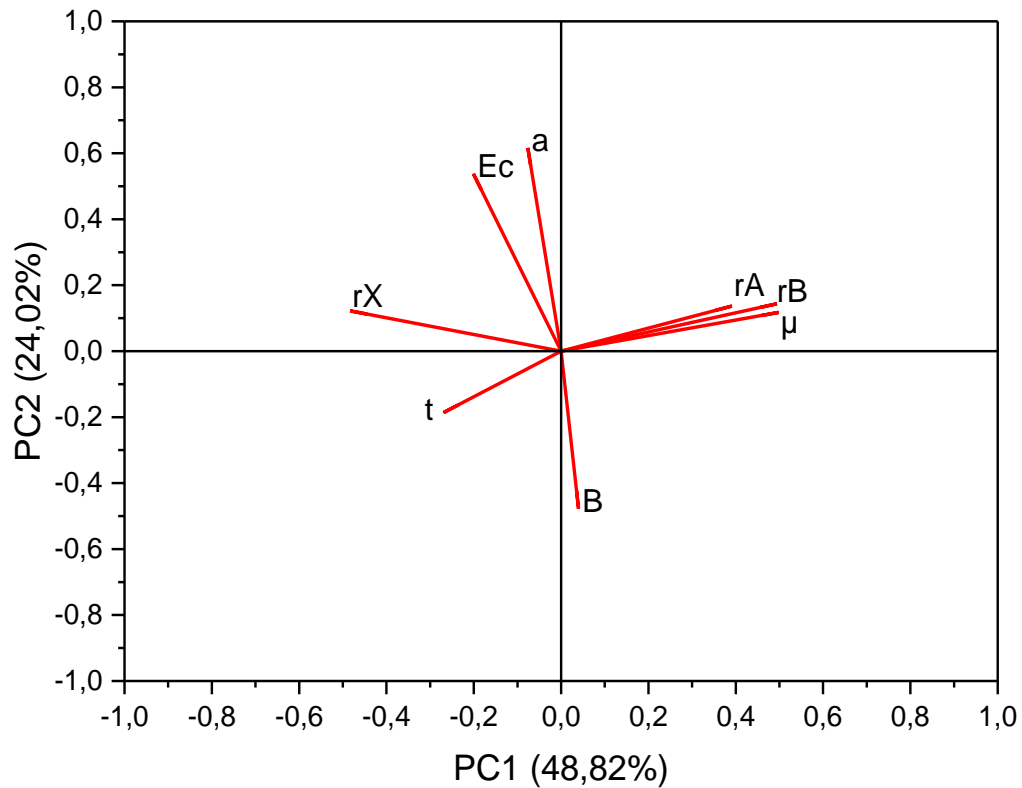
Ce sont les coordonnées dans le nouvel espace qui représente la corrélation entre les variables d'origine en fonction des composants dans le nouvel espace. Le tableau 7 représente les coordonnées des variables que nous étudions.

**Tableau III.6** Les coordonnées des variables

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
A	-0,07632	0,61652	-0,07064	0,4432	-0,48743	-0,41412	-0,05874	0,01126
B	0,04003	-0,47713	-0,24363	0,82126	0,19089	-0,02154	-0,00233	6,07917E-4
rA	0,3913	0,13661	0,55273	0,25559	-0,08706	0,44153	-0,28151	-0,41906
rB	0,49449	0,14401	-0,03992	0,05343	-0,0027	0,22995	-0,12268	0,81385
rX	-0,48333	0,12234	-0,07878	0,15147	-0,26138	0,70343	0,38164	0,11612
T	-0,26953	-0,18621	0,78793	0,12717	0,02041	-0,27584	0,24871	0,34251
E <sub>coh</sub>	-0,20093	0,53797	0,03096	0,14589	0,80357	0,04499	-0,01307	0,00681
μ	0,49831	0,11745	-0,02341	0,02271	0,06258	-0,08536	0,83341	-0,17624

### III.3.8 Loading plot

Les coordonnées des variables peuvent être représentées dans un graphe des variables dit « loading plot » et dont les axes sont les deux axes principaux PC1 et PC2. Figure 8 représente loading plot de notre cas.



**Figure III.3** Loading plot

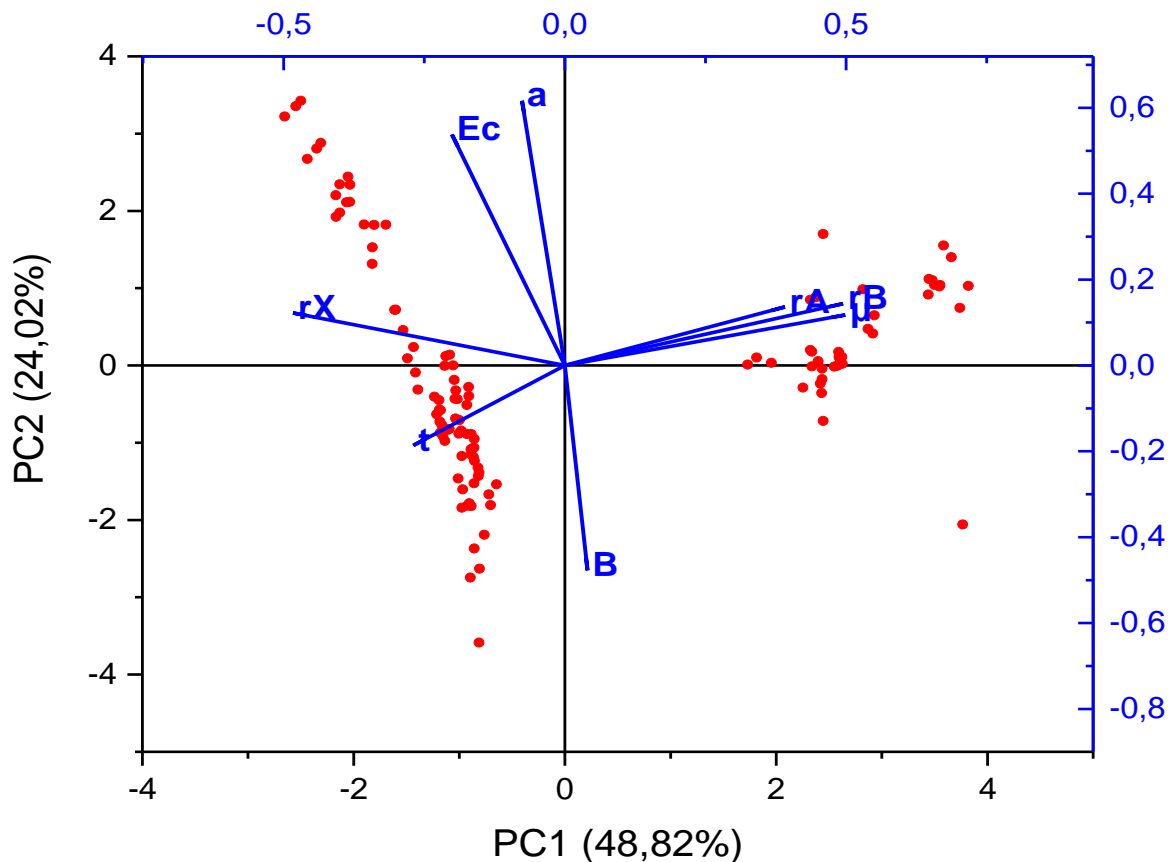
A partir du graphe des variables de la figure 8, nous observons :

- Une corrélation positive entre le paramètre de maille  $a$  et l'énergie de cohésion  $E_{coh}$ .
- une forte corrélation positive entre le rayon ionique de site B  $r_B$  et le facteur d'octaèdre  $\mu$ .
- une corrélation positive entre l'énergie de cohésion  $E_{coh}$  et le rayon ionique de site X  $r_X$ .
- une forte corrélation positive entre le facteur d'octaèdre  $\mu$ , le rayon ionique  $r_A$  et  $r_B$ .
- une corrélation négative entre le module de compressibilité isostatique B, paramétré de maille  $a$  et énergie de cohésion  $E_{coh}$ .
- une corrélation négative entre le facteur de tolérance  $t$  et  $r_A$ ,  $r_B$  et le facteur d'octaèdre  $\mu$ .
- Aucune corrélation ou presque nulle entre :
  - Le facteur de tolérance  $t$  et le paramètre de maille  $a$ .
  - Le facteur de tolérance  $t$  et l'énergie de cohésion  $E_{coh}$
  - Le paramètre de maille  $a$  et le facteur d'octaèdre  $\mu$ .

- L'énergie de cohésion  $E_{\text{coh}}$  et le rayon ionique  $r_A$ .
- Le module de compressibilité isostatique  $B$  et le facteur d'octaèdre  $\mu$ .
- L'énergie de cohésion  $E_{\text{coh}}$  et le rayon ionique  $r_B$ .
- Le module de compressibilité isostatique  $B$  et rayon ionique  $r_X$ .

### III.3.9 Biplot des variables et des individus :

Le biplot est un graphe indiquant la relation entre tous les pérovskites et les paramètres correspondants. Il s'agit donc d'une représentation graphique simultanée des individus et des variables selon les axes principaux PC1 et PC2 (figure 9).



**Figure III.4** Graphe de « Biplot » superposant les individus et leurs propriétés.

Pour cette analyse, le signe de chaque composante principale n'a relativement pas de sens physique. En se basant sur le graphe de figure 9 on peut constater que les matériaux du premier

cluster sont fortement corrélé avec les paramètres  $(r_A, r_B, \mu)$  cela veut dire que ces matériaux ont un grand facteur d'octaèdre  $\mu$  et grand rayon ionique  $r_A$  et  $r_B$ , et inversement corrélé avec les paramètres  $(r_X, t)$ . Tandis que les matériaux du deuxième cluster sont corrélé avec les paramètres  $(r_X, t)$  c'est-à-dire que c'est matériaux possède un grand facteur de tolérance  $t$  et grand rayon ionique  $r_X$  et inversement corrélé avec les paramètres  $(r_A, r_B, \mu)$ .

### **III.4 Conclusion**

Les valeurs calculées des paramètres structuraux optimisés de 120 pérovskites s'accordent avec les données expérimentales disponibles, cela confirme la validité de l'approche théorique et des approximations adoptées. L'ACP nous a permis d'établir des corrélations linéaires entre les propriétés étudiées des pérovskites: Corrélations linéaires positives entre les paramètres  $(r_A, r_B, \mu)$ , et entre  $(a, E_{\text{coh}}, r_X)$ . Corrélations linéaires négatives entre le facteur de tolérance  $t$  et les paramètres  $(r_A, r_B, \mu)$ . Le biplot permet de déceler les composés de fortes et faibles propriétés.

## Conclusion et perspective

Dans le cadre de ce mémoire, nous nous intéressons à l'étude des matériaux pérovskites. Ces composés ont des propriétés très intéressantes. Basé sur la théorie de la fonctionnelle de la densité (DFT) et l'approche des pseudo potentiels et onde plane telle qu'implémentée dans le code CASTEP, nous avons étudiés les propriétés structurales de 120 pérovskites cubiques afin de générer une base de donnés. Ensuite, nous avons exploré ces données en utilisant des techniques de datamining. En effet, l'analyse en composantes principales (ACP) a été appliquée pour identifier les corrélations entre les différentes propriétés de ces matériaux.

Les résultats obtenus nous ont permis d'affirmer selon une analyse dans un espace plan PC1 et PC2, les corrélations suivantes. Corrélations linéaires positives entre les paramètres ( $r_A$ ,  $r_B$ ,  $\mu$ ), et entre ( $a$ ,  $E_{\text{coh}}$ ,  $r_X$ ). Corrélations linéaires négatives entre le facteur de tolérance  $t$  et les paramètres ( $r_A$ ,  $r_B$ ,  $\mu$ ).

Par ailleurs, les résultats de biplot nous a permis de déceler les composés de fortes et faibles propriétés. Grace à ces résultats obtenus, nous pourrons à l'avenir développer un modèle prédictif qui nous permettra de calculer les propriétés à partir de données faciles à trouver en utilisant l'une des méthodes prédictives de datamining. Par exemple on prend le rayon ionique ( $r_A$ ,  $r_B$  ou  $r_X$ ) de pérovskite et à l'aide de technique de régression de moindre carré partiel (PLS) on peut prédire l'énergie de cohésion sans recourir à des calculs ab-initio.

## References

- [1] E. A. R. Assirey, « Perovskite synthesis, properties and their related biochemical and industrial application », *Saudi Pharmaceutical Journal*, vol. 27, n° 6, p. 817-829, sept. 2019, doi: 10.1016/j.jsps.2019.05.003.
- [2] Bastien L, « Data Mining : qu'est ce que l'exploration de données ? », *LeBigData.fr*, janv. 31, 2018. <https://www.lebigdata.fr/data-mining-definition-exemples> (consulté le juin 22, 2021).
- [3] R. Arar, « Etude des propriétés structurales, électroniques et mécaniques des composés fluoro-pérovskites à base de sodium NaXF<sub>3</sub> (X=Mg, Zn) par la méthode FP-LAPW », université Djilali Liabes, Sidi Bel Abbes, 2017.
- [4] K. Rida, « SYNTHÈSE, CARACTÉRISATION ET PROPRIÉTÉS CATALYTIQUES DES OXYDES La<sub>1-x</sub>M<sub>x</sub>CrO<sub>3</sub> ( M = Sr et Ce ) », université Mentouri, Constantine, 2008.
- [5] V. M. Goldschmidt, « Die Gesetze der Krystallochemie », *Naturwissenschaften*, vol. 14, n° 21, p. 477-485, mai 1926, doi: 10.1007/BF01507527.
- [6] J. B. Philipp *et al.*, « Structural and doping effects in the half-metallic double perovskite A<sub>2</sub>CrWO<sub>6</sub> ( A = Sr, Ba, and Ca ) », *Phys. Rev. B*, vol. 68, n° 14, p. 144431, oct. 2003, doi: 10.1103/PhysRevB.68.144431.
- [7] C. Li, X. Lu, W. Ding, L. Feng, Y. Gao, et Z. Guo, « Formability of ABX<sub>3</sub> ( X = F, Cl, Br, I) halide perovskites », *Acta Crystallogr B Struct Sci*, vol. 64, n° 6, p. 702-707, déc. 2008, doi: 10.1107/S0108768108032734.
- [8] F. Dogan, H. Lin, M. Guilloux-Viry, et O. Peña, « Focus on properties and applications of perovskites », *Science and Technology of Advanced Materials*, vol. 16, n° 2, p. 020301, avr. 2015, doi: 10.1088/1468-6996/16/2/020301.
- [9] N. Das et S. Kandimalla, « Application of perovskites towards remediation of environmental pollutants: an overview: A review on remediation of environmental pollutants using perovskites », *Int. J. Environ. Sci. Technol.*, vol. 14, n° 7, p. 1559-1572, juill. 2017, doi: 10.1007/s13762-016-1233-7.
- [10] O. Ourahmoun, « Les cellules solaires à base de matériaux », *Energies Renouvelables*, vol. 21, n° 4, p. 6, 2018.

- [11] P. Dally, « Cellules Solaires à base de Matériaux Pérovskites: De la caractérisation des matériaux à l'amélioration des rendements et de la stabilité », Université Grenoble Alpes, France, 2019.
- [12] F. Ledee, « Cristallisation et fonctionnalisation de pérovskites hybrides halogénées à 2-dimensions pour le photovoltaïque et l'émission de lumière », Université Paris-Saclay, France, 2018.
- [13] E. Schrodinger, « Der stetige Ubergang von der Mikro- zur Makromechanik », *Naturwissenschaften*, vol. 14, n° 28, p. 664-666, juill. 1926, doi: 10.1007/BF01507634.
- [14] M. Born et R. Oppenheimer, « Zur Quantentheorie der Molekeln », *Ann. Phys.*, vol. 389, n° 20, p. 457-484, 1927, doi: 10.1002/andp.19273892002.
- [15] D. R. Hartree, « The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part I. Theory and Methods », *Math. Proc. Camb. Phil. Soc.*, vol. 24, n° 1, p. 89-110, janv. 1928, doi: 10.1017/S0305004100011919.
- [16] V. Fock, « Selfconsistent field mit Austausch für Natrium », *Z. Physik*, vol. 62, n° 11-12, p. 795-805, nov. 1930, doi: 10.1007/BF01330439.
- [17] L. H. Thomas, « The calculation of atomic fields », *Math. Proc. Camb. Phil. Soc.*, vol. 23, n° 5, p. 542-548, janv. 1927, doi: 10.1017/S0305004100011683.
- [18] E. Fermi, « A statistical method for the determination of some proprieta dell'atome », *Rend. Accad. Nat. Lincei*, vol. 6, n° 602-607, p. 32, 1927.
- [19] P. Hohenberg et W. Kohn, « Inhomogeneous Electron Gas », *Phys. Rev.*, vol. 136, n° 3B, p. B864-B871, nov. 1964, doi: 10.1103/PhysRev.136.B864.
- [20] W. Kohn et L. J. Sham, « Self-Consistent Equations Including Exchange and Correlation Effects », *Phys. Rev.*, vol. 140, n° 4A, p. A1133-A1138, nov. 1965, doi: 10.1103/PhysRev.140.A1133.
- [21] D. C. Langreth et M. J. Mehl, « Beyond the local-density approximation in calculations of ground-state electronic properties », *Phys. Rev. B*, vol. 28, n° 4, p. 1809-1834, août 1983, doi: 10.1103/PhysRevB.28.1809.
- [22] E. Engel et S. H. Vosko, « Exact exchange-only potentials and the virial relation as microscopic criteria for generalized gradient approximations », *Phys. Rev. B*, vol. 47, n° 20, p. 13164-13174, mai 1993, doi: 10.1103/PhysRevB.47.13164.

- [23] J. P. Perdew, K. Burke, et M. Ernzerhof, « Generalized Gradient Approximation Made Simple », *Phys. Rev. Lett.*, vol. 77, n° 18, p. 3865-3868, oct. 1996, doi: 10.1103/PhysRevLett.77.3865.
- [24] J. P. Perdew *et al.*, « Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces », *Phys. Rev. Lett.*, vol. 100, n° 13, p. 136406, avr. 2008, doi: 10.1103/PhysRevLett.100.136406.
- [25] U. von Barth et C. D. Gelatt, « Validity of the frozen-core approximation and pseudopotential theory for cohesive energy calculations », *Phys. Rev. B*, vol. 21, n° 6, p. 2222-2228, mars 1980, doi: 10.1103/PhysRevB.21.2222.
- [26] M. D. Segall *et al.*, « First-principles simulation: ideas, illustrations and the CASTEP code », *J. Phys.: Condens. Matter*, vol. 14, n° 11, p. 2717-2744, mars 2002, doi: 10.1088/0953-8984/14/11/301.
- [27] N. Brahim, « Construction et exploitation d'une base de données sur les semiconducteurs Magnétiques Dilués », Working Paper, 2011. Consulté le: juin 19, 2021. [En ligne]. Disponible sur: <http://dspace.univ-tlemcen.dz/handle/112/120>
- [28] M. A. Lajnef, M. B. Ayed, et C. Kolski, « Convergence possible des processus du data mining et de conception-évaluation d'IHM : adaptation du modèle en U », p. 4, 2005.
- [29] H. ZENASNI, « Etude théorique des propriétés Magnétiques, électroniques et structurales des alliages Heusler », PhD Thesis, 2013.
- [30] S. Tufféry, *Data mining et statistique décisionnelle: l'intelligence dans les bases de données*. Editions TECHNIP, 2005.
- [31] K. Pearson, « LIII. On lines and planes of closest fit to systems of points in space », *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 2, n° 11, p. 559-572, nov. 1901, doi: 10.1080/14786440109462720.
- [32] H. Hotelling, « ANALYSIS OF A COMPLEX OF STATISTICAL VARIABLES INTO PRINCIPAL COMPONENTS », p. 25.
- [33] P.-L. Gonzalez, « L'ANALYSE EN COMPOSANTES PRINCIPALES (A.C.P.) », p. 37.
- [34] B. Vincent, « Etude de données d'expression par Analyse en Composantes Principales Fonctionnelles », p. 5, 2002.
- [35] « [http:// www.cabannes.org/acp.pdf-France](http://www.cabannes.org/acp.pdf-France). L'Analyse en composantes principales. »
- [36] B. Escofier et J. Pagès, *Analyses factorielles simples et multiples: objectifs, méthodes et interprétation*. Paris: Dunod, 2008. Consulté le: juin 19, 2021. [En ligne]. Disponible sur:

<http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=304157>

- [37] L. Lebart, M. Piron, et A. Morineau, *Statistique exploratoire multidimensionnelle*. Paris: Dunod, 1995.
- [38] M. Boubchir et H. Aourag, « Materials genome project: The application of principal component analysis to the formability of perovskites and inverse perovskites », *Computational Condensed Matter*, vol. 24, p. e00495, sept. 2020, doi: 10.1016/j.cocom.2020.e00495.
- [39] F. J. Gotor, C. Real, M. J. Dianez, et J. M. Criado, « Relationships between the Texture and Structure of BaTiO<sub>3</sub> and Its Tetragonal  $\bar{A}$  Cubic Transition Enthalpy », p. 5.
- [40] M.-S. Park, J. Giim, S.-H. Park, Y. W. Lee, S. I. Lee, et E. J. Choi, « Physical properties of ZnCNi<sub>3</sub>: comparison with superconducting MgCNi<sub>3</sub> », *Supercond. Sci. Technol.*, vol. 17, n° 2, p. 274-277, févr. 2004, doi: 10.1088/0953-2048/17/2/007.
- [41] P. Tong, Y. P. Sun, X. B. Zhu, et W. H. Song, « Strong spin fluctuations and possible non-Fermi-liquid behavior in AlC<sub>2</sub>Ni<sub>3</sub> », *Phys. Rev. B*, vol. 74, n° 22, p. 224416, déc. 2006, doi: 10.1103/PhysRevB.74.224416.
- [42] J. P. Buban, H. Iddir, et S. Ögüt, « Structural and electronic properties of oxygen vacancies in cubic and antiferrodistortive phases of SrTiO<sub>3</sub> », *Phys. Rev. B*, vol. 69, n° 18, p. 180102, mai 2004, doi: 10.1103/PhysRevB.69.180102.
- [43] R. Wahl, D. Vogtenhuber, et G. Kresse, « SrTiO<sub>3</sub> and BaTiO<sub>3</sub> revisited using the projector augmented wave method: Performance of hybrid and semilocal functionals », *Phys. Rev. B*, vol. 78, n° 10, p. 104116, sept. 2008, doi: 10.1103/PhysRevB.78.104116.
- [44] C. M. I. Okoye, « Structural, elastic and electronic properties of new antiperovskite-type superconductor from first-principles », *Physica B: Condensed Matter*, vol. 405, n° 6, p. 1562-1570, mars 2010, doi: 10.1016/j.physb.2009.12.040.
- [45] C. M. I. Okoye, « Theoretical investigation of electronic structure and optical properties of paramagnetic non-oxide perovskite AlCNi<sub>3</sub> », *Solid State Communications*, vol. 136, n° 11-12, p. 605-610, déc. 2005, doi: 10.1016/j.ssc.2005.09.033.
- [46] « Elements, Atomic Radii and the Periodic Radii ». <http://www.crystallmaker.com/support/tutorials/atomic-radii/> (consulté le juill. 05, 2021).

- [47] R. L. Moreira et A. Dias, « Comment on “Prediction of lattice constant in cubic perovskites” », *Journal of Physics and Chemistry of Solids*, vol. 68, n° 8, p. 1617-1622, août 2007, doi: 10.1016/j.jpcs.2007.03.050.
- [48] W. Travis, E. N. K. Glover, H. Bronstein, D. O. Scanlon, et R. G. Palgrave, « On the application of the tolerance factor to inorganic and hybrid halide perovskites: a revised system », *Chem. Sci.*, vol. 7, n° 7, p. 4548-4556, juin 2016, doi: 10.1039/C5SC04845A.
- [49] M. Boubchir et H. Aourag, « Materials genome project: The application of principal component analysis to the formability of perovskites and inverse perovskites », *Computational Condensed Matter*, vol. 24, p. e00495, sept. 2020, doi: 10.1016/j.cocom.2020.e00495.

## Résumé :

Dans ce travail, nous avons utilisé une étude par simulation ab-initio, en se basant sur la DFT, pour calculer les propriétés structurales de 120 pérovskites afin de générer une base de données. Ensuite, nous avons utilisé la méthode statistique multi variée l'ACP pour étudier ces données, afin de déterminer quelles corrélations existent entre ces propriétés. Les résultats obtenus montrent qu'il existe des corrélations linéaires positives entre les paramètres ( $r_A$ ,  $r_B$ ,  $\mu$ ), et entre ( $a$ ,  $E_{coh}$ ,  $r_X$ ) et corrélations linéaires négatives entre le facteur de tolérance  $t$  et les paramètres ( $r_A$ ,  $r_B$ ,  $\mu$ ). Par ailleurs, les résultats de biplot nous a permis de déceler les composés de fortes et faibles propriétés.

**Mots clés :** pérovskite, DFT, Datamining, Analyse en composantes principale (ACP).

## Abstract:

In this work, we used an ab-initio simulation based on the DFT, to calculate the structural properties of 120 perovskites in order to generate a database. Then, we used the multivariate statistical method the PCA to study the calculated data, in order to determine the correlations between these properties. The obtained results show that there are positive linear correlations between the parameters ( $r_A$ ,  $r_B$ ,  $\mu$ ), and between ( $a$ ,  $E_{coh}$ ,  $r_X$ ) and negative linear correlations between the tolerance factor  $t$  and the parameters ( $r_A$ ,  $r_B$ ,  $\mu$ ). Furthermore, the biplot results allowed us to detect compounds with strong and weak properties.

**Key words:** perovskite, DFT, Datamining, Principal component analysis (PCA).

## الملخص:

قمنا في هذا العمل ، باستخدام دراسة محاكاة ab-initio واستنادًا إلى DFT ، لحساب الخصائص الهيكلية لـ 120 بيروفسكيت من أجل إنشاء قاعدة بيانات. بعد ذلك ، استخدمنا الطريقة الإحصائية متعددة المتغيرات PCA لدراسة الخصائص الهيكلية المحسوبة لهذه المواد ، من أجل تحديد الارتباطات الموجودة بين هذه الخصائص. تظهر النتائج التي تم الحصول عليها أن هناك ارتباطات خطية موجبة بين المتغيرات ( $r_A$  ،  $r_B$  ،  $\mu$ ) ، وبين ( $a$  ،  $E_{coh}$  ،  $r_X$ ) والارتباطات الخطية السلبية بين عامل التسامح  $t$  والمتغيرات ( $r_A$  ،  $r_B$  ،  $\mu$ ) . علاوة على ذلك ، سمحت لنا نتائج biplot باكتشاف المركبات ذات الخصائص القوية والضعيفة.

**الكلمات المفتاحية:** البيروفسكيت, نظرية الكثافة الدالية, التنقيب في البيانات, تحليل المركبات الرئيسية (ACP).