An Efficient Model

for

$sp$-Lone Pair Cations

Submitted by

Woongkyu Jee

for the degree of Doctor of Philosophy of the

University College London

Department of Chemistry

February 2023
Declaration

I, Woongkyu Jee, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.
Abstract

Lone pair cations have an $s^2\,p^0$ valence electron configuration as found in Sn(II), Pb(II) and Bi(III). Non-centrosymmetric external electric fields around such a cation stabilise the two valence electrons as a localised off-centre lone pair lobe formed by hybridisation of $s$ and $p$ type valence orbitals, that cause structural distortions in a wide class of material. Due to the exciting properties derived from the polarisation of the lone pair and the interest in potential applications, materials including lone pair cations have been attracting huge attention in the materials community.

As an approach to study materials including lone pair cations, computational methods have been widely employed, and electronic structure techniques are used most frequently. There is a deficit of physiochemically accurate and computationally affordable models for these materials within classical atomistic approaches.

To address this issue, we have developed an efficient model for simulating the lone pair cations, employing a simple concept from quantum mechanics of a single-electron molecular orbital to treat the electron lone pair. Throughout this thesis, we developed theoretical basis of the lone pair model, including mathematical and technical details used as recipes for the model implementation in our in-house software package, SLAM, which is freely available to download from the GitHub. For validation of our approach, we investigated the model properties, including high order electric multipoles and their polarisabilities, which were compared to their counterparts obtained using an electronic structure method, and our approach has turned out to be effective. Finally, the new lone pair model was employed to simulate nanoclusters of lead monoxide and difluoride as its application, and it was evidenced by the results that the lone pair model is efficient and accurate in predicting and reproducing the nanocluster atomic configurations, obtained using density functional theory.
Acknowledgements

I would like to thank so many people who have been staying with me throughout my PhD. The person who deserves the greatest thanks is my supervisor Professor Scott M. Woodley for giving me great help, insight and guidance for my 20’s project. I would not forget the last five years and half, especially your patience and support, even though sometimes I spread myself too thin and indulged in things out of the research.

I am also grateful to Dr. Alexey A. Sokol, research fellow in our group, for your help on designing mathematical and physical basis of the model developed during my PhD, and of course, the doughnuts and coffee you treated me all the time.

For the people who I had great fun together while working at UCL, Isa and Dong-Gi, my colleagues in the office and at the same time the closest friends I met in London. The fun we had really helped my work-life balance.

In my personal life, I also would like to thank my parents, Soo and Byeol, and my brother, the second Kyu, for their continued support both mental and financial, without them, I would not be able to complete this special task in my life. Lastly, I would like to thank someone who has been together with me, has great patience, encouraged me when I was frustrated, Min Jung, I really thank you from my heart.
Contents

Declaration .................................................................................................................................................. 2
Abstract .......................................................................................................................................................... 3
Acknowledgements ...................................................................................................................................... 4

Chapter 1

Introduction .................................................................................................................................................. 9

1.1 Stereo-chemically active sp-lone pairs ................................................................................................. 9
  1.1.1 Formation of stereo-active lone pair on cations ............................................................................. 10
  1.1.2 Outlook for lone pair cation model ................................................................................................. 13

1.2 Atomistic model and interatomic potential ......................................................................................... 14
  1.2.1 A brief discussion about classical modelling approaches ............................................................... 15

1.3 The model for lone pair cations ........................................................................................................... 18
  1.3.1 Modelling the valence state of lone pair cations ............................................................................ 20
  1.3.2 Ionic picture and model charge parameter .................................................................................... 21
  1.3.3 The core part of model cation ........................................................................................................ 22
  1.3.4 Model energy .................................................................................................................................. 23
  1.3.5 Summary and model assumptions ................................................................................................. 24

1.4 The Application of lone pair ion model ............................................................................................... 25

Chapter 2

Development of the lone pair model ......................................................................................................... 27

2.1 An isolated lone pair cation .................................................................................................................. 29

2.2 The basis set of the lone pair ion model ............................................................................................... 32
  2.2.1 The numeric radial wavefunction ................................................................................................. 33
2.2.2 Calculating the induced electric dipole .......................................................... 36

2.3 Interactions between ions and lone pair electrons ........................................ 39
   2.3.1 Electrostatic interactions between point charges and the lone pair electron density .......... 39
   2.3.2 The short-range potential between ions and the model electron lone pair .......... 47

2.4 Lone pair electrons interacting with lone pair densities .................................. 48
   2.4.1 Localised model orbital approach for lone pair density ..................................... 49
   2.4.2 Development of the Hamiltonian for a system including multiple lone pair cations .......... 50
   2.4.3 The electronic energy of the lone pair model .................................................. 54
   2.4.4 The self-consistent field algorithm for calculating model electronic energy ............... 56
   2.4.5 Convergence analysis of the model SCF algorithm .......................................... 59

2.5 The geometric derivatives of the lone pair model ........................................... 61
   2.5.1 Geometric derivatives of the model electronic energy ........................................ 63
   2.5.2 Development of a fixed-point iteration algorithm for the model derivative calculations .......... 65
   2.5.3 Convergence analysis of the algorithm calculating the model derivatives .................. 68
   2.5.4 Rotation of reference frames in the model derivatives calculations ....................... 73

2.6 Chapter summary .............................................................................................. 76

Chapter 3

Validation of the lone pair model ........................................................................... 78

3.1 Response of a lone pair cation to electric field ............................................. 79
   3.1.1 Electric dipole moment and dipole-polarisability of the lone pair ion model ............... 80
   3.1.2 Quadrupole polarisabilities of the lone pair model .............................................. 92
   3.1.3 Lone pair charge and onsite energy separation parameters .................................... 96

3.2 The effective valence state of a lone pair cation ............................................ 101
   3.2.1 Effective valence state of the lone pair ion model in comparison with DFT ............. 101
   3.2.2 Dipole moments and the relationship between the valence states of LPM and DFT ........ 107
3.2.3 Summary ........................................................................................................... 111

3.3 The effective potentials in the lone pair model ....................................................... 112

3.3.1 Pb(II) cation interacting with a point charge ....................................................... 112
3.3.2 Pb(II) cation interacting with a fluoride anion ..................................................... 117
3.3.3 Pb(II) cation interacting with another Pb(II) cation .......................................... 123

Chapter 4

Application of the New Model......................................................................................... 128

4.1 Structure prediction of nanoclusters ........................................................................... 129

4.2 Methodology ............................................................................................................. 131

4.2.1 Energy and Derivatives of the LPM ....................................................................... 132
4.2.2 Local optimisation ................................................................................................. 133
4.2.3 Global optimisation ............................................................................................... 135
4.2.4 How we assessed the LPM when used to predict nanoclusters ............................... 138

4.3 Nanoclusters of lead monoxide .................................................................................. 141

4.3.1 Structures of (PbO)\text{\textit{N}} clusters (\textit{N} = 6 - 12) ..................................................... 141
4.3.2 (PbO)\text{\textit{N}} LM clusters: comparison of SM and LPM with DFT ............................. 146
4.3.3 (PbO)\text{\textit{N}} GM clusters: second-order energies and their dipole moments ............... 151
4.3.4 Effect of stereo-active lone pair in nanocluster structures .................................... 153
4.3.5 Concluding remarks ............................................................................................. 160

4.4 Nanoclusters of lead fluoride ..................................................................................... 160

4.4.1 Structures of (PbF\text{\textit{2}})\text{\textit{N}} clusters (\textit{N} = 3 - 8) ............................................................. 161
4.4.2 (PbF\text{\textit{2}})\text{\textit{N}} LM clusters: comparison of SM and LPM with DFT ............................. 164
4.4.3 (PbF\text{\textit{2}})\text{\textit{N}} GM clusters: second-order energies and their dipole moments ............... 170
4.4.4 Transferability of LPM for simulating lone pair free cations ................................ 172
4.4.5 Concluding remarks ............................................................................................. 178

Chapter 5
Conclusions and Future Work ........................................................................................................ 181

5.1 Conclusions and summary of completed work ........................................................................ 181

5.2 Suggested future work ............................................................................................................. 185

Appendix A. Dipole-quadrupole and quadrupole-quadrupole polarisabilities of the lone pair model ........................................................................................................................................... 189

Appendix B. Comparing atomic configurations of two different nanoclusters ......................... 192

Appendix C. Parameters used in the study of lead monoxide nanoclusters ............................. 196

Appendix D. Parameters used in the study of lead fluoride nanoclusters ............................... 198

Appendix E. Software for simulating the lone pair model ......................................................... 200

References ..................................................................................................................................... 208

List of Figures ............................................................................................................................... 218

List of Tables ............................................................................................................................... 227
Chapter 1

Introduction

Stereo-chemically active \( sp \)-lone pairs are frequently found in the materials containing the post transition metals with the specific electronic configuration of \( d^{10}s^2p^0 \), which includes \( \text{Pb}^{(II)} \), \( \text{Sn}^{(II)} \) and \( \text{Bi}^{(III)} \). Under suitable conditions, applied electric fields, the lone pair electron density can distort the local environment of the atomic structure (including a change in coordination) about these cations. In an atomistic structure model, it is, therefore, important to include an energy term that provides the driving force which can reproduce these distortions.

1.1 Stereo-chemically active \( sp \)-lone pairs

\( \text{Pb}^{(II)} \), \( \text{Sn}^{(II)} \) and \( \text{Bi}^{(III)} \) have an electronic configuration of \( d^{10}s^2p^0 \) and their ionic compounds typically adopt non-centrosymmetric crystal structures, where the presence of \( sp \)-lone pairs have often caused a reduction in the crystal symmetry. Typical coordination polyhedral shapes – tetrahedral \([T]\), triangular bipyramidal \([TBP]\) and square planar \([SP]\) – adopted by a lone pair cation are summarised in figure 1.1, where the lone pairs are shown as blue pale lobes and labelled “E”. The lone pairs stabilise the \( T \), \( TBP \) and \( SP \) configurations, which would otherwise be trigonal planar,
see-saw and square planar, respectively, i.e. the lone pair drives the latter configurations to distort to the former. Due to the presence of lone pair electron densities, the binary oxide compounds $MO$, where $M = Pb^{2+}, Sn^{2+}$, form the litharge structure [2,3], which has square pyramid building units of $MO_4$ (see figure 1.1 [SP]) with base edges shared by neighbouring units to form the flat layers stacked along the [001] direction.

Alongside the interest in how atomic structures are distorted due to the presence of lone pair cations, there is also a wide range of important applications using materials containing lone pair cations. For example, SnO and PbO are promising candidate materials for photo conductive/voltaic devices [4–6], whereas $Bi_2O_3$ is used as ionic conductor in photovoltaic devices [7–9].

1.1.1 Formation of stereo-active lone pair on cations

In crystal lattice, the lone pair cations could be either stereo-active or inactive. The stereo-active lone pair cations are associated with distorted crystal structures. Considering that the binary oxides of alkaline earth metals adopt the highly symmetric NaCl rocksalt structure and have cations that do not have lone pair electrons, it is reasonable to ascribe the phase change between the rock-salt
and litharge structures to the effect of the lone pairs. Indeed, formation of stereo-active lone pairs in litharge structure has been rationalised by the stabilisation caused by the hybridisation of the occupied $s^2$ and empty $p^0$ valence orbitals localized on lone pair cations that lead to the on-site second order Jahn-Teller effect [10–12]. For instance, Sn$^{2+}$ is known to have a very large polarisability, which is attributed to internal excitation of $s \rightarrow p$ character from its $5s^2$ valence state giving rise to a stereo-active lone pair [13].

Sometimes, lone pair cations remain stereo-inactive, preserving the spherical symmetry of the $s$ orbital and, apparently, the lone pairs do not play structure-directing role. For example, the Pb$^{2+}$ cation in the fluorite structure of PbF$_2$ is stereo-inactive; the cation is coordinated to 8 anions and keeps central symmetry. Based on experimental ionic and thermal conductivity measurements it has been suggested that thermal lattice vibrations average the lone pair distribution around Pb$^{2+}$ cations in this system [14]. Such inactive lone pair could be also found in other materials, notably for chalcogenides, including PbS, which adopts the rock-salt structure rather than that of litharge PbO. The anomalous behaviour of active lone pairs was reviewed by Walsh et al, based on the electronic structure calculations and X-ray emission spectroscopy experiments [15]. According to the study, the preference of a lone pair for stereo-active behaviour is governed by effects of covalency, the extend of which is controlled by the relative energy levels of metal $s$ and ligand $p$ orbitals.

Furthermore, several experimental techniques focussing on the local structure, especially pair distribution function analysis from X-ray and neutron total scattering experiments, show significant local distortions in materials containing lone pair cations, including chalcogenides, bismuth pyrochlores and halide perovskites [16]. Similar experimental observations have also been reported for chalcogenide complexes of Pb$^{2+}$ [17,18]. Significantly, a first principle study for gas
phase SnO dimer has showed that the cation-anion interaction is purely “ionic”, i.e., no orbital substantial hybridisation has been found between Sn and O states and the polarisability of the cation proved to be independent of the anion orbital contributions [19].

The discussion above leads to a question if the stereo-active behaviour of lone pairs is determined not only by the ligand field effects, or the covalency, but also by the crystal field effects how, or in what geometry or which directions, charged particles, molecules or anions, are coordinating a lone pair cation. The answer to the question could be again found in the studies on gas phase molecules including lone pair cations. One of the simplest examples is the fluorides of lone pair cations, which are not stereo-active in condensed phase, related to their high symmetry, which does not offer sufficient space for a lone pair to be pronounced. However, the stereo-active lone pairs are realised in gas phase molecules of PbF$_2$, SnF$_2$ and BiF$_3$, as evidenced by their photoelectron valence-band spectra [20,21]. Furthermore, according to the first principle calculations on Pb$^{2+}$ including complexes from Cambridge Structural Database (CSD), the ionically bonded structures with lone pair are energetically more favoured (by ca. 8-12 kcal/mol) when they are stereo-active [22,23]. In figure 1.2, we compare two types of coordination, and in particular, complexes

---

**Figure 1.2** Schematic diagrams of (A) holodirected and (B) hemidirected coordination. With respect to the centred cation $M$, the black circles and blue arrows refer to the coordination sphere and ligands, respectively.
with a hemidirected coordination relaise active lone pairs – including all structures depicted in figure 1.1. On the other hand, complexes of Pb$^{4+}$ strongly prefer the holodirected coordination with less ionic character as the cation does not have residual electron pair on its $s$ orbital.

### 1.1.2 Outlook for lone pair cation model

As discussed so far, most of the computational studies that explore the structures, properties, and applications of materials including lone pair cations are heavily relying on electronic structure methods, e.g., density functional theory (DFT) [24]; there are many other cases, see Refs [25–30]. There is a lack of computationally affordable models for large scale simulations of materials with lone pair cations. Here, affordable implies the computational cost is of the level typically required by classical atomistic level of approaches.

In this work, we aim to develop an atomistic model, that can effectively describe system containing lone pair cations, which implies that other species, for instance anions in PbO and PbF$_2$, will also be modelled with atomistic approaches. An important aspect of the model is the charge distribution between positively charged cations and negatively charged anions. This description may seem contradicting the picture of of covalency as postulated in some literature and discussed in the previous section. However, there is still substantial evidence, supporting the role of crystal field effects in the stabilisation of stereo-active lone pairs, or $s \rightarrow p$ excitation in internal electric field generated by neighbouring anions. More detailed discussion about this choice will be delivered in section 1.3.2 where we introduce the model charge parameters.
1.2 Atomistic model and interatomic potential

In the field of materials research, classical atomistic methods often are a viable alternative to electronic structure methods, and they can be used to explore materials with lone pair species. In conventional atomistic approaches to modelling ionic materials, ions are typically represented by the rigid-ion model [31] or the polarisable ion models: the shell model [32] by Dick and Overhauser, and the point ion model by Wilson and Madden and their co-workers [33,34]. The rigid-ion model is the simplest model that describes an ion as a point charge, while the latters including ionic polarisabilities: (i) the shell model deals with two point charges, the core and the shell, connected to each other by a harmonic “spring, separation” of the core and the spring simulates the effect of electronic dipole polarisation and (ii) the point ion model includes the effects of both dipole and quadrupole polarisations, described by the electric multipoles.

Amongst the polarisable ion models, the shell model, due to its simplicity, has been widely used for the materials simulations. With the background, the shell model could be employed as an approach to represent the lone pair species, considering that the model contains the dipole-polarisation, which is mainly presented in the lone pair species. However, in vacuum, the shell model shows only a linear response to the external electric field due to its constant intrinsic polarisability (see equation 1.9 – the dipole polarisability of the shell model ion in vacuum), which in some cases may prove to be only a crude approximation. This limitation does not usually cause problems when the shell model is employed to model a crystal lattice, since its in-crystal dipole polarisability becomes non-linear due to the augmenting effect, induced by the local environment (e.g., the potentials effectively modifying the spring) of shell model ions. Furthermore, for a symmetrical lattice environment, where electric fields are relatively low, its usefulness is evident from many studies which employ the shell model [35–40]. However, when the system becomes strongly
asymmetrical and includes highly polarisable species like lone pair cations – for instance, by Tessman et al, the experimentally measured polarisability of Pb$^{2+}$ is 4.9 Å$^3$ compared to other non-lone pair species like Ba$^{2+}$ 2.5 Å$^3$ and Sr$^{2+}$ 1.6 Å$^3$ [41], and other experiments also showed higher polarisability of lone pair cations [42,43] – the shell model polarisability may not be sufficiently accurate anymore. Moreover, some cations of transition metals show polarisation, which is not properly represented by the shell model [44,45].

To tackle this problem, there is a need for developing a new ion model for lone pair cations, which is the aim of our work. The new ion model we wish to develop (i) should give a better physical description than the shell model, and (ii) be computationally more affordable than the full electronic structure methods. Focusing on the lone pair cations first, we introduce the basic concept of molecular orbitals to represent the lone pair, more detail will be delivered in section 1.3. To make our model computationally affordable, we retain the atomistic approach in our treatment of the non-lone pair species, using the classical shell model or the rigid-ion model as appropriate.

1.2.1 A brief discussion about classical modelling approaches

In this section, we recall the main tenets of the atomistic approach using the method of interatomic potentials. In the classical description, for $N$ atoms, where each atom, $i$, is located at $r_i$, the total energy, $E_{\text{total}}$, is typically expressed as function of atomic positions:

$$E_{\text{total}} = E(r_1, r_2, ... r_i ... r_N), \quad (1.1)$$

then it is possible to calculate the atomic force, $F_i$, which is simply equal in magnitude and opposite in sign to the first derivatives of energy with respect to coordinates of atom $i$:
Using standard optimisation techniques, it is possible to relax an initial atomic structure to a local energy minimum. Moreover, using global optimisation algorithms, it is possible to predict local energy minima structures, and physical properties at these local minima from further derivatives of $E_{total}$.

For charged ions, the energy function of atomic coordinates in equation 1.1 is commonly approximated as a sum of pairwise potentials that depend only on the interatomic distances $r_{ij}$:

\[ E_{total} = \sum_{i,j} \frac{q_i q_j}{r_{ij}} + \sum_{i,j} V(r_{ij}). \]  

Equation 1.3 states that the first two terms on the right side refer to the Coulomb interactions (Coulomb’s constant is omitted for brevity but is used in practical calculations) and short-range (typically repulsive) interactions where the summations over $i$ and $j$ over unique ion pairs with charges of $q_i$ and $q_j$, respectively. For some ionic materials, a third term, known as a three-body term, is yet required, however, it is also often neglected.

A wide range of different forms are used for the short-range potential in equation 1.3, but, perhaps, one of the popular forms is the following:

\[ V(r_{ij}) = D_{ij} \exp \left( -r_{ij}/\rho_{ij} \right) - \frac{C_{ij}}{r_{ij}^6}, \]

known as the Buckingham potential where $D_{ij}, \rho_{ij}$ and $C_{ij}$ are species dependent parameters, and the first and second terms represent the Pauli repulsion, or exchange interaction at short-range interionic distances and the dispersion interaction at long-range distances, respectively [46].
Parameters of interatomic potentials should be appropriately tuned first to a target system [47]. By substituting equation 1.4 into 1.3, one arrives at the Coulomb-Buckingham potential, which is ideal for modelling rigid spherical ions since each term is simply a function of interatomic distance without polarisation of ions:

$$E_{total} = \sum_{i,j} \left[ \frac{q_i q_j}{r_{ij}} + D_{ij} \exp \left( -\frac{r_{ij}}{\rho_{ij}} \right) - \frac{C_{ij}}{r_{ij}^6} \right].$$

(1.5)

The rigid ion model is appealing due its simplicity but is missing important physical features, for instance halide ions in electric fields can readily polarise as their valence electrons displace breaking the spherical symmetry about the nucleus (and core electrons) and forming a dipole.

To overcome this deficiency, the polarisable shell model alluded to above is widely employed to reproduce the polarisation of ions in external fields and whose internal potential energy is described by:

$$E = \frac{1}{2} k r_{cs}^2$$

(1.6)

where $k$ is the harmonic spring constant, which is species dependent; $r_{cs}$ is a distance between the ionic core and its shell. The core and shell charges interact Coulombically with surrounding ions, but not with each other, which is accounted for by equation 1.5, whereas the repulsive term of the Buckingham potential typically acts only upon the shells of ions.

Considering the case of uniform external electric field $\mathbf{F}$, the shell is displaced creating a dipole. Assuming the shell is at the equilibrium position while the position of the core is fixed, the energy of the shell is:

$$E = -\frac{1}{2} \frac{q^2}{k} \mathbf{F}^2,$$

(1.7)
and its minus first derivative with respect to the field gives the dipole moment,

$$\mu_i = \frac{q^2}{k} F_i$$  \hspace{1cm} (1.8)

whilst the minus second derivative gives the dipole-polarisability,

$$\alpha_{ij} = \frac{q^2}{k} \delta_{ij}$$  \hspace{1cm} (1.9)

where the subscripts $i$ and $j$ refer to the Cartesian coordinate components \(\{x, y, z\}\).

As pointed out earlier, the shell model polarisability in equation 1.9 is constant in vacuum\(^1\), i.e., independent of the external field, and therefore the response of the model is linear. In fact, there is a variation of the shell model, so called \(k4\) shell model, which partially solves the deficiency of the harmonic shell model by the addition of anharmonicity to the spring [48]. To relate our new model with the behaviour of the harmonic and \(k4\) shell models in external fields, their physical properties including energy, dipolar moment and polarisability will be compared with the lone pair model in chapter 3.

### 1.3 The model for lone pair cations

Considering lone pair behaviour in other chemical systems, beyond those containing lone pair cations, there is a well-known problem of modelling water and ammonia molecules where non-bonding electron pairs on the oxygen and nitrogen atoms spatially occupying empty space outside the atom centre adopting asymmetrical orbital shapes. The water molecule has been the focus of

---

\(^1\) As mentioned earlier in section 1.2 in the presence of other surrounding or coordinating species, e.g., in the crystalline bulk environment, the polarisability of the shell model can be adaptively reduced, i.e., damped.
especial attention due to its role as solvent in many fields of inorganic chemistry and live science but also as a pure substance of interest with multiple condensed phases, all of which rely on molecular dynamics (MD) simulations. For example, TIP5P (and also its many variations) water model employs the point dipole approach to describe the lone pairs [49]. In the model, the point dipoles are created by dummy or ghost charges on the geometrically fixed sites where lone pairs reside. Further, there have been many other approaches, used in MD simulations in the same spirit [50,51], including those based on the classical Drude oscillator for describing lone pairs on atoms in organic molecules [52]. However, to our knowledge, none of these models so far have been employed for modelling lone pair cations.

Turning to systems with lone pair cations, Bellac et al (1995) investigated bulk PbO and SnO in litharge structure, employing a polarisable point dipole method to describe lone pairs on those cations [53]. The shell model has also been employed to study the crystal structure of α-PbO [54] and oxygen diffusion in δ-Bi₂O₃ [55]. Thus, compared to other systems or materials there is obvious paucity of atomistic-level studies in systems containing lone pair cations, evidently due to a lack of an efficient and physically transparent model for lone pair cations.

We want the new model to capture possible deformation of a lone pair from its simple spherical form (for an isolated cation where the lone pair occupies a pure s state) to a deformed shape (the lone pair becoming stereo-active as an sp-hybridisation occurs). Therefore, the point dipole method would not be appropriate since we do not want to put a permanent dipole on a cation. The shell model could be an alternative choice, however, as discussed in section 1.2, we are aiming to capture more sophisticated ionic polarisability with the model. Therefore, we decided to employ a quantum-mechanical molecular orbital approach to represent the lone pair, or the valence state of cations, and related model assumptions will be formulated and discussed below.
1.3.1 Modelling the valence state of lone pair cations

At the beginning of this chapter, it has been pointed out that the lone pair species share the same electronic configuration of $d^{10}s^2p^0$, and that the lone pair is formed by the hybridisation of the occupied s and the empty p orbitals. The inner $d^{10}$ shell is relatively inert to such hybridisation and thus can be treated as a spherical entity, i.e., the individual angular dependencies of $d$ orbitals are cancelled out since they are fully occupied.

We treat a lone pair cation as two different objects: (i) the core, representing the nucleus and core electrons, including the $d^{10}$ shell, treated classically, and (ii) the effective valence state, including electrons that form a lone pair, spanning the occupied s and the empty three p orbitals, which is an eigenstate of an onsite model Hamiltonian. The choice of the model Hamiltonian will be discussed in the next chapter, from section 2.1 to 2.4. The two constituents of the lone pair model are somewhat analogous to those of the core-shell model where the core remains spherical and the valence electrons are represented by the polarisable shell. However, instead of the point charge for the valence electrons, we introduced a single electron molecular orbital as a superposition of s and p atomic functions (a linear combination of atomic orbitals, which is commonly employed by ab initio and semi-empirical approaches) that is occupied by the lone pair.

The rationale for employing a single molecular orbital is that the two electrons are always paired in the ground state of the electron lone pair so that the spin of valence electrons does not need to be included. Furthermore, the minimum number of basis functions is chosen to save the computational effort whilst reproducing appropriately the major lone pair electronic effects.

Using Dirac’s notations, the model molecular orbital can be expressed as:

$$ |\psi\rangle = c_s |s\rangle + c_{p_x} |p_x\rangle + c_{p_y} |p_y\rangle + c_{p_z} |p_z\rangle $$  \hspace{1cm} (1.9)
where \( c_\mu \) with subscript \( \mu = s, p_x, p_y \) and \( p_z \) are the molecular orbital coefficients that following the variational principle minimise the energy of the lone pair, \( \langle \psi | \hat{H} | \psi \rangle \). Hence, for any trial wave function, \( | \psi^{\text{trial}} \rangle \), defined on the same space of the model molecular orbitals parameterised by \{\( c_s, c_{p_x}, c_{p_y}, c_{p_z} \)\} coefficients, the corresponding energy, \( E^{\text{trial}} \), is subject to the following inequality:

\[
E^{\text{trial}} = \langle \psi^{\text{trial}} | \hat{H} | \psi^{\text{trial}} \rangle \geq E^{\text{gs}}
\]

where \( E^{\text{gs}} \) is the exact, or ground state eigenvalue of the given model Hamiltonian, \( \hat{H} \). This Hamiltonian accounts for (i) the internal, or onsite interactions of the lone pair electrons, including the potential energy contributions from the cation nucleus and core electrons, the electrons’ kinetic energy, and electron-electron repulsion; (ii) the interactions with external ions, including non-lone-pair species and other lone pairs. More details will be discussed in chapter 2.

### 1.3.2 Ionic picture and model charge parameter

Throughout this thesis, charge \( Z \) is assigned to the core and charge \( \omega \) to the valence electrons, which are the model parameters. The sum of these charges, \((Z + \omega) = +2\) and \(+3\) for \( \text{Pb}^{2+} \) and \( \text{Bi}^{3+} \), respectively (for brevity, the units of charge, e, is omitted and throughout this thesis), i.e., the net charges of model cations are represented by their formal charges.

As discussed in section 1.1.2, our employment of formal charges implies the model description is under the ionic picture, and charge density deformation due to the interaction between cations and anions (deformation of heteropolar bonds in covalent picture [56]) in asymmetric environments is explained by polarisation, realised by creation of onsite dipoles on each ion. Further, as discussed in section 1.1.1, our choice of the ionic picture is supported by results of \emph{ab initio} calculations on excited states of \( \text{Pb}^{2+} \) complexes and an \( \text{SnO} \) molecule. Focusing on the cation in
external field, for example provided by interactions in a PbO molecule, the formation of lone pair involves an *sp*-hybridisation, or polarisation, which can be interpreted as the “vertical excitation”, or onsite charge transfer of $s \to p$ character, which is induced by electric field generated by the anion. In our model, such process of *sp*-hybridisation will be realised, for instance, by the displacement of negatively charged shell of anion to the core site of cation due to its positive charge, $Z$, resulting in the deformation of lone pair density.

Since the model employs a molecular orbital to represent the effective valence state, or lone pair, one might consider the charge of lone pair, $\omega$, should be -2, i.e., the formal charge of a lone pair. In practice, however, instead of using -2, a slightly lower by magnitude partial charge between -2 and -1 turns out to be effective. Such a shift of a fraction of the charge from the lone pair to the core allows us to keep the model intrinsically consistent by accounting for the electron-electron repulsion within the lone pair implicitly. More details will be given in section 3.2.1 based on comparison of our model with the Hartree approach to many-electron systems.

The validity of the ionic description chosen for the lone pair model will be eventually assessed in its applications, as the most practical application of the concept of ionicity concerns the development of interatomic potentials [57]. In the relevant applications, we have modelled the nanoclusters of PbO and PbF$_2$, results of which will be delivered in chapter 4, where we assess the model performance by reproduction of characteristic nanocluster configurations and their dipole moments obtained with DFT calculations.

### 1.3.3 The core part of model cation

Finally, in our model, we assume that the core not only carries charge $Z$ but it also has size, which can be simply modelled by adding a short-range repulsive interaction between the core and other
ions as described by equation 1.4. Such interactions are often omitted in atomistic simulations of lighter, smaller ions, but cannot be neglected for heavier lone-pair cations. The core of a lone pair cation and other non-lone-pair species (represented by the rigid-ion or polarisable-shell models) will be treated as molecular mechanical (MM) objects. Their interactions can be represented by the Coulomb-Buckingham potential, see equation 1.5. The rationale of introducing rigid core par on lone pair cations is justified by the presence of nucleus and particularly the core electrons, also occupying the inert $d^{10}$ states.

1.3.4 Model energy

Consequently, for a system including a number of lone pair cations and non-lone-pair ions, the energy $E_{total}$ can be separated into two different terms, $E_{MM}^{total}$ and $E_{LP}^{total}$ where the former term represents the energy contribution by molecular mechanical (classical) objects, including the non-lone-pair species and the cores of the lone pair cations while the latter term represents the quantum mechanical energy contribution by the lone pairs. Thus, the total energy of a system in the lone pair ionic model can be expressed as:

$$E_{total} = E_{MM}^{total} + E_{LP}^{total} = \sum_{i,j} \left[ \frac{q_i q_j}{r_{ij}} + D_{ij} \exp \left( -r_{ij}/\rho_{ij} \right) - \frac{C_{ij}}{r_{ij}^6} \right] + \sum_{\alpha} \langle \psi_{\alpha} | \hat{H}_\alpha | \psi_{\alpha} \rangle \quad (1.11)$$

where the summation over $\alpha$ runs over the lone pairs (or lone pair cations), and $\langle \psi_{\alpha} | \hat{H}_\alpha | \psi_{\alpha} \rangle$ is the energy of the $\alpha^{th}$ lone pair with the Hamiltonian, $\hat{H}_\alpha$. 
1.3.5 Summary and model assumptions

So far in this section, we have discussed what is new in our lone pair ionic model and what will be borrowed from pre-existing atomistic methods. For clarity, to avoid any possible confusion throughout the remainder of this thesis, the main assumptions of the model are listed below:

- A single electron molecular orbital is employed to represent the lone pair where the molecular orbital is a linear combination of one $s$ and three Cartesian $p$ atomic orbitals. In total, there are four basis functions, which are chosen as the minimum number necessary to capture the lone pair polarisation effect, whilst minimising the required computational cost.

- A lone pair cation is separated into two objects: (i) the core and (ii) the valence electrons (i.e., the lone pair), carrying charges $Z$ and $\omega$, respectively. As formulated in section 1.3.2, model is under the picture of ionicity, therefore the sum of those two charges is equal to the formal charge of a lone pair cation, e.g., $Z + \omega = +2$ for Pb$^{+2}$. Same scheme is also applied to describe non-lone pair ions, and the interionic charge transfer effect will be modelled as polarisation.

- To represent the effective valence state of lone pair cations, the charge carried by model molecular orbital, $\omega$, will be treated as a partial charge, instead of employing its nominal value of -2. This treatment enable the model to implicitly include the electron-electron repulsion in the valence state.

- The Buckingham potential is used to represent the core size effect in the interactions of a lone pair cation with other ions. In other words, the core part is treated as a rigid ion, which for describing the nucleus and core electrons, including relatively inert $d^{10}$ electrons on a lone pair cation.
Based on these ideas, the development process for the lone pair ionic model is presented in chapter 2. Further validation with physical rationales of the model are delivered in chapter 3.

### 1.4 The Application of lone pair ion model

The lone pair ionic model developed in this work has been tested by tackling a challenging problem of nanocluster structure prediction for the compounds including lone-pair cations [58]. Here, the choice of nanocluster systems for the model application rather than, perhaps, the more obvious crystalline bulk is determined by the much higher structural diversity of the former. In general, a nanocluster has also many more degrees of freedom defining its atomic configuration compared to bulk. We can exploit the power of our model to select most probable structures amongst a great number of structural configurations, i.e., much fewer structures of low energy could be obtained by the model and verify our choice by comparing structures and energies with DFT results.

Through comparison of large numbers of diverse viable atomic structures generated using our lone pair ionic model with their DFT counterparts, we establish veracity of the proposed model. Here, since the DFT energy landscape (i.e., the atomic structures and energies) is not exactly identical with the model energy landscape, our objective is to evaluate how well the newly developed lone pair model approximates the landscape generated by the higher-level theory.

Furthermore, unlike dense systems, ions in a nanocluster are mostly exposed to the surface. Such an environment is also more appropriate and challenging for an investigation of the high polarisability of the lone pair cations (since the stabilisation of lone pairs is not restricted as much as in the bulk environment). Therefore, we will investigate the influence of lone pair formation on the structural distortions.
In this work, we have targeted two different chemical systems, nanoclusters of lead monoxide and lead fluoride, which have been finding various applications as nanomaterials, see references [59,60] for PbO and for PbF₂ [61–63]. Analysing results, we have mainly concentrated on validating our model against DFT, but the polarisable shell model was also employed to investigate the differences caused by using different descriptions of the Pb²⁺ cation. The relevant details, methodology, results and analyses of the nanocluster structure prediction are provided in chapter 4.
Chapter 2

Development of the lone pair model

This chapter is focused on the development of the lone pair ion model, starting from an isolated lone pair cation in vacuum. The model is then enhanced to include an electric field generated from other ions. Thus, the target is able to simulate a lone pair ion with a system of ions that are themselves modelled using more traditional, or standard interatomic potentials. Finally, the model will include more than one lone pair ion in a system to fully describe the interaction occurring between lone pairs. The major objectives of the section are (i) to construct an interatomic potential that can capture the effects resulting from the creation of \( sp \)-lone pairs; (ii) to enable the extraction of electronic information captured by the model; and (iii) to provide sufficient information for the reader to gain an insight into the benefits of using this model, understand how to implement and possibly extend the derived model.

The ground state of lone pair. As discussed in section 1.3, the formation of lone pair electron density is due to the hybridisation of the valence states, i.e. mixing of the initially occupied \( s \) state and the three empty \( p \) states. The model divides the lone-pair cation into two entities: (i) the valence states; and (ii) the core part of a lone pair cation, which includes the nucleus and the core
electrons. The latter is treated as a rigid ion with charge $Z$. The former is treated using a single molecular orbital approach to describe the lone pair with charge $\omega$. The total charge of lone pair cation is simply the sum of these two and is constrained to be the formal charge of a lone pair cation. The valence state of the lone pair cation is modelled as a single electron molecular orbital:

$$\ket{\psi} = c_s \ket{s} + c_{p_x} \ket{p_x} + c_{p_y} \ket{p_y} + c_{p_z} \ket{p_z} \quad (2.1.1)$$

where $c_\mu$ are the molecular orbital coefficients for $\mu = s, p_x, p_y$ and $p_z$. For the Hamiltonian of the lone pair electrons, $\hat{H}$, the energy of lone pair is:

$$E_{LP} = \langle \psi | \hat{H} | \psi \rangle . \quad (2.1.2)$$

Substituting equation 2.1.1 into 2.1.2 we obtain,

$$E_{LP} = \sum_{\mu,\nu} c_\mu c_\nu \langle \mu | \hat{H} | \nu \rangle = \sum_{\mu,\nu} c_\mu c_\nu H_{\mu\nu} , \quad (2.1.3)$$

where $\mu$ and $\nu$ runs over the basis functions, $s, p_x, p_y$ and $p_z$. In order to determine the ground state energy of the lone pair, first it is necessary to calculate the Hamiltonian matrix whose elements are $H_{\mu\nu}$. Since we choose to employ a basis set of four functions, there are $4 \times 4$ matrix elements. The variational principle (Rayleigh-Ritz approximation) implies that the lowest energy combination of the trial function will provide the best approximation of the ground state that can be achieved using the chosen trial function. Here, the eigenvalues are the energies of the four electronic valence states and their corresponding eigenvectors are represented by the molecular orbital coefficients of each state. Accordingly, the lowest energy eigenvalue and corresponding eigenvector is taken as the ground state energy and the coefficients for the corresponding ground state molecular orbital.
2.1 An isolated lone pair cation

A single lone pair cation in vacuum is treated as a four-level system, with the lone pair electrons occupying the ground state $s$ atomic orbital and three, higher energy unoccupied degenerate $p$ atomic orbitals as shown schematically in figure 2.1(a) – more precisely, the energy difference is due to the effects of shielding and penetration, e.g. the electrons in the $s$ orbital penetrates more than the $p$ does. Therefore the $s$ state is more tightly bound to the nucleus and lower in energy.

To represent such a four-level system, we define is energy in state $\sigma$ as $E_\sigma = \langle \sigma |\hat{H}|\sigma \rangle$ where $\hat{H}$ refers to the Hamiltonian of electron lone pair in vacuum and due to the energy difference between the $s$ and $p$ states, the following relation holds:

$$E_s < E_p, \text{ or } \langle s |\hat{H}|s \rangle < \langle p_\mu |\hat{H}|p_\mu \rangle \text{ for } \mu = x, y \text{ and } z. \quad (2.1.4)$$

In *ab initio* theories, the Hamiltonian, $\hat{H}$, for an isolated electron lone pair includes the kinetic energy ($\hat{T}$); the potential energy in the field of cation’s nucleus and core electrons ($\hat{V}$); and the electron-electron interaction term ($\hat{j}$) since there are two electrons in a lone pair. Therefore, the explicit expression for the Hamiltonian is $\hat{H} = \hat{T} + \hat{V} - \hat{j}$. Following our model assumptions, we parameterise the effect of these terms along with introducing one-electron molecular orbital. A fitting parameter is chosen to absorb all these complications, as discussed later in section 3.1.3 and 3.2.1. That is, we choose to define this parameter, $\lambda$, as the energy difference between states $s$ and $p$:

$$\lambda = E_p - E_s = \langle p_\mu |\hat{H}|p_\mu \rangle - \langle s |\hat{H}|s \rangle. \quad (2.1.5)$$

We are interested in calculating energy differences and not the absolute energy values of $E_s$ and $E_p$, therefore, we define the integrals in equation 2.1.5 as:
\[ \langle s | \hat{H}^0 | s \rangle = 0 \quad \text{and} \quad \langle p_\mu | \hat{H}^0 | p_\mu \rangle = \lambda, \quad (2.1.6) \]

and the contribution of the so-called onsite Hamiltonian can, therefore, be summarised as:

\[ H^0_{\mu\nu} \equiv \langle \mu | \hat{H}^0 | \nu \rangle = \begin{cases} \lambda, & \text{if } \mu = \nu = p_x, p_y \text{ and } p_z \\ 0, & \text{otherwise}. \end{cases} \quad (2.1.7) \]

The square matrix \( H^0_{\mu\nu} \) is, therefore, diagonal, with diagonal elements 0, \( \lambda \), \( \lambda \), and \( \lambda \) (assuming the order of the rows and columns are \( s, p_x, p_y \) and then \( p_z \)). In this condition, there is no need to diagonalise the matrix: the ground state is the pure \( s \) atomic orbital with an energy of zero, whereas the other eigenstates are the degenerate \( p_x, p_y \) and \( p_z \) atomic orbitals with energy \( \lambda \).

![Energy diagrams](image)

**Figure 2.1** Energy diagrams of the valence states for an isolated lone pair cation \( M \) (e.g., \( M = \text{Sn or Pb} \)) in vacuum: (a) shows two electrons occupying the \( s \) state is occupied by lone pair electrons and the three \( p \) states left empty after the 2\(^{nd} \) ionisation occurs; (b) shows one electron occupying the \( s \) state after the 3\(^{rd} \) ionisation occurs where the energy difference between \( s \) and \( p \) states is marked with a Greek letter \( \lambda \).

The energy difference \( \lambda \) is a parameter of the lone pair model and is dependent on the species of the lone pair cation. The value of \( \lambda \), employed in the simulations that are described in a later chapter, were referred to the observed ionisation energies of the respective lone pair cations (slightly
revised in the model applications); see energy diagrams in figure 2.1. For example, consider modelling Pb\(^{2+}\), the second and third ionisation energies, IE\(_2\) and IE\(_3\), are,

\[
Pb^+ + IE_2 \ (15.03 \text{ eV}) \rightarrow Pb^{2+} + e
\]
\[
Pb^{2+} + IE_3 \ (31.94 \text{ eV}) \rightarrow Pb^{3+} + e
\]

where the values in the parentheses are experimentally observed taken from the reference [64]. From these ionisation expressions, we can determine the value of \(\lambda\) for the Pb\(^{2+}\) cation as \(\lambda_{\text{Pb}} = 31.94 \text{ eV} - 15.03 \text{ eV} = 16.91 \text{ eV}\); the other values of \(\lambda_M\) (\(M = \text{Sn and Bi}\)) are summarised in table 3.1 in the next chapter.

Strictly speaking, in \textit{ab initio} theories, the given diagrams do not represent the energy differences between the lower \(s^2\) and the upper degenerate \(p^0\) of the lone pair electron configuration, since they do not reflect the complications of electron-electron interactions. However, the approach is correct in the single electron (i.e., non-interacting electron) picture and is also instructive to assess the scale of the model parameter, \(\lambda\); a more detailed analysis of the parameterisation will be discussed in the next chapter, see section 3.1.3.

An alternative approach to determine the value of \(\lambda\) is to use the Koopman’s theorem [65,66] and electronic structure calculations, i.e. compute the energy difference between the HOMO and LUMO states where the technique has been employed in other studies [67–69]. The result would of course depend on the quality of the chosen application of electronic structure, e.g., choice of density functional and basis set.

In summary, the lone pair model requires fitting a parameter \(\lambda\), which accounts for the energy difference between the low energy state \(s\) and the three high energy degenerate \(p\) states. We have investigated various \(\lambda\) parameters in order to find a suitable potential, which can be used in the
model applications. As a result, we found that the promising $\lambda$ parameters exist near the range of experimentally observed ionisation energy, and further details will be discussed in chapter 3.

### 2.2 The basis set of the lone pair ion model

In the previous section the isolated lone pair ion in a vacuum was introduced. Without an external electric field, the lone pair can be represented by the pure $s$ atomic orbital, which, using the notation already introduced, means that $c_s = 1$ and $c_\mu = 0$ for $\mu = p_x, p_y$ and $p_z$ (see equation 2.1.1). Whereas when an electric field is applied, e.g., that generated by the charge of adjacent ions, the spherical symmetry of lone pair density is broken as this minimises the total energy, i.e., being stereo-active. Mathematically, the shape of the polarised, or distorted lone pair density is described by including contributions of $p$ atomic orbitals to the ground state, i.e. there is at least one non-zero coefficient for the $p$ atomic orbitals, $c_\mu \neq 0$ for $\mu = p_x, p_y$ or $p_z$.

For the lone pair model, a dipole is created when some of the charge of the lone pair electrons move from the $s$ to one or more of the $p$ atomic orbitals. Using spherical polar coordinates, each atomic orbital is represented as a product of a radial function and an angular function. With the lone pair ion (position of the core) centred at the origin of the basis-functions that are employed to describe the charge density of the lone pair electrons can be written as:

\[
\begin{align*}
    s &= R_s(r)\Gamma_s(\theta, \phi) \\
    p_x &= R_p(r)\Gamma_x(\theta, \phi) \\
    p_y &= R_p(r)\Gamma_y(\theta, \phi) \\
    p_z &= R_p(r)\Gamma_z(\theta, \phi)
\end{align*}
\]  

(2.2.1)
where we have decided to employ different normalised radial functions, $R_s(r)$ and $R_p(r)$, for $s$ and $p$ atomic orbitals. The real form of angular terms are frequently used in quantum chemistry \[70\], and they are:

$$
\Gamma_s = \frac{1}{2} \left( \frac{1}{\pi} \right)^{1/2}
$$

$$
\Gamma_x(\theta, \phi) = \left( \frac{3}{4\pi} \right)^{1/2} \sin\theta \cos\phi
$$

$$
\Gamma_y(\theta, \phi) = \left( \frac{3}{4\pi} \right)^{1/2} \sin\theta \sin\phi
$$

$$
\Gamma_z(\theta, \phi) = \left( \frac{3}{4\pi} \right)^{1/2} \cos\theta
$$

In equation 2.2.2, the angular term of the $s$ atomic orbital, $\Gamma_s$, is a constant, i.e. the orbital is purely radial or spherical, whereas the angular contributions of the $p$ orbitals provide the directionality and the anisotropic deformation of the lone pair electron density when the polarisation occurs.

**2.2.1 The numeric radial wavefunction**

For the radial parts of the model basis set, the radial wavefunctions of DFT software FHI-AIMS (using numeric atom-centred orbitals) code \[71\], the package default “Tier-2” basis functions, were reproduced in the model. In detail, the $s$ and $p$ radial wavefunctions of Pb$^{2+}$ cation, for their use in the lone pair model, the valence orbitals, ns and np of the Pb(n=6) atom, were taken from the code, and as for other species, Bi$^{3+}$ (n=6) – those are different to Pb – and Sn$^{2+}$ (n=5).

At the beginning, it was tempting to simply employ hydrogenic wavefunctions, given it is desirable to employ differentiable analytical expressions. However, unlike the hydrogen atom, the lone pair cation has two outer electrons. Even though these electrons have the same principal
quantum number, their interaction will lead to change in the radial distribution of their charge density. Technically it is possible to tune the size of hydrogenic wavefunctions, however it requires additional model parameters, and therefore this direction of development was not pursued.

Consequently, the implemented radial wavefunctions of the \( s \) and \( p \) atomic orbitals are presented in figure 2.2. In practice, the radial functions in the lone pair model are represented as cubic polynomials.
splines, which guarantees the continuity and twice differentiability and such numeric treatment of radial functions is also often employed in other studies [72,73]. Being technical, to obtain the spline radial wavefunctions, the datapoints in figure 2.2 were first sampled from the FHI-AIMS code. By default, the code provides its radial wavefunctions logarithmically, which is appropriate to secure the accuracy of radial wavefunctions, especially in the region near zero distance where the radial wavefunctions show oscillating profiles. Afterwards, we pre-processed the data by re-sampling those of each one for every interval of 7, which is for the purpose of reducing the number of integrations required in the lone pair model, i.e. to save the computational expense.

![Diagram of interpolated radial wavefunction](image)

**Figure 2.3** Schematic diagram of an interpolated radial wavefunction, $R(r)$. The red dots are known values of $R$ at discrete points $r = r_{\theta}$ and the black solid lines between nearest neighbour points are the functions $R_{\theta}(r)$ for the interval $[r_{\theta}, r_{\theta+1}]$. Assuming the functions $R_{\theta}(r)$ are cubic polynomials, the entire function is twice differential at their boundaries.

Say, the sampled datapoints are expressed as $r_k$ ($k = 1, 2, ..., k_{\text{max}}$), to decide the right-end cut-off, the $k_{\text{max}}$ was chosen for the integer when the radial wavefunction becomes below $10^{-8}$, i.e. beyond $r = r_{k_{\text{max}}}$ the radial functions are set to zero. Lastly, the clamped cubic-spline algorithm [74] was applied to develop the model radial wavefunctions. The reliability of the approach was

---

2 From the first datapoint $r_1 = 0.00001a_0/Z$ where $Z$ is atomic number of a chosen element; $a_0$ is the Bohr radius; and the $k^{\text{th}}$ datapoint is given by the formula, $r_{k+1} = \eta \cdot r_k$ where $\eta = 1.0123$.

3 The choice of the interval of 7 provides the accuracy of $10^{-7}$ for the normalisation condition of model atomic orbitals.

4 One of the variations of the cubic-spline algorithm, which is able to control the slopes of the left/right ends.
validated by examining the normalisation of the wavefunctions, i.e. $\langle s|s\rangle = \langle p_x|p_x\rangle = \langle p_y|p_y\rangle = \langle p_z|p_z\rangle = 1$, by keeping the errors in those numerical integrations less than $10^{-7}$. To give a bit more detail, the schematic diagram of the model splines is shown in figure 2.3 where the interpolating splines are expressed as:

$$
R_{(k-1)}(r) = a_{k-1}r^3 + b_{k-1}r^2 + c_{k-1}r + d_{k-1} \\
R_{(k)}(r) = a_{k}r^3 + b_{k}r^2 + c_{k}r + d_{k} \\
R_{(k+1)}(r) = a_{k+1}r^3 + b_{k+1}r^2 + c_{k+1}r + d_{k+1}
$$

For the equations in 2.2.3, the polynomial coefficients, $\{a, b, c, d\}$ with subscripts for each interval, could be generated by the employed spline method in practice.

Accurate reproduction of the fast oscillating parts of radial wavefunctions, with multiple nodes in the core region should not be necessary as the short-range interaction between lone pair cations and other ions mainly occur outside the core regions. Moreover, treating these oscillating parts is numerically more challenging as a finer grid is required. Therefore, alternatively, it should be possible to smooth out the oscillations by following the pseudo-orbital approach, which however would require a substantial additional effort. Indeed, even with the chosen prototype atomic orbitals, we have achieved sufficient accuracy as seen from their normalisation, and left development of pseudo-orbitals for future developments.

### 2.2.2 Calculating the induced electric dipole

When employing the new model, it is important to be able to calculate the dipole on the lone pair cation. Here, we will consider the simplest case of applying a constant electric field to an isolated lone pair cation and demonstrate how to calculate the time-independent transition dipole moment.
To account for the interaction between the applied electric field of $\mathbf{F}$ and the model electron lone pair, the Hamiltonian, defined earlier in equation 2.1.6, requires an addition term. That is:

$$\hat{H} = \hat{H}^0 + \hat{H}^1 \quad (2.2.4)$$

where $\hat{H}^0$ is the model onsite Hamiltonian (see equations 2.1.5 and 2.1.6) and

$$\hat{H}^1 = -\omega (\mathbf{F} \cdot \mathbf{r}). \quad (2.2.5)$$

In equation 2.2.5, $\omega$ is the charge of model lone pair electrons, and $\mathbf{r}$ is the displacement of the electrons from the centre of model cation. In Cartesian coordinates, the dot product within the parenthesis becomes $xF_x + yF_y + zF_z$. To calculate the energy of the lone pair, it is necessary to calculate the matrix elements, $H_{\mu\nu}$, which are:

$$H_{\mu\nu} = \langle \mu | \hat{H}^0 + \hat{H}^1 | \nu \rangle = \langle \mu | \hat{H}^0 | \nu \rangle + \langle \mu | \hat{H}^1 | \nu \rangle. \quad (2.2.6)$$

The first term $\langle \mu | \hat{H}^0 | \nu \rangle$ is already known (see equation 2.1.6). Using equation 2.2.6, the second term, $\langle \mu | \hat{H}^1 | \nu \rangle$, has three contributions:

$$\langle \mu | \hat{H}^1 | \nu \rangle = -\omega \{ \langle \mu | x | \nu \rangle F_x + \langle \mu | y | \nu \rangle F_y + \langle \mu | z | \nu \rangle F_z \}. \quad (2.2.7)$$

Thus, there are three volume integrals to compute for each of the sixteen matrix elements. Fortunately, many of these integrate to zero and, using symmetry arguments, the remaining six matrix elements are identical:

$$\langle s|x|p_x \rangle = \langle p_x|x|s \rangle = \langle s|y|p_y \rangle = \langle p_y|y|s \rangle = \langle s|z|p_z \rangle = \langle p_z|z|s \rangle \neq 0. \quad (2.2.8)$$
Employing spherical polar coordinates, \( \mathbf{r} = (r, \theta, \phi) \), each integral is separated into a product of three integrals, one in \( r, \theta \) and \( \phi \). It is the integrations with respect to \( \theta \) and \( \phi \) that determine whether the volume integral will vanish.

\[
\langle s | x | p_x \rangle = \int_0^\infty \int_0^\pi \int_0^{2\pi} R_s(r) x R_p(r) \Gamma_x(\theta, \phi) r^2 \sin \theta \, dr \, d\theta \, d\phi
\]

\[
= \frac{\sqrt{3}}{4\pi} \int_0^\infty R_s(r) R_p(r) r^3 \, dr \int_0^\pi \sin^3 \theta d\theta \int_0^{2\pi} \cos^2 \phi \, d\phi
\]

\[
= \frac{\sqrt{3}}{3} \int_0^\infty R_s(r) R_p(r) r^3 \, dr
\]

\[
= u.
\]

In the equations above, the basis functions of the \( s \) and the \( p_x \) are taken from equations 2.2.1 and 2.2.2, and the kernel \( x \) is expressed as \( r \sin \theta \cos \phi \). The final integral result is a scalar, \( u \), which depends on the choice of radial wavefunctions, \( R_s(r) \) and \( R_p(r) \). Consequently, the non-zero integrals in equation 2.2.8 implies that the constant electric field drives the (electric dipole) transition of charges, \( s \to p \), and, thus, creates a dipole moment, i.e., polarisation of lone pair is modelled.

Given our choice of radial functions (expressed as splines), the radial integral is actually a sum of integrals:

\[
u = \frac{\sqrt{3}}{3} \int_0^\infty R_s(r) R_p(r) r^3 \, dr = \frac{\sqrt{3}}{3} \sum_{k=1}^{k_{\text{max}}-1} \int_{r(k)}^{r(k+1)} R_{s; (k)}(r) R_{p; (k)}(r) r^3 \, dr
\]

where the original integral range of \([0, \infty]\) is replaced by the set of piecewise intervals, \([r(k), r(k+1)]\); and the radial wavefunctions, \( R_{s; (k)}(r) \) and \( R_{p; (k)}(r) \), are the cubic polynomials defined within the interval of \( r \in [r(k), r(k+1)] \) (see figure 2.3 and equations 2.2.3). Note that after this section, it is assumed that any radial integration will be completed in a similar way to that
shown in equation 2.2.10, i.e. for simplicity, we will not explicitly write radial integrations as a piecewise summation.

2.3 Interactions between ions and lone pair electrons

In equation 2.2.4, the term $\mathbf{H}^e$ represents a simple constant electric field. However, it is unlikely that a lone pair cation is only exposed to such uniform fields, as the electric fields generated by surrounding ions are likely to be much more complicated. Consider the field generated by an approaching molecule. Each atom in the molecule has a charge distribution that will generate a uniform and higher order field effects. Consequently, we need to improve our model to account for the effect of this more complex electric field on the polarisation of the lone pair.

2.3.1 Electrostatic interactions between point charges and the lone pair electron density

For a two-particle system consisting of one lone pair cation and one anion, consider the effect that the presence of the anion has on the distribution of the lone pair electron density. For large interatomic distances between these two particles, it is sufficient to treat the anion as a point charge. Likewise, for an $(N + 1)$-particle system consisting of one lone pair cation and $N$ ions that have an interatomic distance from the lone pair cation that are sufficiently large, we can treat all surrounding ions as point charges. Accordingly, $\mathbf{H}^1$, can be expressed as,

$$\mathbf{H}^1 = \sum_{A} \frac{\omega q_A}{r_A}. \hspace{1cm} (2.3.1)$$

In equation 2.3.1, the terms within summation over $A$, follows the Coulomb’s law where $\omega$ and $q_A$ represent charges of the model electron lone pair and the $A^{th}$ ion, respectively; and for the
separation distance, \( r_A \), i.e., \( r_A = |\mathbf{r} - \mathbf{R}_A| \), \( \mathbf{r} \) and \( \mathbf{R}_A \) are the position vectors of the lone pair electrons and the point charge representing the ion, respectively. By substituting equation 2.3.1 into the integrals \( \langle \mu | \hat{\mathbf{A}}^1 | \nu \rangle \) we get

\[
\langle \mu | \hat{\mathbf{A}}^1 | \nu \rangle = \sum_{A} \left( \mu | \frac{\omega q_A}{r_A} | \nu \right) = \left( \mu | \frac{\omega q_1}{r_1} | \nu \right) + \left( \mu | \frac{\omega q_2}{r_2} | \nu \right) + \left( \mu | \frac{\omega q_3}{r_3} | \nu \right) + \cdots \tag{2.3.2}
\]

where

\[
\left( \mu | \frac{\omega q_A}{r_A} | \nu \right) = \int_{0}^{\infty} \int_{0}^{\pi} \int_{0}^{2\pi} \mu(r, \theta, \phi) \frac{\omega q_A}{|\mathbf{r} - \mathbf{R}_A|} v(r, \theta, \phi) r^2 \sin \theta \, dr \, d\theta \, d\phi \tag{2.3.3}
\]

and \( \mu \) and \( v \) are the model basis functions.

**Figure 2.4** Schematic illustration of the distance \( |\mathbf{r} - \mathbf{R}_A| \) where the vector, \( \mathbf{r} \), refers to the locations of the lone pair electrons of the model cation at the origin.

Equation 2.3.3 can be solved exactly, in the specific choice of reference frame that is centred on the core of the lone pair cation and rotated such that the \( z \)-axis goes through the point charge, as shown in figure 2.4.
Let us first consider the volume integral for when the functions \{\mu, \nu\} are \{s, p_x\}, and by using the law of cosines, \(|\mathbf{r} - \mathbf{R}_A| = \sqrt{r^2 + R_A^2 - 2rR_A\cos\theta}\), equation 2.3.3 becomes

\[
\left\langle s \left| \frac{\omega q_A}{r_A} \right| p_z \right\rangle = \frac{\sqrt{3}}{2} \omega q_A \int_0^\infty R_s(r)R_p(r)r^2 \int_0^\pi \frac{\sin\theta \cos\theta}{\sqrt{r^2 + R_A^2 - 2rR_A\cos\theta}} \, dr \, d\theta. \tag{2.3.4}\]

Note that, as both \(s\) and \(p_x\) are not functions of \(\phi\), the integration over \(\phi\) in equation 2.3.3 gave a factor of \(2\pi\).

Before integrating over \(\theta\), it is convenient to use the substitution \(w = \cos\theta \Rightarrow dw = -\sin\theta \, d\theta\), so that,

\[
\int_0^\pi \frac{\sin\theta \cos\theta}{\sqrt{r^2 + R_A^2 - 2rR_A\cos\theta}} \, d\theta = \int_{-1}^{+1} \frac{w}{\sqrt{r^2 + R_A^2 - 2rR_Aw}} \, dw. \tag{2.3.5}\]

Note that the change in the direction of the limits changes the sign of the integration. Now using the substitution \(t = r^2 + R_A^2 - 2rR_Aw \Rightarrow dt = -2rR_A \, dw\), \(w = (r^2 + R_A^2 - t)/2rR_A\) gives,

\[
\int_{-1}^{+1} \frac{w}{\sqrt{r^2 + R_A^2 - 2rR_Aw}} \, dw = \frac{1}{4r^2R_A} \int_{|r-R_A|^2}^{(r+R_A)^2} (r^2 + R_A^2 - t) t^{-\frac{1}{2}} \, dt. \tag{2.3.6}\]

Feeding equation 2.3.6 into 2.3.4 and solving the integrals over \(t\) gives:\footnote{The integrals over \(t\) can be easily solved by using the Mathematica software, which supports symbolic calculations of mathematical expressions [75].}

\[
\left\langle s \left| \frac{\omega q_A}{r_A} \right| p_z \right\rangle = \frac{\sqrt{3}}{8R_A^2} \omega q_A \int_0^\infty R_s(r)R_p(r) \int_{|r-R_A|^2}^{(r+R_A)^2} \frac{(r^2 + R_A^2 - t) t^{-\frac{1}{2}} \, dt \, dr}{(r^2 + R_A^2 - t) t^{-\frac{1}{2}} \, dt \, dr}
\]

\[
= \frac{\sqrt{3}}{8R_A^2} \omega q_A \int_0^\infty R_s(r)R_p(r) \left( \frac{4}{3} \right) \left[ R_A^3 + r^3 - (R_A^2 + rR_A + r^2)(|R_A - r|) \right] \, dr. \tag{2.3.7}\]
The integral over \( r \) needs to be separated into two parts, for the ranges of \( r \in [0, R_A] \) and \( r \in [R_A, \infty] \), due to the term \(|R_A - r|\). Thus, the result in equation 2.3.7 becomes:

\[
\frac{\omega q_A}{T_A} |p_z\rangle = \frac{\sqrt{3} \omega q_A}{8 R_A^2} \left[ 4 \int_0^{R_A} R_s(r) R_p(r) \left( R_A^3 + r^3 - (R_A^2 + r R_A + r^2) (R_A - r) \right) dr + 4 \int_{R_A}^{\infty} R_s(r) R_p(r) \left( R_A^3 + r^3 + (R_A^2 + r R_A + r^2) (R_A - r) \right) dr \right], (2.3.8)
\]

which can be further abbreviated as:

\[
\frac{\omega q_A}{T_A} |p_z\rangle = \frac{\sqrt{3} \omega q_A}{8 R_A^2} \left[ \int_0^{R_A} R_s(r) R_p(r) T_+ \, dr + \int_{R_A}^{\infty} R_s(r) R_p(r) T_- \, dr \right], (2.3.9)
\]

where \( T_\pm = \frac{4}{3} \{ R_A^3 + r^3 - (R_A^2 + r R_A + r^2) \times \pm (R_A - r) \} \).

In equation 2.3.9, the terms, \( T_+ \) and \( T_- \), are selectively chosen for the conditions of inequalities, \( r < R_A \) and \( R_A < r \), respectively.

The demonstration shown from 2.3.3 to 2.3.9 is only for the case, \( \{\mu, \nu\} = \{s, p_z\} \). Accordingly, for the other possible combinations of \( \mu \) and \( \nu \), the integration should be appropriately modified.

The modifications could be done by following the analogous process what has been shown above.

As one further example of \( \{\mu, \nu\} = \{s, s\} \), one can simply modify \( \theta, \phi \) dependent terms for the \( s \) atomic orbitals and also the radial wavefunctions in equation 2.3.4. Consequently, the result of equation 2.3.9 will be in the different form with different expressions of \( T_\pm \) terms.

Assuming the calculations of the integrals, for all possible pairs of \( \{\mu, \nu\} \), are done, the Hamiltonian matrix of the model lone pair electrons, interacting with the point charge on the \( z \)-axis, will be in a form of...
\[
H = \begin{pmatrix}
H_{ss} & 0 & 0 & H_{sp_z} \\
0 & H_{px}p_x + \lambda & H_{py}p_y + \lambda & 0 \\
0 & 0 & 0 & H_{pz}p_z + \lambda \\
H_{pz}p_z & 0 & 0 & 0 \\
\end{pmatrix}
\]  
(2.3.10)

where the elements of matrix \( H \) are \( H_{\mu\nu} = \langle \mu | \hat{H}^0 | \nu \rangle + \left\langle \mu | \frac{\omega q_A}{r_A} \right| \nu \rangle \);

the zero elements are due to their integrals over angular terms, \( \theta \) and \( \phi \), which are given as a scalar product of even and odd functions and vanish; and \( \lambda \) is the parameter of the lone pair model, representing the energy separation between the \( s \) and the three \( p \) atomic orbitals, defined by the model onsite Hamiltonian, \( \hat{H}^0 \) (see equation 2.1.6).

According to the diagonalisation result of the matrix \( H \) in 2.3.10 (the matrix is real and symmetric, due to the use of real spherical harmonics, see equation 2.2.2), the ground state is represented by the \( sp_z \)-hybridisation\(^6\), and the other three states consist of the high energy anti-state of the \( sp_z \)-hybridisation and the two rests remain degenerate as the pure \( p_x \) and \( p_y \) atomic orbitals (details will be discussed in chapter 3). More precisely, such ground state of the lone pair is valid, if and only if the inequality holds, \( H_{mm} + \lambda > H_{ss} \) (\( m = p_x, p_y \) and \( p_z \)); otherwise the ground state will become the degenerate \( p \) states. Such inequality, however, is guaranteed by the model parameter \( \lambda \) with positive value\(^7\).

---

\(^6\) Since the point charge is placed on the \( z \)-axis, the ground state of \( sp_z \)-hybridised lone pair can be expected even without the diagonalisation. Furthermore, this hybridisation implies that the transition of lone pair charge, \( s \to p \), is driven by electric fields generated by a point charge.

\(^7\) In fact, the inequality also relies on the choice of the model radial wavefunctions. In detail, if the \( s \) atomic orbital is much more diffuse, i.e. if the radius is larger than the \( p \) atomic orbital, and with sufficiently small \( \lambda \) parameter, the inequality may not hold (the inequality would be reversed). However, for the employed model radial wavefunctions and tested \( \lambda \) parameters, the reversed case has not been observed.
**Generalisation of the model electrostatic interactions.** By using the integral method, demonstrated above, now it is able to calculate the electrostatic interaction between a point charge and the model electron lone pair. However, the method has a major limitation of which it can only deal with a point charge placed on the z-axis.

In general circumstance, it is more likely that ions are arbitrarily scattered in space as what is expressed in equation 2.3.2. To improve the method, the key idea is utilising the rotation of reference frame, which is a technique typically employed in quantum chemistry [76,77]. As aforementioned, about the centre of a lone pair cation, it is convenient to rotate the reference frame and such that the z-axis goes through a point charge. Afterwards, in the rotated reference frame (so-called the *local* reference frame) the calculations of the matrix elements – the integrals accounting for the electrostatic interaction between the model electron lone pair and the point charge – can be easily carried out.

In figure 2.5, the rotation is described where the original (solid lines) and the local reference frame (broken lines) are with Cartesian axes of \(\{x, y, z\}\) and \(\{x', y', z'\}\), respectively; and here, the core of lone pair cation is at the centre (black circle) and the point charge (grey circle) is on the \(z'\)-axis of the local reference frame.

**Figure 2.5** Schematic illustration of two Cartesian reference frames, \(\{x, y, z\}\) and \(\{x', y', z'\}\), which share the same origin. The axes for the former and latter are shown using solid and broken lines, respectively. The black and grey balls represent core of lone pair cation and a point charge, respectively.
To generalise the rotation method, it is necessary to find the transformation that bridges the original reference frame and a local reference frame. For our own design of the method, we first established the expressions for the model atomic orbitals, \( s, p_x, p_y, \) and \( p_z \), in the local reference frame, which are expressed as:

\[
\varphi_{\mu} = \sum_{h}^{4} P_{h\mu} \varphi_{h}'.
\] (2.3.11)

In equation 2.3.11, we introduced the new notations for the convenience of expression where the atomic orbitals in the global and the local reference frame are expressed as, \( \{ s, p_x, p_y, p_z \} = \{ \varphi_1, \varphi_2, \varphi_3, \varphi_4 \} \) and \( \{ s', p'_x, p'_y, p'_z \} = \{ \varphi'_1, \varphi'_2, \varphi'_3, \varphi'_4 \} \), respectively, where the atomic orbitals with the prime symbol, ‘, emphasises that those are in the local reference frame, and \( P_{h\mu} \) are the elements of the transformation matrix, which is

\[
P = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & P_{22} & P_{23} & P_{24} \\
0 & P_{32} & P_{33} & P_{34} \\
0 & P_{42} & P_{43} & P_{44}
\end{pmatrix}.
\] (2.3.12)

In the matrix \( P \), the elements in the first row and column are zero except the first element, \( P_{11} = 1 \). This is because, for any rotations applied to the original reference frame, the relation, \( \varphi_1 = \varphi'_1 \), does always hold since both \( \varphi_1 \) and \( \varphi'_1 \) refer to the \( s \) atomic orbital, which does not have directionality. As an example of \( \mu = 1 \) to make the expression, \( \varphi_1 = P_{11} \varphi'_1 + P_{12} \varphi'_2 + P_{13} \varphi'_3 + P_{14} \varphi'_4 \), same with, \( \varphi_1 = \varphi'_1 \), the matrix elements, \( P_{1,1-4} \), have to be, \( \{ P_{11}, P_{12}, P_{13}, P_{14} \} = \{ 1, 0, 0, 0 \} \). On the other hand, the \( p \) atomic orbitals are directly under the influence of rotations due to their directionalities in space. For instance, the representation of \( \varphi_2 \) (the \( p_x \) atomic orbital
in global reference frame) in a local reference frame can be expressed as $P_{22}\varphi_2' + P_{32}\varphi_3' + P_{42}\varphi_4'$

where $P_{12}\varphi_1'$ is omitted since $P_{12}$ is zero.

By using the relation in 2.3.11, the volume integral, $\langle \varphi_\mu | \hat{H} | \varphi_\nu \rangle$, in the global reference frame can be expressed as the form in the local reference frame.

$$
\langle \varphi_\mu | \hat{H} | \varphi_\nu \rangle = \sum_{h,l} P_{h\mu}P_{l\nu} \langle \varphi_h' | \hat{H}' | \varphi_l' \rangle 
$$

(2.3.13)

where the prime, $'$, is also on the Hamiltonian, $\hat{H}'$, to specify that term is expressed as a form in the local reference frame. For our case of interest, the lone pair electrons interacting with a point charge, the locally expressed Hamiltonian, $\hat{H}'$, is nothing different but merely treats the point charge on the local $z'$-axis.

Now, coming back to the problem of calculating the electrostatic interactions between the lone pair electrons and a collection of point charges (see equation 2.3.2), the electrostatic interaction by each charge can be calculated in its local reference frame, i.e. the relation given in 2.3.13 is used for each volume integral. Furthermore, this process is repeated for all possible pairs of $\{\mu, \nu\}$ to construct the Hamiltonian matrix, $H_{\mu\nu}$. Afterwards, using a standard matrix diagonalisation technique, it is able to determine the ground state of the lone pair electron density.

Note that throughout the rest of this thesis, as long as it is not required, any volume integrals will be completed in an analogous way to the demonstration in this section, for brevity, incorporating the rotation technique described above.

---

8 Here, instead of $\{\mu, \nu\}$ to represent the atomic orbitals as previously used, the alternative representation, $\{\varphi_\mu, \varphi_\nu\}$, is particularly employed here merely to distinguish the atomic orbitals in a different reference frame.
2.3.2 The short-range potential between ions and the model electron lone pair

So far, we have developed the lone pair ion model that accounts for the electrostatic interactions with a collection of point charges. The current model, however, is only valid when the lone pair cation is at large interatomic distances from the ions. Accordingly, to make a further advancement, it is necessary to introduce an additional interaction that is in effect when ions are in close proximity to the lone pair electron density.

At such close distance, the Pauli repulsion occurs between the electron lone pair and the charge cloud of the ion [78]. In the level of molecular mechanical theory, this type of interaction is typically modelled by using a so-called short-range potential that has been frequently represented by the first part of the Buckingham potential (see equation 1.4) [79–82], and due to the simplicity of the potential we decided to take it as the prototype in our model. Consequently, we can write the potential or the Hamiltonian of model electron lone pair interacting with an ion as:

$$H_{eYMM}^{\text{MM}} = D_Y \exp\left(-\frac{r_Y}{\rho_Y}\right) + \frac{\omega q_Y}{r_Y}. \quad (2.3.14)$$

In the above equation, the exponential term represents the short-range interaction with species dependent parameters, $D_Y$ and $\rho_Y$, which will be further discussed in chapter 3, especially about the model responses with respect to those parameters; and the second term is the electrostatic interaction as in equation 2.3.1. Here, note that we renamed the Hamiltonian with $H_{eYMM}$ for later convenience\(^9\), where the superscript MM is added to emphasise that the potential represents the

---

\(^9\) To distinguish the potential with the other potential that represents the interactions with lone pairs of different cations.
interaction with non-lone pair species, which are described in the level of molecular mechanical (MM) theory\textsuperscript{10} (more details will be delivered shortly in section 2.4).

With the potential in equation 2.3.14, the Hamiltonian of the model electron lone pair, interacting with a collection of ions, is expressed as:

\[ \hat{H} = \hat{H}^0 + \sum_A \hat{H}_{\text{MM}}^A \]  

(2.3.15)

where the term, \( \hat{H}^0 \), is again the onsite Hamiltonian. Moreover, in order to determine the ground state of the lone pair, calculation of the \( \mathbf{H} \) matrix must precede its diagonalisation, where the matrix elements are:

\[ H_{\mu\nu} = \langle \mu | \hat{H} | \nu \rangle = \langle \mu | \hat{H}^0 | \nu \rangle + \sum_A \langle \mu | \hat{H}_{\text{MM}}^A | \nu \rangle. \]  

(2.3.16)

### 2.4 Lone pair electrons interacting with lone pair densities

The electronic energy of a single lone pair cation, interacting with a number of non-lone pair ions, can be written as:

\[ E_{\text{LP}} = \langle \psi | \hat{H} | \psi \rangle \]  

(2.4.1)

where \( \psi \) is the molecular orbital of the model electron lone pair, see equation 2.1.1, and for the circumstance, the Hamiltonian, \( \hat{H} \), will be in the form in equation 2.3.15. By using the formula in equation 2.4.1, it is able to describe a system including a single lone pair cation that could be, for instance, employed to study a single lone pair cation as a defect in an ionic cluster. Here, our

\textsuperscript{10} As we assumed at the beginning, see section 1.2.1, the non-lone pair species are represented by conventional models, e.g. the rigid-ion model or the shell model.
objective, however, is to develop a more generalised version of the model that is able to manage a system, including more than one lone pair cation.

For this further generalisation, we can first revise the equation 2.4.1 into the following form (as in equation 1.11),

\[ E_{LP}^{total} = \sum_{\alpha} E_{\alpha;LP} = \sum_{\alpha} \langle \psi_{\alpha} | \hat{H}_{\alpha} | \psi_{\alpha} \rangle. \tag{2.4.2} \]

In equation 2.4.2, the letter \( \alpha \) runs over all lone-pair cations (or the electron lone pairs), and the letter as subscripts in \( E_{\alpha;LP} \), \( \psi_{\alpha} \) and \( \hat{H}_{\alpha} \), specify that those terms are the electronic energy, the molecular orbital and the Hamiltonian of the \( \alpha^{th} \) lone pair, respectively.

### 2.4.1 Localised model orbital approach for lone pair density

From equation 2.4.2, one might tell that we decide to employ the strategy of which the individual lone pairs are represented by their own Hamiltonian matrices, i.e. for \( N \) number of lone pair cations, there will be \( N \) number of \( 4 \times 4 \) matrices where a matrix represents a lone pair density. This subsystem approach would be advantageous to minimise the computational cost, especially required for the matrix diagonalisation process. In detail, a typical computational cost for a diagonalisation algorithm is known as \( O(n^3) \), and for \( N \) number of lone pair cations, the required computational cost for the diagonalisation processes will be \( N \times O(n^3) \) with the subsystem

---

11 In this context, the sentence implies that a matrix contains the ground state information of a lone pair where the lowest eigenvalue and its eigenvector correspond to the energy and the molecular orbital coefficients, respectively.

12 The big \( O \) notation is used, which indicates the time-complexity of an algorithm.
approach. On the other hand, for the same number of lone pair cations, if the total lone pair densities is expressed by using a single matrix with the larger size of \((4N \times 4N)\), the required cost will become \(O((N \times n)^3)\) where the cost will be much more expensive as the number of the lone pair cations, \(N\), increases.

Such a scheme of treating each atom as a subsystem is analogous to the absolutely localised molecular orbital (ALMO) – opposed to being delocalised over the entire system, e.g., canonical Hartree-Fock MOs – method often used in the electronic structure approaches [83–87]. The ALMO approach was originally designed to treat weakly interacting fragments of the entire system separately in order to improve the efficiency of the self-consistent field (SCF) process. By design, wavefunctions of individual molecules are variationally optimised, and charge transfer, or delocalisation, between molecules, or localised fragments, is prohibited. Similar ideas have also been pursued in semi-empirical models, e.g., the multistate empirical valence bond (MS-EVB) method by Voth et al where each fragment, a water molecule in their studies, is described by the linear combination of basis states, has its own Hamiltonian, and the lowest energy of entire system is dependent upon the states of subsystems [88,89]. The key methodological features in all of these methods: (i) the requirement of SCF method to optimise subsystems’ MOs; and (ii) prohibited charge transfer between molecules; are also common with our lone pair model, see discussion in sections 2.4.2 and 3.3.2.

2.4.2 Development of the Hamiltonian for a system including multiple lone pair cations

Assuming a system contains non-lone pair ions and also a number of lone pair cations, the Hamiltonian, \(\mathbf{H}_\alpha\) in equation 2.4.2 can be decomposed into the three terms as:
\[ \hat{\mathcal{H}}_\alpha = \hat{\mathcal{H}}^0_\alpha + \sum_A \hat{\mathcal{H}}^{MM}_{\alpha A} + \sum_{\beta \neq \alpha} \hat{\mathcal{H}}^{Lp}_{\alpha \beta}. \quad (2.4.3) \]

In equation 2.4.3, the first term on the right-hand side, \( \hat{\mathcal{H}}^0_\alpha \), refers to the onsite potential of the lone pair, which is modelled to represent the energy separation between the \( s \) and the three \( p \) atomic orbitals, see equation 2.1.6.

The second term, \( \hat{\mathcal{H}}^{MM}_{\alpha A} \), within the summation over \( A \), refers to the potential generated by non-lone pair species. Here, as aforementioned (see equation 2.3.15), we decided to describe the non-lone pair species as either a rigid-ion or a shell model ion, and for the system including multiple lone pair cations, this term has to contain the additional electrostatic interactions between the lone pairs and the core parts of the other lone pair cations\(^{13} \). Consequently, the term, \( \hat{\mathcal{H}}^{MM}_{\alpha A} \), should take several different forms depending on which type of species comes by the subscript \( A \), and therefore the possible expressions of \( \hat{\mathcal{H}}^{MM}_{\alpha A} \) are:

\[ \hat{\mathcal{H}}^{MM}_{\alpha A} = D_{\alpha A} \exp\left(-r^c_{\alpha A}/\rho_{\alpha A}\right) + \frac{\omega q^c_{\alpha A}}{r^c_{\alpha A}}, \quad (2.4.4a) \]

\[ \hat{\mathcal{H}}^{MM}_{\alpha A} = D_{\alpha A} \exp\left(-r^s_{\alpha A}/\rho_{\alpha A}\right) + \frac{\omega q^c_{\alpha A}}{r^c_{\alpha A}} + \frac{\omega q^s_{\alpha A}}{r^s_{\alpha A}}, \quad (2.4.4b) \]

and \( \hat{\mathcal{H}}^{MM}_{\alpha A} = \frac{\omega Z_{\alpha A}}{r^c_{\alpha A}}. \quad (2.4.4c) \)

\(^{13}\text{Recall that we have decided to separate the model lone pair cation into two parts: (i) the core part and (ii) the valence part (or the lone pair); and the core part is treated as a molecular mechanical entity, see section 1.3.5.}\)
Here, the equation, 2.4.4a or 2.4.4b, represents the cases when the species A refers to a non-lone pair ion, modelled by the rigid-ion model or the shell model; and equation 2.4.4c represents the case when the species A refers to the core parts of other lone pair cations\(^{14}\). Especially, for the interaction with the core part, the potential is kept purely electrostatic, which is because, the core part (including the nucleus and the core electrons) is surrounded by the lone pair electron density. Therefore, it is not typical that a core electron density overlaps with the lone pair electron densities of other cations. The third term in equation 2.4.3, \( \tilde{H}_{\alpha\beta}^{\text{LP}} \), within the summation over \( \beta \), represents the interaction between the \( \alpha \)th lone pair and \( \beta \)th lone pair where the condition, \( \beta \neq \alpha \), is to exclude the self-interaction.

**The interaction between two lone pair electron densities.** Continuing from the passage above, the third potential, \( \tilde{H}_{\alpha\beta}^{\text{LP}} \), represents the electric fields generated by the \( \beta \)th lone pair electron densities while the \( \alpha \)th lone pair electrons is under its’ influence. In this respect, by assuming that the lone pair electrons are under the influence of the average field, generated by the other lone pair densities, i.e. in the spirit of the mean-field approximation [90], the potential can be written as:

\[
\tilde{H}_{\alpha\beta}^{\text{LP}} = \frac{1}{2} \langle \psi_\beta | \Phi(\mathbf{r}_{\alpha\beta}) | \psi_\beta \rangle = \frac{1}{2} \sum_{\lambda,\sigma} c^\beta_{\lambda/\sigma} c^\beta_{\lambda/\sigma} \langle \lambda_\beta | \Phi(\mathbf{r}_{\alpha\beta}) | \sigma_\beta \rangle .
\]  

(2.4.5)

In equation 2.4.5, \( \psi_\beta \) refers to the molecular orbital of the \( \beta \)th lone pair density with its molecular orbital coefficients, \( c^\beta_{\lambda/\sigma} \), whose basis functions are \( \lambda_\beta / \sigma_\beta \), \( (\lambda, \sigma = s, p_x, p_y, p_z) \); the kernel of the integral, \( \Phi(\mathbf{r}_{\alpha\beta}) \), is the effective potential, which represents the interaction between the \( \alpha \)th lone

\(^{14}\) The indices, \( \alpha \) and \( \beta \), cannot be used for the same lone pair cation.
pair electrons and the $\beta^{th}$ lone pair electrons; and the leading product, $\frac{1}{2}$, is added to get rid of doubly counted interactions when the variables $\alpha$ and $\beta$ are interchanged.

Figure 2.6 Schematic diagram shows two lone pair cations whose cores are at $\mathbf{r}_1$ and $\mathbf{r}_2$ where $\mathbf{r}_a$ and $\mathbf{r}_\beta$ refer to the positions of their electron lone pairs, respectively; $r_{\alpha\beta}$ is the separation distance between lone pair electrons $r_{\alpha\beta}$ and $r_{\beta1}$ indicate several distances as in the diagram.

In equation 2.4.5, the effective potential, $\Phi(\mathbf{r}_{\alpha\beta})$, is a function of the distance, $r_{\alpha\beta}$, between the $\alpha^{th}$ and the $\beta^{th}$ lone pair electrons where the associated geometry is described in figure 2.6.

As the prototype function for the effective potential, we employed,

\[
\Phi(\mathbf{r}_{\alpha\beta}) \equiv \Phi(\mathbf{r}_\alpha) + \sum_{k=x,y,z} \left( \frac{\partial \Phi(\mathbf{r}_{\alpha\beta})}{\partial r_{\beta k}} \right)_{r_{\beta k}=0} \quad (2.4.6)
\]

where $\Phi(\mathbf{r}_{\alpha\beta}) = D_{\alpha\beta} \exp(-r_{\alpha\beta}/\rho_{\alpha\beta}) + \frac{\omega^2}{r_{\alpha\beta}}$

and $r_{\beta k}$ ($k = x, y, z$) refers to the Cartesian coordinates of the lone pair electrons of the $\beta^{th}$ lone pair cation. In equation 2.4.6, the potential, $\Phi(\mathbf{r}_{\alpha\beta})$, is given as the multipole expansion [91] of the function, $\Phi(\mathbf{r}_{\alpha\beta})$, whose form is same as the other prototype employed in equation 2.3.14 or 2.4.4a, including both the short-range repulsive potential that comes with the species dependent parameters, $\{D_{\alpha\beta}, \rho_{\alpha\beta}\}$ (the model response to these parameters will be further discussed in
chapter 3), and the long-range Coulomb interactions where \( \omega \) is the charge of model electron lone pair, assuming the same type of lone pair cations are interacting.

For the rationale of choosing prototype potential, \( \Phi(r_{\alpha\beta}) \), one can find that it contains up to the second-order term of the multipole-expansion, as the first-order dipolar term is essential to capture the effect of dipoles created by other lone pairs [92,93]. Moreover, the employment of the multipole-expansion method is computationally advantageous, since the direct integral calculation\(^1\) with such a kernel, \( \Phi(r_{\alpha\beta}) \), is known as the bottleneck in practical calculations [94,95].

Another rationale for the adopted approach is that the lone pair cations are unlikely to be nearest neighbouring each other. If the distance between two lone pair cations is as close as the first coordination shell by anions, the effective potential may not be accurate, however even in the close-packed environment that contains the cations, Pb\(^{2+}\) and Sn\(^{2+}\), their closest cation-cation distances are \(~3.8\) Å, which is much further than the first coordinating oxygen anions at \(~2.2\) Å [2,96].

### 2.4.3 The electronic energy of the lone pair model

With the established equations for describing the interactions, the substitution of equation 2.4.3 into equation 2.4.2 gives the electronic energy – by the model electron lone pairs – which is,

\[
E_{\text{LP}}^{\text{total}} = \sum_{\alpha} E_{\alpha;\text{LP}} = \sum_{\alpha} \langle \psi_{\alpha} | \hat{H}_{\alpha} | \psi_{\alpha} \rangle = \sum_{\alpha} \langle \psi_{\alpha} | \hat{H}_{\alpha}^0 + \sum_{\lambda} \hat{H}_{\alpha\lambda}^{\text{MM}} + \sum_{\beta \neq \alpha} \hat{H}_{\alpha\beta}^{\text{LP}} | \psi_{\alpha} \rangle
\]

\[(2.4.7)\]

where the total electronic energy by the lone pair (LP) densities, \( E_{\text{LP}}^{\text{total}} \), is given as the sum of energy contribution by each lone pair density, \( E_{\alpha;\text{LP}} = \langle \psi_{\alpha} | \hat{H}_{\alpha} | \psi_{\alpha} \rangle \).

\(^{15}\) The potential in equation 2.4.6 has to be integrated twice over the \( \alpha \)^{th} and the \( \beta \)^{th} lone pair molecular orbitals to get their interaction energy.
For the purpose of demonstration, the above expression can be further expanded as:

\[
E_{\text{LP}}^{\text{total}} = \sum_{\alpha} E_{\alpha;\text{LP}} = \sum_{\alpha} \langle \psi_{\alpha} | \hat{H}_\alpha | \psi_{\alpha} \rangle = \sum_{\alpha} \langle \psi_{\alpha} | \hat{H}_\alpha^0 + \sum_A \hat{H}_{\alpha A}^{\text{MM}} + \sum_{\beta \neq \alpha} \hat{H}_{\alpha \beta}^{\text{LP}} | \psi_{\alpha} \rangle
\]

\[
= \sum_{\alpha} \sum_{\mu, \nu} c_{\mu}^\alpha c_{\nu}^\alpha \left\{ \langle \mu_{\alpha} | \hat{H}_\alpha^0 | \nu_{\alpha} \rangle + \sum_A \langle \mu_{\alpha} | \hat{H}_{\alpha A}^{\text{MM}} | \nu_{\alpha} \rangle + \frac{1}{2} \sum_{\beta \neq \alpha} \sum_{\lambda, \sigma} c_{\lambda}^\beta c_{\sigma}^\beta \langle \mu_{\alpha \lambda} \lambda_{\beta} | \Phi(r_{\alpha \beta}) | \nu_{\alpha \sigma} \rangle \right\}
\]

(2.4.8)

where the bras and kets of the molecular orbitals, \( \psi_{\alpha} \), are expanded, and the term, \( \hat{H}_{\alpha \beta}^{\text{LP}} \), is replaced with its detailed expression, given in equation 2.4.5. Here, the molecular orbital coefficients and atomic orbitals of different electron lone pairs are distinguished by using, \( \{c_{\mu/\nu}^\alpha, \mu_{\alpha} / \nu_{\alpha} (\mu, \nu = s, p_x, p_y, p_z)\} \) and \( \{c_{\lambda/\sigma}^\beta, \lambda_{\beta} / \sigma_{\beta} (\lambda, \sigma = s, p_x, p_y, p_z)\} \), which are for the \( \alpha^{\text{th}} \) and the \( \beta^{\text{th}} \) lone pair densities; and the integrals over the kernel, \( \Phi(r_{\alpha \beta}) \), show that it has to be integrated twice over the basis functions of the \( \alpha^{\text{th}} \) \( (\mu_{\alpha} / \nu_{\alpha}) \) and the \( \beta^{\text{th}} \) \( (\lambda_{\beta} / \sigma_{\beta}) \).

For simplicity and the later use, we introduced the short-hand notations\(^{16}\), and equation 2.4.8 is further abbreviated as

\[
E_{\text{LP}}^{\text{total}} = \sum_{\alpha} E_{\alpha;\text{LP}} = \sum_{\alpha} \langle \psi_{\alpha} | \hat{H}_\alpha | \psi_{\alpha} \rangle
\]

\[
= \sum_{\alpha} \sum_{\mu, \nu} c_{\mu}^\alpha c_{\nu}^\alpha \left\{ H^{0}_{\alpha;\mu \nu} + \sum_A H^{\text{MM}}_{\alpha A;\mu \nu} + \frac{1}{2} \sum_{\beta \neq \alpha} \sum_{\lambda, \sigma} c_{\lambda}^\beta c_{\sigma}^\beta \langle \mu_{\alpha \lambda} \lambda_{\beta} | \Phi_{\alpha \beta;\lambda \sigma} | \nu_{\alpha \sigma} \rangle \right\}.
\]

(2.4.9)

\(^{16}\) \( H^{0}_{\alpha;\mu \nu} = \langle \mu_{\alpha} | \hat{H}_{\alpha}^0 | \nu_{\alpha} \rangle, H^{\text{MM}}_{\alpha A;\mu \nu} = \langle \mu_{\alpha} | \hat{H}_{\alpha A}^{\text{MM}} | \nu_{\alpha} \rangle \) and \( \Phi_{\alpha \beta;\lambda \sigma} = \langle \lambda_{\beta} | \Phi(r_{\alpha \beta}) | \sigma_{\beta} \rangle \).
2.4.4 The self-consistent field algorithm for calculating model electronic energy

For the total electronic energy, \( E_{LP}^{total} \), in equation 2.4.9, its lowest value is only guaranteed when the molecular orbitals, \( \psi_\alpha \), are in their respective ground states, or have the lowest energies. In the energy expression, the problem is that the molecular orbitals, or their coefficients, giving the lowest total electronic energy, can only be expressed implicitly \(^\text{17}\). A change occurring in one of the electron lone pairs will cause changes in the others, as they would respond and lower their energies in the new potential, which in turn will have an effect on the original lone pair where the change has occurred first. This effect can be seen clearly from the equation including the term accounting for interactions between lone pairs \( \alpha \) and \( \beta \),

\[
c_\alpha^\beta c_\beta^\alpha \langle \mu_\alpha | \mathcal{F}_{\alpha\beta} | \nu_\alpha \rangle,
\]

which shows that the energy of one of lone pairs, e.g., \( \alpha, E_{\alpha;LP} \), is determined implicitly by unknown orbital coefficients of all other lone pairs indexed here by \( \beta \), \( c_{\mu/\nu}^\beta \), and vice versa.

In the absence of analytical solutions, to find unknown ground states for all electron lone pairs, it is necessary to introduce an appropriate numerical technique. We choose to make use of a so-called self-consistent field (SCF) method, which has been widely used not only in electronic structure calculations but also employed by micro iterative hybrid quantum mechanical and molecular mechanical (QM and MM) approaches \([97–99]\). Such hybrids are also found in the lone pair model that includes the QM and MM parts for the electron lone pairs and the other rigid or polarisable shell ions, respectively.

As discussed in section 2.4.1, the lone pair model employs the concept of absolutely localised molecular orbital (ALMO) to represent lone pair electron densities on each cation; and therefore

\(^\text{17}\) There will be a set of the molecular orbital coefficients for lone pairs, which gives the lowest energy state. However, only utilising the energy expression in 2.4.9 is not enough to determine such set of coefficients.
instead of employing the canonical SCF method used in typical electronic structure methods, it is convenient to employ the model adapted SCF method. In equation 2.4.9, the model electronic energy, $E_{\text{LP}}^{\text{total}}$, is a sum of the energies of each subsystem, i.e., the energy of a lone pair, $E_{\alpha;\text{LP}}$. Its minimised value and corresponding molecular orbital coefficients, $c_\mu^\alpha$ and $c_\nu^\alpha$, could be obtained by solving the diagonalisation problem for subsystem $\alpha$, or its Hamiltonian matrix, $H_\alpha$, whose elements are (analogously with the ALMO method, see Ref [100]):

$$H_{\alpha;\mu\nu} = H_{\alpha;\mu\nu}^0 + \sum_\lambda H_{\alpha;\mu\nu}^{MM} + \frac{1}{2} \sum_{\beta \neq \alpha} \sum_{\lambda, \sigma} c_\lambda^\beta c_\sigma^\beta \langle \alpha_\mu | \mathcal{G}_{\alpha_\beta;\lambda_\sigma} | \alpha_\nu \rangle,$$ (2.4.10)

where the expression is obtained by dropping the summations over $\alpha$ and $\{\mu, \nu\}$ including the coefficients, $c_\mu^\alpha$ and $c_\nu^\alpha$, from equation 2.4.9.

In equation 2.4.10, the molecular orbital coefficients of other subsystems, $c_\lambda^\beta$ and $c_\sigma^\beta$, are only implicitly expressed, and therefore in the spirit of the original SCF method [101] the coefficients (or the molecular orbitals) can first be guessed. For instance, let $\psi_\alpha^{\text{init}}$ denote the initially guessed molecular orbitals, and with those molecular orbital coefficients it is able to compute the Hamiltonian matrices, $H_\alpha^{(0)}$, using equation 2.4.10. Subsequently, by diagonalising the matrices, we can determine the energies of the ground states, $E_{\alpha;\text{LP}}^{(0)}$, and their molecular orbitals, $\psi_\alpha^{(0)}$. Continuously, using the newly determined molecular orbitals, $\psi_\alpha^{(0)}$, it is possible to calculate, $H_\alpha^{(1)}$, and determine, $E_{\alpha;\text{LP}}^{(1)}$ and $\psi_\alpha^{(1)}$. By iterating the process above, we can generate a sequence of approximations, $\{E_{\alpha;\text{LP}}^{(2)}, \psi_\alpha^{(2)}\}, \{E_{\alpha;\text{LP}}^{(3)}, \psi_\alpha^{(3)}\}, \ldots$, and the iteration will be stopped when the differences between that current and the previous step result becomes negligibly small. Consequently, the entire process is summarised in the following Algorithm I.
The algorithm describes the SCF method adapted in the model, which shows how the electronic energy is calculated for \( N \) number of lone pair densities. The algorithm starts with a guessed set of lone pair molecular orbitals, \( \psi_{\alpha}^{\text{init}} \). In practice, the initially guessed molecular orbitals are set to the pure \( s \) atomic orbitals, i.e. \( \psi_{\alpha}^{\text{init}} = s \) for all \( \alpha \), which is a reasonable choice since the pure \( s \) atomic orbital represents the ground state of lone pair density when a lone pair cation is isolated in a vacuum.

According to Algorithm I, with the guessed set of molecular orbitals, it first calculates \( E_{\alpha;\text{LP}}^{(0)} \), \( E_{\text{LP}}^{\text{total}(0)} \) and \( \psi_{\alpha}^{(0)} \) by using equation 2.4.9. In the while loop, the algorithm calculates the Hamiltonian matrices, \( \mathbf{H}_{\alpha}^{(k)} \), using the molecular orbitals, \( \psi_{\alpha}^{(k)} \), with equation 2.4.10. Afterwards, the following matrix diagonalisation step determines the ground state energies, \( E_{\alpha;\text{LP}}^{(k+1)} \), and the molecular orbitals, \( \psi_{\alpha}^{(k+1)} \). As the last step of an iteration, the total electronic energy, \( E_{\text{LP}}^{\text{total}(k+1)} \), is simply updated by the summation of \( E_{\alpha;\text{LP}}^{(k+1)} \) for all \( \alpha \). The process is repeated as the loop (see the while statement in Algorithm I) runs over, and the algorithm generates a sequence of the approximations for \( \{E_{\alpha;\text{LP}}^{(k)} , \psi_{\alpha}^{(k)} \} \). Here, the loop breaks when the convergence conditions are achieved where the
tolerances are specified as $\tau_1$ (dimensionless) and $\tau_2$ (eV). In practice, both tolerances, $\tau_1$ and $\tau_2$, are set to $10^{-9}$.

2.4.5 Convergence analysis of the model SCF algorithm

Here, one might worry about the convergence of Algorithm I, since we could not prove that the algorithm must converge absolutely. Nonetheless, similar variations of the model SCF algorithm have been widely employed in the applications of \textit{ab initio} based electronic structure methods [86,100]; and other semi-empirical models [88,89]. Furthermore, here we provide several test examples of the use of model SCF algorithm\textsuperscript{18} to validate our approach and show its practicality.

The test samples are nanoclusters of $(\text{PbO})_N$ with various sizes of $N = 2, 3, 4, 8$ and 12. As the test result, in figure 2.7, the electronic energies of these samples were calculated and the energy

\textbf{Figure 2.7} Differences in the electronic energies, $E_{LP}^{\text{total}(k+1)} - E_{LP}^{\text{total}(k)}$, calculated by using the lone pair model, and the results are presented in logarithmic scale as a function of SCF iterations, $k$. The calculations were carried out for $(\text{PbO})_N$ nanoclusters with $N = 2, 3, 4, 8$ and 12, and a few of their structures are presented with the ball-and-stick model, Pb (black) and O (red) atoms. Note that the omitted bigger size nanoclusters are in similar structures where the unit of (PbO) are simply further concatenated as strands.

\textsuperscript{18} More details about the software used in this experiment can be found in Appendix E.
differences, $E_{LP}^{\text{total}}(k+1) - E_{LP}^{\text{total}}(k)$, for each size of $N$ are presented on a logarithmic scale as a function of SCF iteration number. Note that for the calculations carried out in this experiment, the oxygen ions were treated as point charges, $q = -2$; the parameters of the lone pair model were, \( \{\omega, \lambda\} = \{-1.64, 14.09\} \), which are the charge of the electron lone pair, $\omega$, and the onsite potential parameter, $\lambda$, see equation 2.1.6; and the effective potential parameters were, \( \{D_{LP}, \rho_{LP}\} = \{422, 0.098\} \), used to describe interactions between lone pairs, see equation 2.4.6.

As seen from the plot in figure 2.7, energy convergence profiles show linear decrease on a logarithmic scale for all sizes of $N$, and therefore SCF increments of the electronic energies decay exponentially as the SCF progresses. The trend of linear convergence is a consequence of choosing localised nonoverlapping atomic orbitals, representing the model lone pairs in close similarity to the ALMO approach, where similar behaviour has been found from the SCF convergence analysis of two interacting ALMOs by Lethe and Balint-Kurti, see Ref [102]. In detail, one of the convergence tolerances, $\tau_2$ (eV) in Algorithm I, was set to $10^{-9}$ in this experiment (marked as the breaking line in the plot), and the convergences were achieved at the 7th iteration for $N = 2$, and for the rest, $N \geq 3$, were achieved at the 9th ($N = 3$) and the 10th ($N = 4, 8$ and 12) iterations. In this respect, the model SCF algorithm is efficient, considering that the required number of iterations for $N = 3$ is same with those for $N = 4, 8$ and 12, even though there is maximum four times difference in the size of the nanocluster. For practical use of the algorithm, e.g. calculating electronic energies of more complicated atomic configurations than the samples employed in this experiment (where the atoms are simply concatenated on the plane), the required number of SCF iterations would vary depending on how ions are distributed in space. Moreover, in the worst-case scenario, the algorithm might suffer to meet the convergence criteria, caused by many different reasons. For
instance, by non-realistic atomic configurations\textsuperscript{19} or by the employment of ill-defined model parameters\textsuperscript{20}. Without those issues, we found that the required number of SCF iterations was typically ranging from 10 to 20 in practice for nanoclusters including maximum 20 lone pair cations, where the required number of SCF iterations were not largely deviated out from the observed results in this experiment.

The performance of the model SCF method could be further improved by introducing an acceleration support, for instance, employing the direct inversion in the iterative subspace (DIIS) method, widely used in the electronic structure methods [103–105]. However, so far, we found that the current version of implementation was effective, and this has been proved many times by using the model approach in the applications, characterising the nanocluster structures including lone pair cations where the details can be found in chapter 4.

2.5 The geometric derivatives of the lone pair model

So far, we have established the electronic energy of the lone pair electron densities, which is essentially a function of the coordinates of nuclei – assuming the rest of parameters used in the lone pair model are constants – and now our interest moves to calculating the derivatives with respect to the positions of nuclei that give the forces, see equation 1.1. and 1.2. Such geometric derivatives are important to determine the motions of nuclei, and therefore essential to explore the potential energy landscape of the model later in the applications.

\textsuperscript{19} A nanocluster includes too close interatomic distances or even some of ions are on top of each other.

\textsuperscript{20} In fact, this is not just only for the lone pair ion model but also for many other semi-empirical approaches relying on model parameters.
To quickly remind, the total energy of a system, including a number of lone pair cations and non-lone pair ions, is (as in equation 1.11):

\[ E_{\text{total}} = E_{\text{MM}}^{\text{total}} + E_{\text{LP}}^{\text{total}} \]  \hspace{1cm} (2.5.1a)

where

\[ E_{\text{MM}}^{\text{total}} = \sum_{i,j} \left[ \frac{q_i q_j}{r_{ij}} + D_{ij}\exp\left(-r_{ij}/\rho_{ij}\right) - \frac{C_{ij}}{r_{ij}^6} \right] \]  \hspace{1cm} (2.5.1b)

and

\[ E_{\text{LP}}^{\text{total}} = \sum_{\alpha} \langle \psi_{\alpha} | \hat{\Pi}_{\alpha} | \psi_{\alpha} \rangle. \]  \hspace{1cm} (2.5.1c)

In equation 2.5.1a, the total energy, \( E_{\text{total}} \), is decomposed into two parts, \( E_{\text{MM}}^{\text{total}} \) and \( E_{\text{LP}}^{\text{total}} \). The former term represents the energy contribution by the interactions of molecular mechanical (MM) species where they are ions modelled by using the rigid-ion model or the shell model, and the term also covers the core part of the lone pair cations, which is also modelled by using the rigid-ion model in the current approach. To describe the interactions of those species, we employed the Coulomb-Buckingham potential in equation 2.5.1b, and it is straightforward to derive the expression for its analytical derivatives [106,107].

The latter term, on the other hand, represents the electronic energy, \( E_{\text{LP}}^{\text{total}} \), of the model electron lone pairs whose analytical derivatives are more complex to derive. Without the analytical expressions of the derivatives, one simple alternative is using the finite-difference method (FDM) of which calculating the energy with respect to a small but finite displacement \( \Delta x \) in a parameter \( x \) (coordinate of nuclei), and the derivatives can be approximated by using the following equation:

\[ \frac{\partial E_{\text{LP}}^{\text{total}}}{\partial x} \approx \frac{E_{\text{LP}}^{\text{total}}(x + \Delta x) - E_{\text{LP}}^{\text{total}}(x - \Delta x)}{2\Delta x}. \]  \hspace{1cm} (2.5.2)
The advantage of FDM is that the required mathematics is much simpler than that of analytical method. However, it has a major flaw that the displacement \( \Delta x \) must be carefully adjusted for every case to avoid any possible numerical errors. Also, the required number of computations will be large if there are many atoms/ions in a system. Consequently, it is inevitable to find the analytical expressions for the derivatives if the mathematics is tractable in support of an efficient algorithm.

### 2.5.1 Geometric derivatives of the model electronic energy

To derive the analytical derivatives of the electronic energy, \( E_{\text{LP}}^{\text{total}} \), we start from its energy expression in equation 2.4.9, assuming that we have a converged set of the lone pair electron densities, i.e., molecular orbital coefficients of the ground states are known. In this condition, we can first proceed by treating \( E_{\text{LP}}^{\text{total}} \) as a function of a number of parameters \( x_1, x_2, x_3, \ldots \), which are the nuclear coordinates of lone pair ions and other non-lone pair ions in space. Here, the objective is the evaluation of \( \frac{\partial E_{\text{LP}}^{\text{total}}}{\partial x_i} \) at the origin \( x_1 = x_2 = \cdots = 0 \), which is, for the algebraic convenience, to cancel out the derivatives with respect to the atomic orbitals. The same technique was used by Pople et al, which can be found in Ref [108]. Accordingly, the direct differentiation of equation 2.4.9 with respect to parameter \( x \) gives,

\[
\frac{\partial E_{\text{LP}}^{\text{total}}}{\partial x} = \sum_\alpha \frac{\partial E_{\alpha;\text{LP}}}{\partial x} = \sum_\alpha \left \langle \psi_\alpha \left | \frac{\partial \tilde{H}_\alpha}{\partial x} \right | \psi_\alpha \right \rangle \\
= \sum_\alpha \sum_{\mu,\nu} c_\alpha c_\nu \left \langle \sum_A \frac{\partial H_{\alpha \mu;\lambda A;\nu}^{\text{MM}}}{\partial x} \right \rangle + \sum_\beta \sum_{\lambda,\sigma} \frac{\partial c_\lambda^\beta}{\partial x} \left \langle c_\sigma^\beta (\mu_\alpha | \Phi_{\alpha \beta;\lambda \sigma} | \nu_\alpha) + \frac{1}{2} c_\lambda^\beta c_\sigma^\beta \left \langle \mu_\alpha \left | \frac{\partial \Phi_{\alpha \beta;\lambda \sigma}}{\partial x} \right | \nu_\alpha \right \rangle \right \rangle .
\]

(2.5.3)
In equation 2.5.3, the derivatives of \( c_{\mu}^a \) and \( c_{\nu}^a \) vanish due to the Hellmann-Feynman theorem [109], and also the derivatives of the onsite Hamiltonian matrix, \( \partial H_{\alpha;\mu\nu}^0 / \partial x \), become zero and omitted, since its elements are treated as constant, see equation 2.1.6. The surviving derivatives are, \( \partial H_{\alpha;\mu\nu}^{MM} / \partial x \), \( \partial \Phi_{\alpha;\lambda e} / \partial x \) and \( \partial c_{\lambda}^B / \partial x \), where the first two terms account for the change in the interaction between a lone pair and a non-lone pair species and a lone pair function with respect to the changes in parameter \( x \), which are straightforward to calculate.

On the other hand, it is challenging to calculate the third term, derivatives of molecular orbital coefficients, \( \partial c_{\lambda}^B / \partial x \), which is the derivative of eigenvectors obtained by solving the diagonalisation problem of the Hamiltonian matrix of a lone pair. The presence of such terms can be understood as the model basis set and, therefore, the eigenvector depends on the coordinates of nuclei; hence, the changes in the positions of nuclei affects the molecular orbitals in order to minimise the energy of a system [110,111], i.e., a change occurring in the nuclear coordinate of a lone pair cation induces changes in electron lone pairs of other lone pair cations.

**The coupled-perturbed Hartree-Fock equation for the model derivative calculations.** As continuation of the passage above, we found the key to deal with the derivatives of molecular orbital coefficients analogous to the coupled perturbed Hartree-Fock (CPHF) equation, see Ref [112]. The form of equation obtained in our model is written as:

\[
\frac{\partial c_{\nu}^{a;n}}{\partial x} = \sum_{m \neq n} \sum_{\mu,\gamma} \frac{c_{\mu}^{a;m}}{E_{\alpha;LP}^m - E_{\alpha;LP}^n} \left( c_{\nu}^{a;n} \frac{\partial E_{\alpha;LP}^n}{\partial x} \delta_{\mu\nu} - c_{\nu}^{a;n} \frac{\partial H_{\alpha;\mu\nu}}{\partial x} \right)
\]  

(2.5.4)

where the superscripts \( m \) and \( n \) are introduced to indicate the \( m^{th} \) and \( n^{th} \) states of the \( \alpha^{th} \) lone pair electron density, e.g., \( E_{\alpha;LP}^m \) and \( c_{\nu}^{a;iX} (X = m, n \text{ and } Y = \mu, \nu, \gamma) \) denote the energies (or
eigenvalues) and the molecular orbital coefficients (or the eigenvector of each eigenvalue); \( \delta_{\mu \nu} \) is the Kronecker delta; and the terms, \( \partial E_{\alpha;LP}^n / \partial x \) and \( \partial H_{\alpha;\mu \nu} / \partial x \), can be further expressed by dropping out the summations over \( \alpha \) and \( \{\alpha, \mu, \nu\} \) from equation 2.5.3, which are:

\[
\frac{\partial E_{\alpha;LP}}{\partial x} = \sum_{\mu, \nu} c_{\mu}^\alpha c_{\nu}^\alpha \left[ \sum_{A} \frac{\partial H_{\alpha;\lambda \mu \nu}^{MM}}{\partial x} + \sum_{\beta \neq \lambda, \sigma} \left( \frac{\partial c_{\lambda}^\beta}{\partial x} c_{\sigma}^\beta \left( \mu_{\alpha} \Phi_{\alpha \beta;\lambda \sigma} | v_{\alpha} \right) + \frac{1}{2} \frac{c_{\lambda}^\beta c_{\sigma}^\beta}{c_{\mu}^\alpha} \left( \mu_{\alpha} \frac{\partial \Phi_{\alpha \beta;\lambda \sigma}}{\partial x} | v_{\alpha} \right) \right) \right],
\]

(2.5.5)

\[
\frac{\partial H_{\alpha;\mu \nu}}{\partial x} = \sum_{A} \frac{\partial H_{\alpha;\lambda \mu \nu}^{MM}}{\partial x} + \sum_{\beta \neq \lambda, \sigma} \left( \frac{\partial c_{\lambda}^\beta}{\partial x} c_{\sigma}^\beta \left( \mu_{\alpha} \Phi_{\alpha \beta;\lambda \sigma} | v_{\alpha} \right) + \frac{1}{2} \frac{c_{\lambda}^\beta c_{\sigma}^\beta}{c_{\mu}^\alpha} \left( \mu_{\alpha} \frac{\partial \Phi_{\alpha \beta;\lambda \sigma}}{\partial x} | v_{\alpha} \right) \right) .
\]

(2.5.6)

Note that for the variable, \( n \), since we are mainly interested in the derivatives of the molecular orbital coefficients that represent the ground state, and thus \( n \) is picked for such a state.

### 2.5.2 Development of a fixed-point iteration algorithm for the model derivative calculations

To make use of the CPHF equation in 2.5.4, we first found that the equation is coupled with equation 2.5.3 that shows the derivatives of the model electronic energy. In detail, the derivative of molecular orbital coefficients, \( \partial c_{\gamma}^{\alpha;n} / \partial x \), can only be implicitly expressed since the terms have a dependency on \( \partial E_{\alpha;LP} / \partial x \) and \( \partial H_{\alpha;\mu \nu} / \partial x \), see equation 2.5.5 and 2.5.6, respectively.

For this coupled relation, we found that the derivatives of molecular orbital coefficients can be solved by a fixed-point iterative method. To give a short description of the method, it starts with an initially guessed set of \( \left( \partial c_{\gamma}^{\alpha;n} / \partial x \right)^{\text{init}} \) to calculate \( \left( \partial H_{\alpha;\mu \nu} / \partial x \right)^{(0)} \) and \( \left( \partial E_{\alpha;LP}^{n} / \partial x \right)^{(0)} \). Afterwards those calculated terms are fed back into the CPHF equation to calculate \( \left( \partial c_{\gamma}^{\alpha;n} / \partial x \right)^{(0)} \) for use in the next iteration. The operation can be repeated and generate a sequence of approximations.
for \((\partial c_{\alpha:y}^n/\partial x)^{(1)}, (\partial c_{\alpha:y}^n/\partial x)^{(2)}, \ldots\), until a certain termination criterion – when the difference between the current and the previous of the derivatives of molecular orbital coefficients becomes negligible – is achieved. Once the criterion is achieved, the derivative, \(\partial E_{\alpha:LP}^n/\partial x\), is summed over for all \(\alpha\) to determine the derivatives of the model electronic energy, \(\partial E_{LP}^{total}/\partial x\). Consequently, the entire process is summarised in Algorithm II.

**Algorithm II.** The fixed-point iteration method for calculating the model derivatives of the electronic energy.

\[
\text{init} \quad (\partial c_{\gamma}^{\alpha:n}/\partial x)^{\text{init}} = 0 \text{ for all } (\alpha, \gamma, n); \text{ set } k = 0; \\
\text{calculate } (\partial H_{\alpha;\mu\nu}/\partial x)^{(k)}, (\partial E_{\alpha;LP}^n/\partial x)^{(k)} \text{ with } (\partial c_{\gamma}^{\alpha:n}/\partial x)^{\text{init}} \text{ (eq 2.5.5 and 2.5.6)} \\
\text{while } 1/N \sum_{\alpha}^N \sqrt{[(\partial c_{\gamma}^{\alpha:n}/\partial x)^{(k+1)} - (\partial c_{\gamma}^{\alpha:n}/\partial x)^{(k)}]^2} < \tau \\
\text{for } \alpha = 1, 2, \ldots N \\
\text{calculate } (\partial c_{\alpha:y}^n/\partial x)^{(k)} \text{ with } (\partial H_{\alpha;\mu\nu}/\partial x)^{(k)}, (\partial E_{\alpha;LP}^n/\partial x)^{(k)} \text{ (eq 2.5.4)} \\
\text{for } \alpha = 1, 2, \ldots N \\
\text{calculate } (\partial H_{\alpha;\mu\nu}/\partial x)^{(k+1)}, (\partial E_{\alpha;LP}^n/\partial x)^{(k+1)} \text{ with } (\partial c_{\alpha:y}^n/\partial x)^{(k)} \text{ (eq 2.5.5 and 2.5.6)} \\
k = k + 1 \\
\text{end} \\
\text{calculate } \partial E_{LP}^{total}/\partial x = \sum_{\alpha}^N \partial E_{\alpha;LP}/\partial x
\]

In the algorithm, it starts with setting the initial derivatives, \(\partial c_{\gamma}^{\alpha:n}/\partial x = 0\), for all \((\alpha, \gamma, n)\). Considering the possible range of the molecular orbital coefficients are, \(c_{\gamma}^{\alpha:n} \in [-1, +1]\), thus such initial guess is reasonable since the possible converging directions of the derivatives, \(\partial c_{\gamma}^{\alpha:n}/\partial x\), can be toward either the positive or negative. In the while loop, it generates the sequence of approximations for \((\partial c_{\alpha:y}^n/\partial x)^{(k)}\), and as the convergence criterion, the average least squares of the derivatives is used with the tolerance \(\tau\). In practice, the tolerance was set to \(10^{-9}\), and the algorithm proved to be efficient and accurate. For the unique convergence of the algorithm, further details will be discussed shortly in section 2.5.3.
Accuracy of the model derivatives and the fixed-point iteration algorithm. In order to validate the accuracy of the model derivatives, calculated by Algorithm II, we carried out a simple numerical experiment. In the experiment, we calculated the geometric derivatives of atoms in a test sample of (PbO)$_4$ nanocluster where the atomic structure is shown in figure 2.8.21.

Furthermore, taking the calculated derivatives as the analytical derivatives, we compared the result with two different controls: (i) same derivatives but calculated by using the FDM method with the displacement of $\Delta x = 0.0025 \text{ Å}$, see equation 2.5.2, and (ii) a constraint is imposed of which the derivatives were calculated without the contributions of the derivatives of the molecular orbital coefficients.22 Consequently, the result is summarised in table 2.1. In the table, the second column shows the derivatives of each atom toward the Cartesian components, $x$, $y$ and $z$, calculated by the model algorithm (or the model analytical derivatives), and the values in the third and the fourth columns, show the relative errors of those controls with respect to the analytical result. Taking into account that the errors of (i) FDM are less then 0.05%, whose results are close to the analytical derivatives, such small figures validate the accuracy of the model derivative calculation algorithm.

---

21 The same parameters for the lone pair model was employed as in the model SCF experiment.

22 The derivatives of molecular orbital coefficients were set to zeroes, i.e., neglected, to investigate their effect.
For the errors by the (ii) Constraint in the fourth column, the errors are systematic as -156% and -22% for O and Pb atoms. Such consistent figures are due to the cuboid configuration of the nanocluster (see figure 2.8) where O and Pb atoms are alternately located at the vertices, i.e., each species, O and Pb, is exposed in the same environment, and the same amount of systematic corrections are required for the species, O and Pb. Here, recall that the derivative calculations were carried out without the contributions by the derivatives of the molecular orbital coefficients, those errors imply that the displacements in the positions of nuclei affect the molecular orbitals of lone pairs, which is in line with the interpretation made earlier in this section. Consequently, this systematic error makes the result of the model analytical derivatives more reliable.

### 2.5.3 Convergence analysis of the algorithm calculating the model derivatives

Despite we have established the method to calculate the derivatives of the model electronic energy (see Algorithm II), it is still questionable since we have employed an iterative method and therefore whether the algorithm guarantees the unique convergence of the derivatives, i.e., whether the derivatives of the molecular orbital coefficients are unique. This is an important question that if there were more than one possible convergence results by the algorithm, the fact would imply that the

---

**Table 2.1** The analytical geometric derivatives of the atoms in the (PbO)₄ nanocluster – see figure 2.8 – calculated by Algorithm II, and the relative errors by the two other controls, (i) and (ii).

<table>
<thead>
<tr>
<th>Atom</th>
<th>Analytical derivatives (eV/Å)</th>
<th>(i) FDM err (%)</th>
<th>(ii) Constraint err (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>x</td>
<td>y</td>
<td>z</td>
</tr>
<tr>
<td>O1</td>
<td>-8.71.E-01</td>
<td>-2.48.E-01</td>
<td>-8.56.E-01</td>
</tr>
<tr>
<td>O2</td>
<td>1.01.E+00</td>
<td>4.71.E-01</td>
<td>-5.59.E-01</td>
</tr>
<tr>
<td>O3</td>
<td>-4.46.E-01</td>
<td>8.27.E-01</td>
<td>8.18.E-01</td>
</tr>
<tr>
<td>O4</td>
<td>3.06.E-01</td>
<td>-1.05.E+00</td>
<td>5.95.E-01</td>
</tr>
<tr>
<td>Pb1</td>
<td>-1.35.E+00</td>
<td>2.51.E+00</td>
<td>2.48.E+00</td>
</tr>
<tr>
<td>Pb2</td>
<td>-2.64.E+00</td>
<td>-7.49.E-01</td>
<td>-2.60.E+00</td>
</tr>
<tr>
<td>Pb3</td>
<td>3.06.E+00</td>
<td>1.43.E+00</td>
<td>-1.69.E+00</td>
</tr>
<tr>
<td>Pb4</td>
<td>9.31.E-01</td>
<td>-3.19.E+00</td>
<td>1.81.E+00</td>
</tr>
</tbody>
</table>
electronic forces are inappropriate to be used in applications due to potential issues, e.g., the atomic forces may vibrate depending on the atomic configurations of the lone pair cations, and consequently the algorithm unexpectedly slows down during the energy optimisation processes, or even in the worst case scenario, the algorithm may fail to determine a converged set of the derivatives of molecular orbital coefficients.

**The uniqueness of model derivatives.** To make a long story short, the model algorithm guarantees the unique convergence of the derivatives for the molecular orbital coefficients, and needless to say it also guarantees the unique derivatives of the model electronic energy.

To start the proof, it is necessary to rearrange the coupled-equations, 2.5.3 and 2.5.4, into a form of linear system where the unknown variables are, of course, the derivatives of molecular orbital coefficients, \( \partial c_\alpha^\nu / \partial x \). Obtain the expression for \( \partial E_{\alpha;\text{LP}} / \partial x \) first, by dropping out the summation over \( \alpha \) from equation 2.5.3, and its rearrangement gives,

\[
\frac{\partial E_{\alpha;\text{LP}}}{\partial x} = \sum_{\mu, \nu} \sum_{\beta} \sum_{\lambda, \sigma} c_\mu^\alpha c_\nu^\alpha \left( \frac{\partial c_\lambda^\beta}{\partial x} \right) c_\sigma^\beta \langle \mu_\alpha | \Phi_{\alpha;\beta;\lambda, \sigma} | \nu_\alpha \rangle \\
+ \sum_{\mu, \nu} c_\mu^\alpha c_\nu^\alpha \left\{ \sum_A \frac{\partial H_{\alpha;A;\mu, \nu}^{MM}}{\partial x} + \sum_{\beta} \sum_{\lambda, \sigma} \frac{1}{2} c_\lambda^\beta c_\sigma^\beta \left[ \mu_\alpha \left| \frac{\partial \Phi_{\alpha;\beta;\lambda, \sigma}}{\partial x} \right| \nu_\alpha \right] \right\}.
\]

In the expression above, the term denoted by \( L_\alpha \) is a constant that varies by the choice of \( \alpha \). Accordingly, one can abbreviate it further as:

\[
\frac{\partial E_{\alpha;\text{LP}}}{\partial x} = \sum_{\mu, \nu} \sum_{\beta} \sum_{\lambda, \sigma} c_\mu^\alpha c_\nu^\alpha c_\lambda^\beta \langle \mu_\alpha | \Phi_{\alpha;\beta;\lambda, \sigma} | \nu_\alpha \rangle \times \left( \frac{\partial c_\lambda^\beta}{\partial x} \right) + L_\alpha \quad (2.5.7)
\]
where there are variables, \( \lambda = \{c_s, c_{p_x}, c_{p_y}, c_{p_z}\} \), and for \( N \) number of lone pairs cations, \( \alpha = {1, 2, ..., N} \), opening the summations in equation 2.5.7 yields \( N \) number of linear equations:

\[
\begin{align*}
& a_{11} \partial_x c_s^1 + a_{12} \partial_x c_{p_x}^1 + a_{13} \partial_x c_{p_y}^1 + a_{14} \partial_x c_{p_z}^1 + a_{15} \partial_x c_s^2 + \cdots + a_{1N} \partial_x c_{p_z}^N + \mathcal{L}_1 = \partial_x E_1 \\
& a_{21} \partial_x c_s^1 + a_{22} \partial_x c_{p_x}^1 + a_{23} \partial_x c_{p_y}^1 + a_{24} \partial_x c_{p_z}^1 + a_{25} \partial_x c_s^2 + \cdots + a_{2N} \partial_x c_{p_z}^N + \mathcal{L}_2 = \partial_x E_2 \\
& \vdots \\
& a_{N1} \partial_x c_s^1 + a_{N2} \partial_x c_{p_x}^1 + a_{N3} \partial_x c_{p_y}^1 + a_{N4} \partial_x c_{p_z}^1 + a_{N5} \partial_x c_s^2 + \cdots + a_{N,N} \partial_x c_{p_z}^N + \mathcal{L}_N = \partial_x E_N
\end{align*}
\]  

(2.5.8)

Here, the short hand notations, \( \partial E_{\alpha;LP}/\partial x \equiv \partial x E_{\alpha} \) and \( \partial c_{\lambda}^\beta /\partial x \equiv \partial x c_{\lambda}^\beta \), are introduced for simplicity; and \( a_{ij} \) are the constant coefficients with \( i = 1, 2, ..., N \) and \( j = 1, 2, ..., 4N \). Furthermore, to obey the condition, \( \beta \neq \alpha \) (over the summations, see equation 2.5.7), some of the coefficients must vanish, e.g. \( a_{11} = a_{12} = a_{13} = a_{14} = 0 \), which is the case of \( \alpha = \beta = 1 \), and assumed to be completed.

At this stage, the linear equations in 2.5.8 cannot be solved since the problem is to determine \( 4N \) number of unknowns, \( \partial_x c_{\lambda}^\beta \), with \( N \) number of equations. To set the number of equations equal to the unknowns, we can exploit the CPHF equation in 2.5.4, which can be first rearranged as:

\[
\frac{\partial c^\alpha;_{n}}{\partial x} = \sum_{m \neq n} \sum_{\mu, \nu} \left( \frac{c^\alpha;_{m}}{E^m_{\alpha;LP}} \frac{c^\alpha;_{n}}{E^n_{\alpha;LP}} \delta_{\mu, \nu} \right) \times \left( \frac{\partial E^{n}_{\alpha;LP}}{\partial x} \right) + \sum_{m \neq n} \sum_{\mu, \nu} \frac{c^\alpha;_{m}}{E^m_{\alpha;LP}} \frac{c^\alpha;_{n}}{E^n_{\alpha;LP}} \left( \frac{\partial H^{\alpha;\mu,\nu}}{\partial x} \right). 
\]

Again, there are two constant terms, \( C_{\gamma}^\alpha;_{n} \) and \( R_{\gamma}^\alpha;_{n} \); and the variable, \( n \), can be dropped out by assuming \( n \) refers to the ground states. Thus, the above equation can be simplified as:

\[
\partial_x c^\alpha;_{\gamma} = C^\alpha;_{\gamma} \times \partial_x E_{\alpha} + R^\alpha;_{\gamma} \quad \Rightarrow \quad \partial_x E_{\alpha} = \frac{1}{C^\alpha;_{\gamma}} \left( \partial_x c^\alpha;_{\gamma} - R^\alpha;_{\gamma} \right) 
\]

(2.5.9)
where the same short-hand notations are introduced as in equation 2.5.8. In equation 2.5.9, there is a variable \( \gamma = \{ s, p_x, p_y, p_z \} \), and therefore, the term, \( \partial_x E_\alpha \), can have four different forms:

\[
\partial_x E_\alpha = \frac{1}{\epsilon_s^\alpha} (\partial_x c_s^\alpha - R_s^\alpha) = \frac{1}{\epsilon_{p_x}^\alpha} (\partial_x c_{p_x}^\alpha - R_{p_x}^\alpha) = \frac{1}{\epsilon_{p_y}^\alpha} (\partial_x c_{p_y}^\alpha - R_{p_y}^\alpha) = \frac{1}{\epsilon_{p_z}^\alpha} (\partial_x c_{p_z}^\alpha - R_{p_z}^\alpha). 
\]  

(2.5.10)

Here, for instance, by using the relation above, the case of \( \alpha = 1 \) in equations 2.5.8, which is:

\[
a_{11} \partial_x c_s^1 + a_{12} \partial_x c_{p_x}^1 + a_{13} \partial_x c_{p_y}^1 + a_{14} \partial_x c_{p_z}^1 + a_{15} \partial_x c_s^2 + \cdots + a_{1,4N} \partial_x c_s^N + L_1 = \partial_x E_1,
\]

can be expressed as four different forms (see the top four equations in 2.5.11), and as well as for the rest of \( \alpha = 2, 3, \ldots, N \). Accordingly, the equations in 2.5.8 now can be expanded as a linear system that has \( 4N \) number of linear equations with \( 4N \) number of unknowns. That is:

\[
a_{11} \partial_x c_s^1 + a_{12} \partial_x c_{p_x}^1 + a_{13} \partial_x c_{p_y}^1 + a_{14} \partial_x c_{p_z}^1 + a_{15} \partial_x c_s^2 + \cdots + a_{1,4N} \partial_x c_s^N + L_1 = 1/c_s^1 (\partial_x c_s^1 - R_s^1)
\]

\[
a_{11} \partial_x c_s^1 + a_{12} \partial_x c_{p_x}^1 + a_{13} \partial_x c_{p_y}^1 + a_{14} \partial_x c_{p_z}^1 + a_{15} \partial_x c_s^2 + \cdots + a_{1,4N} \partial_x c_s^N + L_1 = 1/c_{p_x}^1 (\partial_x c_{p_x}^1 - R_{p_x}^1)
\]

\[
a_{11} \partial_x c_s^1 + a_{12} \partial_x c_{p_x}^1 + a_{13} \partial_x c_{p_y}^1 + a_{14} \partial_x c_{p_z}^1 + a_{15} \partial_x c_s^2 + \cdots + a_{1,4N} \partial_x c_s^N + L_1 = 1/c_{p_y}^1 (\partial_x c_{p_y}^1 - R_{p_y}^1)
\]

\[
a_{11} \partial_x c_s^1 + a_{12} \partial_x c_{p_x}^1 + a_{13} \partial_x c_{p_y}^1 + a_{14} \partial_x c_{p_z}^1 + a_{15} \partial_x c_s^2 + \cdots + a_{1,4N} \partial_x c_s^N + L_1 = 1/c_{p_z}^1 (\partial_x c_{p_z}^1 - R_{p_z}^1)
\]

\[\vdots\]

\[
a_{N1} \partial_x c_s^1 + a_{N2} \partial_x c_{p_x}^1 + a_{N3} \partial_x c_{p_y}^1 + a_{N4} \partial_x c_{p_z}^1 + a_{N5} \partial_x c_s^2 + \cdots + a_{N,4N} \partial_x c_s^N + L_N = 1/c_s^N (\partial_x c_s^N - R_s^N)
\]

\[
a_{N1} \partial_x c_s^1 + a_{N2} \partial_x c_{p_x}^1 + a_{N3} \partial_x c_{p_y}^1 + a_{N4} \partial_x c_{p_z}^1 + a_{N5} \partial_x c_s^2 + \cdots + a_{N,4N} \partial_x c_s^N + L_N = 1/c_{p_x}^N (\partial_x c_{p_x}^N - R_{p_x}^N)
\]

\[
a_{N1} \partial_x c_s^1 + a_{N2} \partial_x c_{p_x}^1 + a_{N3} \partial_x c_{p_y}^1 + a_{N4} \partial_x c_{p_z}^1 + a_{N5} \partial_x c_s^2 + \cdots + a_{N,4N} \partial_x c_s^N + L_N = 1/c_{p_y}^N (\partial_x c_{p_y}^N - R_{p_y}^N)
\]

\[
a_{N1} \partial_x c_s^1 + a_{N2} \partial_x c_{p_x}^1 + a_{N3} \partial_x c_{p_y}^1 + a_{N4} \partial_x c_{p_z}^1 + a_{N5} \partial_x c_s^2 + \cdots + a_{N,4N} \partial_x c_s^N + L_N = 1/c_{p_z}^N (\partial_x c_{p_z}^N - R_{p_z}^N)
\]

(2.5.11)

At this stage, the unknowns, i.e., the derivatives of molecular orbital coefficients, \( \partial_x c^\beta_\lambda \), can be directly calculated by solving the linear system above. However, there are a few steps left to reduce
the linear system into the form that the model algorithm exploits. To show this, for each linear equation in 2.5.11, one can first multiply $C_s^\alpha$ and add $R_s^\beta$ to the both hand sides with appropriately chosen $\alpha$ and $\gamma$. For instance, multiplying $C_s^1$ and add $R_s^1$ to the first equation in 2.5.11 gives:

$$b_{11} \partial_x c_s^1 + b_{12} \partial_x c_{p_x}^1 + b_{13} \partial_x c_{p_y}^1 + b_{14} \partial_x c_{p_z}^1 + b_{15} \partial_x c_s^2 + \cdots + b_{1,4N} \partial_x c_{p_x}^N + L_1 + R_s^1 = \partial_x c_s^1$$

(2.5.12)

where $b_{11} = C_s^1 a_{11}, b_{12} = C_s^1 a_{12}, \ldots, b_{1,4N} = C_s^1 a_{1,4N}$. This similar operation can be carried out for each equation in 2.5.11, and the result can be summarised as the following matrix form:

$$B c + d = c$$

(2.5.13)

where

$$B = \begin{pmatrix}
    b_{11} & b_{12} & \cdots & b_{1,4N} \\
    b_{21} & b_{22} & \cdots & b_{2,4N} \\
    \vdots & \vdots & \ddots & \vdots \\
    b_{4N,1} & b_{4N,2} & \cdots & b_{4N,4N}
\end{pmatrix} \in \mathbb{R}^{4N \times 4N},$$

$$c^T = \left( \partial_x c_s^1, \partial_x c_{p_x}^1, \partial_x c_{p_y}^1, \partial_x c_{p_z}^1, \ldots, \partial_x c_{p_x}^N \right) \in \mathbb{R}^{4N} \text{ and }$$

$$d^T = \left( L_1 + R_s^1, L_1 + R_{p_x}^1, L_1 + R_{p_y}^1, L_1 + R_{p_z}^1, \ldots, L_N + R_{p_x}^N \right) \in \mathbb{R}^{4N}.$$ 

The matrix equation in 2.5.13 is in a standard form to apply an iterative method. For using the Jacobi iteration method, equation 2.5.13 can be converted as:

$$B c^{(k)} + d = c^{(k+1)}.$$  

(2.5.14)

For the given form above, the iterative method is applied and generates a sequence of approximations, $c^{(1)}, c^{(2)}, \ldots$ with initially guessed $c^{(0)}$ until the approximation meets the termination criteria, i.e., when the difference between $c^{(k+1)}$ and $c^{(k)}$ becomes negligible – essentially identical with the criteria in Algorithm II (see the while statement).
In this section, we have shown that the two coupled-equations, 2.5.3 and 2.5.4, used to calculate the derivatives of the model electronic energy, can be reduced into a system of linear equations in 2.5.11. Strictly speaking, the unique solutions for the derivatives exist if and only if the linear system is consistent and independent. However, after many uses of the model derivative calculations, to our best knowledge we have not yet encountered any cases that break the conditions, and we conclude that the model approach, in Algorithm II, is again reliable in general and validated.

2.5.4 Rotation of reference frames in the model derivatives calculations

In section 2.3.1, we have discussed the rotations of reference frames, which was introduced to solve the volume integrals in the model calculations. Likewise, to calculate the model derivatives, the similar type of rotations is more frequently required to solve a greater number of different volume integrals, viz. simply a gradient of an integral yields three different integrals for its Cartesian components. Therefore, it would be instructive to go through the technique, employed in the lone pair model.

To demonstrate, let’s take an example that we want to calculate the following integrals:

$$
\nabla H_{\mu \nu} = \langle \mu | \nabla \hat{H} | \nu \rangle = \left\{ \left( \mu \left| \frac{\partial \hat{H}}{\partial x} \right| \nu \right), \left( \mu \left| \frac{\partial \hat{H}}{\partial y} \right| \nu \right), \left( \mu \left| \frac{\partial \hat{H}}{\partial z} \right| \nu \right) \right\} \tag{2.5.15}
$$

where $\mu$ and $\nu$ are the model basis functions, representing the atomic orbitals; and the integrals in the curly parentheses are the gradient of the Hamiltonian matrix element where the variables, {x, y, z}, refer to the nuclear coordinate of a lone pair cation. Here, if the Hamiltonian, $\hat{H}$, represents the potential by a point charge arbitrarily placed in space, first it is necessary to rotate the original reference frame to the local reference frame where the z'-axis goes through the point
charge – as used earlier in section 2.3, the prime indicates that any items with the symbol is in the local reference frame throughout this section.

The key point is that the integrals above, i.e., the gradient of $H_{\mu \nu}$, can be calculated in the local reference frame in the first place, and their results can be inverted back to the original reference frame. As in section 2.3, it is convenient to distinguish the Hamiltonian in the original reference and the local reference with $\hat{H}$ and $\hat{H}'$; also using the notations for the atomic orbitals, $\varphi_{\mu/\nu}$ and $\varphi'_{\mu/\nu}$, where the subscripts vary from 1 to 4, i.e. $\{\varphi_1, \varphi_2, \varphi_3, \varphi_4\}$ and $\{\varphi'_1, \varphi'_2, \varphi'_3, \varphi'_4\}$ refer to the model lone pair atomic orbitals, $\{s, p_x, p_y, p_z\}$. By using the notation, one can express the gradient in equation 2.5.15 as in the local reference – note that the Einstein’s summation convention is employed for the repeated indices particularly in this section – which is:

$$\{\left(\varphi_{\mu} \left| \frac{\partial \hat{H}}{\partial x} \right| \varphi_{\nu}\right), \left(\varphi_{\mu} \left| \frac{\partial \hat{H}}{\partial y} \right| \varphi_{\nu}\right), \left(\varphi_{\mu} \left| \frac{\partial \hat{H}}{\partial z} \right| \varphi_{\nu}\right)\} \equiv \{P_{h\mu}P_{l\nu}\left(\varphi'_h \left| \frac{\partial \hat{H}'}{\partial x} \right| \varphi'_l\right), P_{h\mu}P_{l\nu}\left(\varphi'_h \left| \frac{\partial \hat{H}'}{\partial y} \right| \varphi'_l\right), P_{h\mu}P_{l\nu}\left(\varphi'_h \left| \frac{\partial \hat{H}'}{\partial z} \right| \varphi'_l\right)\}$$

(2.5.16)

where $P_{h\mu}$ and $P_{l\nu}$ are the elements of the transformation matrix $P$ given in 2.3.12; and instead of using the equal sign but the symbol, $\equiv$, is used to specify the two gradients (or vectors) are equivalent but expressed in different coordinate systems.

As demonstrated in the section 2.3, the calculations of the gradient in the local reference frame are straightforward, and here the main interest is in how to bring the locally calculated results back to the original reference. For simplicity, using the following shorthand notations,

$$\{\partial_x H'_\mu, \partial_y H'_\mu, \partial_z H'_\mu\} \equiv \{P_{h\mu}P_{l\nu}\left(\varphi'_h \left| \frac{\partial \hat{H}'}{\partial x} \right| \varphi'_l\right), P_{h\mu}P_{l\nu}\left(\varphi'_h \left| \frac{\partial \hat{H}'}{\partial y} \right| \varphi'_l\right), P_{h\mu}P_{l\nu}\left(\varphi'_h \left| \frac{\partial \hat{H}'}{\partial z} \right| \varphi'_l\right)\},$$

(2.5.17a)
\[ \{ \partial_x H_{\mu \nu}, \partial_y H_{\mu \nu}, \partial_z H_{\mu \nu} \} \equiv \{ \left( \varphi_\mu \left| \frac{\partial \hat{H}}{\partial x} \right| \varphi_\nu \right), \left( \varphi_\mu \left| \frac{\partial \hat{H}}{\partial y} \right| \varphi_\nu \right), \left( \varphi_\mu \left| \frac{\partial \hat{H}}{\partial z} \right| \varphi_\nu \right) \}, \]  

(2.5.17b)

The local expression in 2.5.17a can be inverted back to the original expression in 2.5.17b by the formula,

\[
\begin{pmatrix}
\partial_x H_{\mu \nu} \\
\partial_y H_{\mu \nu} \\
\partial_z H_{\mu \nu}
\end{pmatrix}
= (P')^T
\begin{pmatrix}
\partial_x H'_{\mu \nu} \\
\partial_y H'_{\mu \nu} \\
\partial_z H'_{\mu \nu}
\end{pmatrix}
\]  

(2.5.18)

where the matrix, \( P' \in \mathbb{R}^{3\times3} \), is the lower block diagonal matrix of the transformation matrix \( P \), obtained by cutting off the first column and the row – containing dummy elements to deal with the rotation of \( s \) atomic orbital – from \( P \).

As the extension of the method, using the same shorthand notation \( \partial_{mn}^2 H_{\mu \nu} \) where \( m \) and \( n \) are \{\( x, y, z \)\}, the second derivatives of \( H_{\mu \nu} \) can be obtained by the formula,

\[
\begin{pmatrix}
\partial_{xx}^2 H_{\mu \nu} & \partial_{xy}^2 H_{\mu \nu} & \partial_{xz}^2 H_{\mu \nu} \\
\partial_{yx}^2 H_{\mu \nu} & \partial_{yy}^2 H_{\mu \nu} & \partial_{yz}^2 H_{\mu \nu} \\
\partial_{zx}^2 H_{\mu \nu} & \partial_{zy}^2 H_{\mu \nu} & \partial_{zz}^2 H_{\mu \nu}
\end{pmatrix}
= (P')^T
\begin{pmatrix}
\partial_{xx}^2 H'_{\mu \nu} & \partial_{xy}^2 H'_{\mu \nu} & \partial_{xz}^2 H'_{\mu \nu} \\
\partial_{yx}^2 H'_{\mu \nu} & \partial_{yy}^2 H'_{\mu \nu} & \partial_{yz}^2 H'_{\mu \nu} \\
\partial_{zx}^2 H'_{\mu \nu} & \partial_{zy}^2 H'_{\mu \nu} & \partial_{zz}^2 H'_{\mu \nu}
\end{pmatrix}(P').
\]  

(2.5.19)

The transformation in equation 2.5.19 is particularly required to calculate the geometric derivatives for a lone pair cation with respect to other lone pair cations due to the effective potential – describing the interactions between the model electron lone pairs, see equation 2.4.6 – which is already once differentiated.
2.6 Chapter summary

As summary of this chapter, we would like to highlight the following key points in order to clarify some features of the lone pair model. Especially, what have been newly developed and which ones already exist and were integrated to the lone pair model.

(i) We have employed the single molecular orbital approach to represent the effective valence state of lone pair species, which is formulated by the two-level system consisting of the energetically low-lying $s$ state and the higher degenerate three $p$ states. The energy difference is parameterised by introducing the onsite potential.

(ii) To represent the lone pair molecular orbital, we introduced the basis set of one $s$ and three $p$ atomic orbitals taken from the FHI-AIMS code in the DFT level of theory. For use in the lone pair model, we numerically reproduced the radial wavefunctions as splines.

(iii) We treated the model lone pair cation as two parts: (i) the valence part (electron lone pair), represented by the introduced basis set; and (ii) the core part, including the nuclei and the core electrons. Each part carries their own charges, $\omega$ and $Z$, respectively, and their sum is equal to the formal charge of a lone pair cation. Here, the core part implicitly contains the residual or inner $d^{10}$ electrons. We assume that the inner electrons are relatively inert to the $sp$-hybridisation occurring in the valence level of a lone pair cation, and therefore, we decided to model the core part by using the rigid-ion model.

(iv) In order to integrate the model lone pair cations into a system including non-lone pair ions, described under the classical approaches (the rigid-ion model or the shell model is used), we developed a Hamiltonian of the model lone pair electrons. The Hamiltonian contains the potentials of electron lone pair, interacting with external species, including non-lone
pair species and other lone pair electron cations. Here, for the interactions amongst lone pairs, it requires the SCF method to determine their ground states, and consequently, we present the model adapted SCF algorithm.

(v) In order to calculate the derivatives of the model electronic energy (due to the electron lone pairs), we have developed an approach that requires the support of CPHF method. The approach is summarised as an iterative method, and we proved that the method is reliable by the convergence analysis of the newly developed algorithm.
In this chapter, we are going to discuss the following features of the lone pair model (LPM): (i) response of an isolated lone pair cation to electric field, including the model induced electric multipole moments and polarisabilities. Here, a comparison is made with properties obtained using the polarisable shell model and DFT. (ii) In the LPM, the effective electronic space of a lone pair cation consists of one s and three p states, and for its description the one-electron approach is employed. In order to validate the model, we have investigated the highest occupied molecular orbital (HOMO) of a lone pair cation in the LPM and DFT. (iii) Lastly, the sensitivity of the LPM parameters of the employed effective potentials – interaction of lone pair electrons with non-lone-pair ions and other lone pair electron densities – was tested against DFT. Results of this investigation prove to be instructive, especially for choosing the LPM parameters within physically sensible value ranges before their use in the LPM applications.
3.1 Response of a lone pair cation to electric field

In general, polarisation of an ion by an external electric field is characterised by the induced electric multipole moments. In the presence of an electric field, which consists of uniform field $F_i$ and uniform field gradient $F_{ij}$, the energy of the ion can be expanded as a Taylor series, using traceless quadrupoles defined by Buckingham [114,115] (note that the Cartesian summation convention is used for the repeated indices in this section),

$$E(F_i, F_{ij}) = E^0 - \mu_i^0 F_i - \frac{1}{2} \alpha_{ij} F_i F_j - \frac{1}{6} \beta_{ijk} F_i F_j F_k - \cdots$$

$$- \frac{1}{3} Q_{ij}^0 F_i - \frac{1}{3} A_{ijk} F_i F_j F_k - \frac{1}{6} C_{ijkl} F_i F_j F_k F_l - \cdots \quad (3.1.1)$$

where $i$ and $j$ refer to the Cartesian components $\{x, y, z\}$; $\alpha_{ij}$ and $\beta_{ijk}$ are dipole-polarisability and the first hyperpolarisability; $A_{ijk}$ and $C_{ijkl}$ are dipole-quadrupole and quadrupole-quadrupole polarisabilities; $E^0$, $\mu_i^0$ and $Q_{ij}^0$ are the energy, permanent dipole and quadrupole moments, respectively, which describe a free ion.

A lone pair cation in its ground state has a spherical charge density (the valence state has a pure $s$ character), and therefore has zero permanent moments $\mu_i^0$ and $Q_{ij}^0$. Thus, the energy in equation 3.1.1 is determined by the field induced electric moments:

$$\mu_i = - \frac{\partial E}{\partial F_i} = \mu_i^0 + \alpha_{ij} F_j + \frac{1}{2} \beta_{ijk} F_j F_k + \frac{1}{3} A_{ijk} F_k + \cdots \quad (3.1.2a)$$

$$Q_{ij} = -3 \frac{\partial E}{\partial F_{ij}} + \delta_{ij} \sum_k \frac{\partial E}{\partial F_{kk}} = Q_{ij}^0 + A_{klj} F_k + C_{ijkl} F_l + \cdots \quad (3.1.2b)$$

and polarisabilities:

$$\alpha_{ij} = \frac{\partial \mu_i}{\partial F_j}, \quad \beta_{ijk} = \frac{\partial \alpha_{ij}}{\partial F_k}, \quad (3.1.2c)$$
\[ A_{kij} = \frac{\partial Q_{ij}}{\partial F_k} \quad \text{and} \quad C_{ijkl} = \frac{\partial^2 Q_{ij}}{\partial F_k \partial F_l}. \quad (3.1.2d) \]

**3.1.1 Electric dipole moment and dipole-polarisability of the lone pair ion model**

In section 2.2.2, the response of a lone pair to a constant electric field was already briefly discussed in the context of integration of numerical radial wavefunctions, employed in the LPM. Recall, in the constant electric field \( \mathbf{F} \), a lone pair Hamiltonian is,

\[ \hat{H} = \hat{H}^0 - \omega (\mathbf{F} \cdot \mathbf{r}) = \hat{H}^0 - \omega (xF_1 + yF_2 + zF_3) \quad (3.1.3) \]

where \( \omega \) is the total charge of a lone pair, \( \mathbf{F} = \{F_1, F_2, F_3\} \), and \( \mathbf{r} = \{x, y, z\} \) are the Cartesian coordinates of the electron lone pair. The first step to determine the ground state of the Hamiltonian is to evaluate the Hamiltonian matrix elements in the basis of the LPM atomic orbitals (indexed by \( \mu \) and \( \nu \)),

\[ H_{\mu\nu} = \langle \mu | \hat{H} | \nu \rangle = \langle \mu | \hat{H}^0 + \hat{H}^1 | \nu \rangle = \langle \mu | \hat{H}^0 | \nu \rangle + \langle \mu | \hat{H}^1 | \nu \rangle \quad (3.1.4a) \]

where

\[ \langle \mu | \hat{H}^1 | \nu \rangle = -\omega \{ \langle \mu | x | \nu \rangle F_1 + \langle \mu | y | \nu \rangle F_2 + \langle \mu | z | \nu \rangle F_3 \}. \quad (3.1.4b) \]

As demonstrated previously – see equations 2.2.8-10 – the nonvanishing integrals on the right-hand side of equation 3.1.4b are:

\[ \langle s | x | p_x \rangle = \langle p_x | x | s \rangle = \langle s | y | p_y \rangle = \langle p_y | y | s \rangle = \langle s | z | p_z \rangle = \langle p_z | z | s \rangle = u \quad (3.1.5) \]

where \( u \) is determined by the radial wavefunctions of the LPM basis set.

Using the onsite interaction contribution, \( \langle \mu | \hat{H}^0 | \nu \rangle \), see equation 2.1.6, the final Hamiltonian matrix of the electron lone pair in the constant electric field is:
\[ H = H^0 + H^1 = \begin{pmatrix} 0 & -\omega F_1 u & -\omega F_2 u & -\omega F_3 u \\ -\omega F_1 u & \lambda & 0 & 0 \\ -\omega F_2 u & 0 & \lambda & 0 \\ -\omega F_3 u & 0 & 0 & \lambda \end{pmatrix} \]. \quad (3.1.6)\]

Diagonalising matrix \( H \) yields the lowest \( (E_-) \) and the highest \( (E_+) \) eigenvalues,

\[ E_{\pm} = \frac{1}{2} \left( \lambda \pm \sqrt{\lambda^2 + 4u^2(F_1^2 + F_2^2 + F_3^2)\omega^2} \right). \quad (3.1.7) \]

The remaining two states of a pure \( p \) character do not hybridise with the \( s \) state and remain degenerate with the eigenvalue of \( \lambda \), which accounts for the energy difference between the \( s \) and the three \( p \) states of an isolated lone pair cation.

For the ground state energy \( E_- \), the LPM dipole moment and dipole-polarisability are:

\[ \mu_i = -\frac{\partial E_-}{\partial F_i} = \frac{2u^2\omega^2 F_i}{(\lambda^2 + 4u^2(F_1^2 + F_2^2 + F_3^2)\omega^2)^{1/2}}, \quad (3.1.8a) \]

\[ \alpha_{ij} = \frac{\partial \mu_i}{\partial F_j} = \frac{2u^2\omega^2 \left( -4u^2\omega^2 F_i F_j + \delta_{ij}(\lambda^2 + 4u^2(F_1^2 + F_2^2 + F_3^2)\omega^2) \right)}{\left(\lambda^2 + 4u^2(F_1^2 + F_2^2 + F_3^2)\omega^2\right)^{3/2}} \quad (3.1.8b) \]

The results in the equations have a few notable differences with the shell model in vacuum, according to which (equations 1.7-1.9):

\[ E_{SM} = -\frac{1}{2} \frac{q^2}{k} F^2, \quad (3.1.9a) \]

\[ \mu_{SM;i} = \frac{q^2}{k} F_i, \quad (3.1.9b) \]

\[ \alpha_{SM;ij} = \frac{q^2}{k} \delta_{ij}. \quad (3.1.9c) \]
First, the dipole polarisability in the LPM is a function of applied field, whereas in the shell model, it is independent of the field. Second, differentiating $\alpha_{ij}$ with respect to $F_k$ will give non-zero values$^{23}$, i.e., the model cation shows finite first hyperpolarisability; by further differentiation, finite second, third and higher hyperpolarisabilities can be derived in the same manner.

To elucidate the field dependent properties of the LPM, for simplicity, we choose the field along the $z$-axis and express the energy of the ground state, the induced dipole moment, dipole-polarisability, the first and the second hyperpolarisabilities as functions of the field magnitude $F$, which are:

$$E(F) = \frac{1}{2}\left(\lambda + \frac{\sqrt{\lambda^2 + 4u^2F^2\omega^2}}{}\right).$$  \hspace{1cm} (3.1.10a)

$$\mu(F) = -\frac{\partial E}{\partial F} = \frac{2u^2\omega^2F}{(\lambda^2 + 4u^2F^2\omega^2)^{1/2}},$$  \hspace{1cm} (3.1.10b)

$$\alpha(F) = \frac{\partial \mu}{\partial F} = \frac{2u^2\lambda^2\omega^2}{(\lambda^2 + 4u^2F^2\omega^2)^{3/2}},$$  \hspace{1cm} (3.1.10c)

$$\beta(F) = \frac{\partial \alpha}{\partial F} = -\frac{24u^4\lambda^2\omega^4F}{(\lambda^2 + 4u^2F^2\omega^2)^{5/2}},$$  \hspace{1cm} (3.1.10d)

$$\gamma(F) = \frac{\partial \beta}{\partial F} = -\frac{24\lambda^2(u^4\lambda^2\omega^4 - 16u^6F^2\omega^4)}{(\lambda^2 + 4u^2F^2\omega^2)^{7/2}}.$$  \hspace{1cm} (3.1.10e)

For the detailed analysis, the ground state energy in equation 3.1.10a is expanded as a Taylor series around $F = 0$:

$^{23}$ Whereas that of the shell model will become zero.
\[ E = E(0) - \mu(0)F - \frac{1}{2}\alpha(0)F^2 - \frac{1}{6}\beta(0)F^3 - \frac{1}{24}\gamma(0)F^4 - \cdots \]

\[
= \frac{1}{2} \left( \lambda - \sqrt{\lambda^2} \right) - 0 \cdot F - \frac{12 u^2 \lambda^2 \omega^2}{2 (\lambda^2)^{3/2}} F^2 - \frac{1}{6} 0 \cdot F^3 + \frac{1}{24} \frac{24 \lambda^2 (u^4 \lambda^2 \omega^4)}{(\lambda^2)^{7/2}} F^4 \\
+ \frac{1}{120} 0 \cdot F^5 + O(F^6)
\]

\[
= 0 - 0 \cdot F - \frac{u^2 \omega^2}{\lambda} F^2 - 0 \cdot F^3 + \frac{u^4 \omega^4}{\lambda^3} F^4 + \frac{1}{120} 0 \cdot F^5 + O(F^6).
\]

In equation 3.1.11, the first term is the ground state energy of the electron lone pair, which is zero in the absence of the field. In this series only the odd terms (even powers of \( F \)) survive, with coefficients proportional to dipole-polarisability, \( \alpha \), and the second hyperpolarisability, \( \gamma \) (see the 3\(^{rd} \) and 5\(^{th} \) terms). In contrast, the even terms (odd powers of \( F \)), whose coefficients are determined by the dipole moment, \( \mu \), and the first hyperpolarisability, \( \beta \) (see 2\(^{nd} \) and 4\(^{th} \) terms) vanish in the absence of external field.

**Perturbation expression.** The energy expansion 3.1.11 can also be obtained using the time-independent perturbation theory [116]. Here, the ground state energy of a system in a weak perturbing potential is written as a convergent series:

\[
E = E^{(0)} + E^{(1)} + E^{(2)} + \cdots + E^{(n)} + \cdots \tag{3.1.12}
\]

where \( E^{(n)} \) is the \( n^{th} \) order correction to the energy.

In equation 3.1.12, the first term \( E^{(0)} \) is the ground state energy of the unperturbed system that we set to zero, i.e., \( E^{(0)} = E_s^{(0)} = 0 \). Calculation of the next order terms is straightforward and only requires knowledge of the eigenvalues and eigenfunctions of the unperturbed Hamiltonian, \( \hat{H}^0 \), as
well as the matrix elements of the perturbing potential, $\mathbf{H}^1$, which we have already evaluated. We choose to direct external field along the $z$-axis, which simplifies equation 3.1.4b to $\langle \mu | H^1 | \nu \rangle = \langle \mu | -\omega F z | \nu \rangle$ with only one nonvanishing pair of off-diagonal symmetrical matrix elements $\langle s|z|p_z \rangle = \langle p_z|z|s \rangle = u$. Thus, commonly used the first and the second order energy corrections, $E^{(1)}$ and $E^{(2)}$, are obtained as,

$$E^{(1)} = \langle s | \mathbf{H}^1 | s \rangle = \langle s | -\omega F z | s \rangle = 0 \quad \text{and}$$

$$E^{(2)} = \sum_{m \neq s} \frac{\langle s | \mathbf{H}^1 | m \rangle^2}{E_s^{(0)} - E_m^{(0)}} = \frac{\langle s | -\omega F z | p_z \rangle^2}{E_s^{(0)} - E_p_z^{(0)}} = -\frac{u^2 \omega^2}{\lambda} F^2.$$  \hspace{1cm} (3.1.13)

Continuously, the higher order corrections, $E^{(3)}$--$(5)$, can be systematically obtained, with the following notation for simplicity,

$$H_{nm} \equiv \langle n | \mathbf{H}^1 | m \rangle \quad \text{and} \quad E_{nm} \equiv E_n^{(0)} - E_m^{(0)}.$$  \hspace{1cm} (3.1.14)

the third order corrections is,

$$E^{(3)} = \sum_{n_2 \neq s} \sum_{n_3 \neq s} \frac{H_{s n_2} H_{n_2 n_3} H_{n_3 s}}{E_{s n_2} E_{n_2 n_3} E_{n_3 s}} - H_{ss} \sum_{n_3 \neq s} \frac{|H_{s n_3}|^2}{E_{n_3 s}^2} = 0,$$  \hspace{1cm} (3.1.15)

which vanishes due to the zero terms, $H_{ss} = 0$ and $H_{n_3 n_2} = 0$; and the fourth order is,

$$E^{(4)} = \sum_{n_2 \neq s} \sum_{n_3 \neq s} \sum_{n_4 \neq s} \frac{H_{s n_4} H_{n_4 n_3} H_{n_3 n_2} H_{n_2 s}}{E_{s n_2} E_{n_2 n_3} E_{n_3 n_4} E_{n_4 s}} - E^{(2)} \sum_{n_4 \neq s} \frac{|H_{s n_4}|^2}{E_{n_4 s}^2}$$

$$- 2H_{ss} \sum_{n_3 \neq s} \sum_{n_4 \neq s} \frac{H_{s n_4} H_{n_4 n_3} H_{n_3 s}}{E_{s n_3}^2 E_{n_3 n_4} E_{n_4 s}} + H_{ss}^2 \sum_{n_4 \neq s} \frac{|H_{s n_4}|^2}{E_{n_4 s}^3}$$  \hspace{1cm} (3.1.16)

where $H_{n_4 n_3} = H_{n_3 n_2} = 0$, and the nonvanishing terms give,
\[ E^{(4)} = -E^{(2)} \frac{\left| H_{sp_z} \right|^2}{E_{sp_z}^2} = \left( \frac{u^2 \omega^2}{\lambda} F^2 \right) \times \left( \frac{u^2 \omega^2}{\lambda^2} F^2 \right) = \frac{u^4 \omega^4}{\lambda^3} F^4. \] (3.1.17)

Lastly, the fifth order correction, \( E^{(5)} \), vanishes due to the zero terms, \( H_{n_5n_4} = H_{n_4n_3} = H_{n_3n_2} = 0 \), similar to the third order contribution, \( E^{(3)} \). Finally, up to the 5th order corrections, equation 3.1.12 becomes,

\[
E = 0 + 0 \cdot F + \left( -\frac{u^2 \omega^2}{\lambda} F^2 \right) + 0 \cdot F^3 + \left( \frac{u^4 \omega^4}{\lambda^3} F^4 \right) + 0 \cdot F^5 + O(F^6),
\] (3.1.18)

which is identical to equation 3.1.11.

**Orbital occupancy.** Finite higher order polarisabilities (see equations 3.1.10) imply that the polarisation in the LPM, unlike the shell model, is nonlinear. A further insight in the polarisation process can be obtained by analysing the ground state molecular orbital of the LPM. The changes in the lone pair molecular orbital are quantified by the atomic orbital occupancies (or the molecular orbital coefficients).

For the field along the \( z \)-axis, the model electron lone pair shows the ground state molecular orbital (the lowest energy eigenstate) has contributions only from the \( s \) and \( p_z \) atomic orbitals, and their respective coefficients are given below:

\[
c_s = \frac{C}{\sqrt{1 + C^2}} \quad \text{and} \quad c_{p_z} = \frac{1}{\sqrt{1 + C^2}} \quad \text{where} \quad C = \frac{\lambda + \sqrt{\lambda^2 + 4u^2F^2\omega^2}}{2uF\omega}. \]
(3.1.19)

In the limits of low and high field, the coefficients are changing as:

\[
c_s \to \text{sgn}(u\omega) \quad \text{and} \quad c_{p_z} \to 0 \quad \text{as} \quad F \to 0, \]
(3.1.20a)

\[
c_s \to \frac{1}{\sqrt{2}} \text{sgn}(u\omega) \quad \text{and} \quad c_{p_z} \to \frac{1}{\sqrt{2}} \quad \text{as} \quad F \to \infty
\]
(3.1.20b)
where \( \text{sgn} \) is the sign function, whose definition is \( \text{sgn}(x) = x/|x| \). Since the lone pair charge \( \omega \) is negative, the sign of \( c_s \) is determined by the sign of \( u \), which depends on the choice of radial functions employed in the LPM.

When there is no electric field \((F \to 0)\), the LPM density does not show any polarisation, and the ground state consists of a pure \( s \) atomic orbital. On the other hand, when the field magnitude becomes large \((F \to \infty)\), a contribution of the \( p_z \) atomic orbital is no longer zero, and both occupancies of \( c_s \) and \( c_{p_z} \) converge to \( 1/\sqrt{2} \) in a 1:1 ratio. This result implies, in extremely large static electric field, the LPM polarisation plateaus (saturates) instead of resulting in the dissociation of the electron lone pair. In contrast, the shell model becomes extremely polarised, with the shell moving away off the core to infinity, due to its linear response to the field (in practice, such dissociation of the shell is usually not allowed as the core-shell separations are constrained to remain within a certain distance cut-off).

**Comparison of the lone pair model with the shell model.** The feature of the LPM, which has the saturation limit of the lone pair hybridisation, is advantageous in practical simulations. Without the saturation effect, for example, molecular dynamics and Monte Carlo simulations using the shell model might probe unrealistic atomic configurations, with small interionic distances, that are characterised by extremely large fields, which might result in structural collapse \([117,118]\). Therefore, employing the shell model to represent highly polarisable species like lone pair cations would be particularly prone to result in undesirable extreme core-shell separations. In contrast, the limited polarisation effect of the LPM would be able to allow more robust simulations.
In order to ameliorate the undesirable behaviour of the shell model, the anharmonic shell model (so called \(k_4\), mentioned in section 1.2) is typically used, which has a revised form of the shell spring potential:

\[
E = \frac{1}{2} k_2 r_{cs}^2 + \frac{1}{24} k_4 r_{cs}^4
\]  

with a quartic spring constant \(k_4\). Addition of the quadratic or the anharmonic term – compare with equation 1.6 – results in the non-linear response of the shell polarisation that could have similar dampening effect, shown by the LPM.

For the anharmonic shell with charge, \(q\), the equilibrium shell displacement, \(r_{cs}^{eq}\), in external field \(F\), can be obtained by solving the following equation,

\[
Fq - \frac{\partial E}{\partial r_{cs}} = 0 \rightarrow Fq - k_2 r_{cs} - \frac{1}{6} k_4 r_{cs}^3 = 0,
\]  

which is a cubic equation that gives the real solution where the net force acting on the shell by the spring and the field becomes zero. That is:

\[
r_{cs}^{eq} = \frac{2k_2 - (-3Fk_4^2 q + L)^{2/3}/k_4}{(-3Fk_4^2 q + L)^{1/3}}
\]  

where \(L = \sqrt{k_4^3(8k_2^3 + 9F^2 k_4^2 q^2)}\).

Continuously, the energy is\(^{24}\):

\(^{24}\) The work done by the field \(F\) and the anharmonic spring to the shell.
\[ E_{k4} = - \int_{r_{\text{eq}}}^{r_{\text{eq}}} \left(Fq - k_2 r - \frac{1}{6} k_4 r^3 \right) dr \]

\[ = \frac{(2k_2 - (-3Fk_4^2 q + L)^{2/3}/k_4)}{24(-3Fk_4^2 q + L)^{4/3}} \]

\[ \times \left\{ -24Fq(-3Fk_4^2 q + L) + (2k_2 k_4 - (-3Fk_4^2 q + L)^{2/3})^3/k_4^2 \right. \]

\[ - 12k_2(-3Fk_4^2 q + L)^{2/3}(-2k_2 k_4 + (-3Fk_4^2 q + L)^{2/3})/k_4 \} , \]

and the dipole moment and dipole-polarisability can be obtained by differentiating the energy with respect to the field\(^25\) (see formulae in 3.1.2), which are:

\[ \mu_{k4} = - \frac{\partial E_{k4}}{\partial F} = \frac{(-2k_2 k_4 + (-3Fk_4^2 q + L)^{2/3})}{L(-3Fk_4^2 q + L)^{7/3}} \]

\[ \times 2k_4^2 q(24Fk_4^3 k_4^2 q + 27F^3 k_4^3 q^3 - 4k_4 L - 9F^2 k_4 q^2 L), \]

\[ \alpha_{k4} = \frac{\partial \mu_{k4}}{\partial F} = \frac{(2k_2 k_4 + (-3Fk_4^2 q + L)^{2/3})}{L(-3Fk_4^2 q + L)^{10/3}} \]

\[ \times 4k_4^4 q^2(-18Fk_4^3 k_4^2 q - 27F^3 k_4^3 q^3 + 2k_4^2 L + 9F^2 k_4 q^2 L). \]

To highlight advantages and disadvantages of the three different models considered: (i) the lone pair model (LPM), (ii) harmonic shell model (SM), and (iii) anharmonic shell model (k4), we have compared their energy, dipole moment and dipole-polarisability as a function of electric field shown in figure 3.1. The analytical expressions of those three functions by different models, (i) - (iii), are given in equations 3.1.9, 3.1.10 and 3.1.24. Furthermore, for a fair comparison, the same

\(^25\) In practice, the Mathematica software was used to carry out the integration and deriving the dipole and dipole-polarisability.
Figure 3.1 Observable properties of a lone pair cation (LPM modelled Sn$^{2+}$ is used for demonstration): (a) energy, $E$, (b) dipole moment, $\mu$, and (c) dipolar polarisability, $\alpha$, as a function of uniform electric field, $F$, for three different models: harmonic shell model (SM)$^a$, anharmonic shell model ($k4)^b$ and lone pair model (LPM)$^c$ – see legend.

With the same charge for the models, $\omega = q = -2$, the following parameters were employed to satisfy the conditions in 3.1.25:

- $^a$ spring constant of the shell model, $k = 16.333$ (eV/Å$^2$);
- $^b$ spring constant of the anharmonic shell model, $k_2 = 16.333$ (eV/Å$^2$) and $k_4 = 100.861$ (eV/Å$^4$); and
- $^c$ LPM parameter from the onsite Hamiltonian, $\lambda = 15.87$ (eV), see equation 2.1.6.
charge\textsuperscript{26}, \( \omega = q = -2 \), dipole-polarisability, \( \alpha \), and the second hyperpolarisability, \( \gamma \), in the limit of zero external field, were used for both LPM and the shell model\textsuperscript{27};

\[
\left( \frac{\partial^2 E_{\text{SM}}}{\partial F^2} \right)_{F=0} = \left( \frac{\partial^2 E_{\text{LPM}}}{\partial F^2} \right)_{F=0} \Leftrightarrow \alpha_{\text{SM}}(0) = \alpha_{\text{LPM}}(0), \tag{3.1.25a}
\]

\[
\left( \frac{\partial^4 E_{k4}}{\partial F^4} \right)_{F=0} = \left( \frac{\partial^4 E_{\text{LPM}}}{\partial F^4} \right)_{F=0} \Leftrightarrow \gamma_{k4}(0) = \gamma_{\text{LPM}}(0); \tag{3.1.25b}
\]

the LPM parameter of \( \lambda = 15.87 \) (eV) – the LPM onsite potential, see equation 2.1.6 – was chosen, taken from the difference between the second and third ionisation potentials of Sn atom, given in table 3.1 where the details were discussed in section 2.1.

For all three functions shown in figure 3.1.(a)-(c), the three different models are in agreement in the weak field region, approximately below 2.5 V/Å, but start noticeably diverging of each other above the region. Of the two shell models, the harmonic shell model shows the largest deviation from the LPM as seen in the energy and dipole moment graphs, 3.1.(a) and (b), due to its constant dipole-polarisability shown in graph (c). The dipole-polarisability of the \( k4 \) anharmonic shell model shows a similar saturation effect to that of LPM whose effect is expected by the occupancy analysis of the orbitals in equations 3.1.20. However, unlike the rapid saturation profile of the LPM dipole-polarisability, dropped nearly to zero at 20 V/Å, the dipole-polarisability of the \( k4 \) anharmonic shell is still far away from zero but showing a long tail, i.e., slowly convergent.

Here, the strongest applied field of 20 V/Å represents an upper limit on the electric fields in a realistic system, where it would be associated with low coordinated sites around point defects or

\textsuperscript{26} The charge of LPM electron lone pair, \( \omega \), and the shell, \( q \).

\textsuperscript{27} Note that the first hyperpolarisability, \( \beta \), of the both LPM and \( k4 \) models are zero in the limit of zero field.
surfaces [119–121]. The local electric fields in perfect ionic crystals would be an order of magnitude lower, for instance in MgO, it ranges from 0 to 2 V/Å [119]; however, such strong fields may need to be considered in particular environments, for example, when modelling nanoclusters exposing their surfaces, indeed, in a cation-anion dimer carrying charges of $q = \pm 2$, each ion generates field of 12 V/Å at its separation distance of 2 Å; and such a field may cause a break off of the shell (or the excessive stretching of the harmonic spring) during energy minimisation process.

<table>
<thead>
<tr>
<th>Species ($M$)</th>
<th>2$^{nd}$ IE / eV</th>
<th>3$^{rd}$ IE / eV</th>
<th>3$^{rd}$ IE − 2$^{nd}$ IE / eV</th>
<th>HOMO-LUMO $M^{2+}$/eV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sn</td>
<td>14.63</td>
<td>30.50</td>
<td>15.87</td>
<td>8.376</td>
</tr>
<tr>
<td>Pb</td>
<td>15.03</td>
<td>31.94</td>
<td>16.91</td>
<td>10.368</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Species ($M$)</th>
<th>3$^{rd}$ IE / eV</th>
<th>4$^{th}$ IE / eV</th>
<th>4$^{th}$ IE − 3$^{rd}$ IE / eV</th>
<th>HOMO-LUMO $M^{3+}$/eV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bi</td>
<td>25.56</td>
<td>45.3</td>
<td>19.74</td>
<td>12.519</td>
</tr>
</tbody>
</table>

**Table 3.1** Ionisation potentials, or energies (IE) of lone pair cations. Experimental values taken from the reference [64], and the relevant HOMO-LUMO gaps were calculated by using the FHI-AIMS code, using provided intermediate basis set, at the DFT-PBEsol level of theory.

In table 3.1, we presented relevant ionisation energies of lone pair species, and the DFT calculated HOMO-LUMO energy differences of isolated lone pair cations, Sn$^{2+}$, Pb$^{2+}$ and Bi$^{3+}$, which have the valence electron configuration of $s^2\,p^0$. Recall that we decided to determine the model parameter $\lambda$ (eV) by using the difference in the ionisation energies, 3$^{rd}$ IE − 2$^{nd}$ IE for Sn$^{2+}$ and Pb$^{2+}$; and 4$^{th}$ IE − 3$^{rd}$ IE for Bi$^{3+}$ based on the model single electron picture (see section 2.1, also figure 2.1); and furthermore, that the HOMO-LUMO energy differences are also presented as a reference based on the multiple electron picture, which could be used as the $\lambda$ parameters (or at least provide a useful test that the values obtained from ionisation energies are on the same scale with those) that model the energy difference between the lower $s^2$ and upper $p^0$ state – see section 3.1.3.
3.1.2 Quadrupole polarisabilities of the lone pair model

Next we focus on the quadrupole moment, dipole-quadrupole and quadrupole-quadrupole polarisabilities of the LPM. In general, these quadrupole related properties are observed in the presence of nonzero electric field gradient, \( F_{ij} \), see equation 3.1.1 – and to investigate the model properties, the Hamiltonian in 3.1.3 should be revised as:

\[
\hat{H} = \hat{H}^0 - \omega (\tau_i F_i + \tau_j F_{ij}) \tag{3.1.26}
\]

Using the basis set of the LPM, similar to the volume integrals with respect to the electric field in 3.1.5, there arise two further LPM parameters, \( \nu \) and \( \omega \), by the electric field gradient, which are determined by the radial wavefunction:

\[
\langle s|x^2|s\rangle = \langle s|y^2|s\rangle = \langle s|z^2|s\rangle = \nu, \tag{3.1.27a}
\]

\[
\langle p_x|x^2|p_x\rangle = \langle p_y|y^2|p_y\rangle = \langle p_z|z^2|p_z\rangle = 3\omega, \tag{3.1.27b}
\]

\[
\langle p_x|y^2|p_x\rangle = \langle p_x|z^2|p_x\rangle = \langle p_y|x^2|p_y\rangle = \langle p_y|z^2|p_y\rangle = \langle p_z|x^2|p_z\rangle = \langle p_z|y^2|p_z\rangle = \langle p_z|xz|p_z\rangle = \langle p_z|zx|p_z\rangle \tag{3.1.27c}
\]

and the rest of integrals are zero due to the orbital symmetry. Thus, the Hamiltonian matrix becomes:

\[
H = H^0 + H^1 = \begin{pmatrix}
D_1 & -\omega F_1 u & -\omega F_2 u & -\omega F_3 u \\
-\omega F_1 u & \lambda + D_2 & D_5 & D_6 \\
-\omega F_2 u & D_5 & \lambda + D_3 & D_7 \\
-\omega F_3 u & D_6 & D_7 & \lambda + D_4
\end{pmatrix}
\]

where \( D_1 = -(F_{11} + F_{22} + F_{33})\omega \nu \), \( D_2 = -(3F_{11} + F_{22} + F_{33})\omega \nu \),

\[
D_3 = -(F_{11} + 3F_{22} + F_{33})\omega \nu, \quad D_4 = -(F_{11} + F_{22} + 3F_{33})\omega \nu,
\]

\[
D_5 = -(F_{12} + F_{21})\omega \nu, \quad D_6 = -(F_{13} + F_{31})\omega \nu \quad \text{and} \quad D_7 = -(F_{23} + F_{32})\omega \nu.
\]
Finding the ground state energy as function of \( F_i \) and \( F_{ij} \) from the Hamiltonian matrix above is quite cumbersome, since the eigenvalue equation of the matrix involves finding the lowest root of a quartic equation with general coefficients (although all four roots and eigenvalues of \( H \) are real as the matrix is real and symmetrical). Therefore, we can instead consider particular cases of which only required fields are adaptively applied in order to calculate the model quadrupole moment, dipole-quadrupole and quadrupole-quadrupole polarisabilities.

For instance, to obtain an element of dipole-quadrupole polarisability, \( A_{123} \), one can choose fields only with \( F_1 \) and \( F_{23} \) components (the rest of field elements are set to zero), and the Hamiltonian matrix, equation 3.1.28, simplifies to:

\[
H = \begin{pmatrix}
0 & -\omega F_1 u & 0 & 0 \\
-\omega F_1 u & \lambda & 0 & 0 \\
0 & 0 & \lambda & -\omega F_{23} w \\
0 & 0 & -\omega F_{23} w & \lambda \\
\end{pmatrix}
\]

(3.1.29)

The characteristic polynomial of the matrix can be easily factorised, and consequently the eigenvalues, \( E_i \), and the eigenvectors, \( v_i = \{c_s, c_{p_x}, c_{p_y}, c_{p_z}\} \), of the matrix are:

\[
\{E_1, E_2, E_3, E_4\} = \\
\left\{ \frac{1}{2} \left( \lambda - \sqrt{\lambda^2 + 4u^2F_1^2\omega^2} \right), \frac{1}{2} \left( \lambda + \sqrt{\lambda^2 + 4u^2F_1^2\omega^2} \right), \lambda - \omega F_{23} w, \lambda + \omega F_{23} w \right\},
\]

(3.1.30)

\[
v_1 = \left\{ \frac{\lambda + \sqrt{\lambda^2 + 4u^2F_1^2\omega^2}}{2uF_1 q}, 1, 0, 0 \right\}, \quad v_2 = \left\{ \frac{\lambda - \sqrt{\lambda^2 + 4u^2F_1^2\omega^2}}{2uF_1 q}, 1, 0, 0 \right\},
\]

\[
v_3 = \{0, 0, 1, 1\} \text{ and } v_4 = \{0, 0, 1, -1\}.
\]
Note that the eigenvalues (or energies) are not ordered by their value, and the eigenvectors are not normalised for simplicity.

In 3.1.30, none of the energies, $E_{1-4}$, includes dependency on a product of the field and its gradient, which implies $A_{123} = 0$ (cf. equations 3.1.2b and 3.1.2c). In the same manner, all other elements of dipole-quadrupolar polarisability were investigated and turned out to be zero in the limit of zero external field and field gradient, $A_{kl}^j(F_k = 0, F_{ij} = 0) = 0$ for all $i, j$ and $k$ (see Appendix A for further details)

The energy expressions in equations 3.1.30 are particularly useful to understand why the dipole-quadrupole polarisability of the LPM is zero. The ground state energy $E_1$ (by taking $F_1 \to 0$ and $F_{23} \to 0$, $E_1$ gives the lowest value) is does not depend on the applied field gradient, $F_{23}$, but does depend on the field, $F_1$, (similar to the case of uniform electric field – see equation 3.1.10a). The corresponding eigenvector, $v_1$, shows the lone pair hybridisation where the first and the second elements related to the occupancies of $s$ and $p_x$ atomic orbitals. The second state with the energy of $E_2$ is the anti-state of $E_1$ whose polarisation is against the field according to its eigenvector, $v_2$. On the other hand, the other excited states with eigenvalues of $E_3$ and $E_4$ are determined only by the field gradient, do not show any $sp$-hybridisation, and remain as pure $p$ character, see $v_3$ and $v_4$. Consequently, the induction of the electric dipole transition, $s \to p$, by the field gradient is turned out to be forbidden, and such transition associated with the lone pair hybridisation is mainly affected by uniform electric fields

28 Instead of the manual repetitions for estimating, $A_{ijk}$, the time-independent perturbation theory will be useful for this case. The analysis in the appendix also contains the result of quadrupole-quadrupole polarisability, $C_{ijkl}$.

29 Pure electric field gradient cannot induce the transition but will do have the effect when it comes together with a nonzero uniform electric field. Such field would be, for instance, generated by an ion located adjacent to a lone pair cation.
Using the same approach that has been used to calculate $A_{klj}$, the quadrupole-quadrupole polarisability, $C_{ijkl}$, is also turned out to be zero for all $i, j, k$ and $l$ in the vanishing uniform electric field gradient ($F_{ij} \rightarrow 0$). In zero uniform field ($F_i = 0$), the ground state energy is function of field gradient only ($F^{ij}_{\rightarrow 0}$), and the state only contains a contribution from the $s$ atomic orbital that is the lone pair does not polarise. Indeed, the $H$ matrix becomes block diagonal (see equation 3.1.28):

$$
H = \begin{pmatrix}
D_1 & 0 & 0 & 0 \\
0 & \lambda + D_2 & D_3 & D_4 \\
0 & D_5 & \lambda + D_3 & D_6 \\
0 & D_6 & D_7 & \lambda + D_4
\end{pmatrix}
$$

(3.1.31)

Accordingly, in a weak field gradient, the ground state energy will be $D_1$ with the corresponding eigenvector of $\{1,0,0,0\}$, i.e., the ground state will retain pure $s$ character, and the other three excited states will span the $p$ orbital subspace. This is due to the forbidden nature of the $s \rightarrow p$ transition by the electric field gradient (compare to the equations 3.1.27 above),

$$
\langle s|r_ir_j|p_x \rangle = \langle p_x|r_ir_j|s \rangle = \langle s|r_ir_j|p_y \rangle = \langle p_y|r_ir_j|s \rangle = \langle s|r_ir_j|p_z \rangle = \langle p_z|r_ir_j|s \rangle = 0 \quad \text{for all } i, j \text{ and } k.
$$

(3.1.32)

In order to introduce the effect of non-zero dipole-quadrupolar and quadrupole-quadrupole polarisabilities, it is necessary to expand the basis set of the LPM to $d$ orbitals. Indeed, the dipole polarisability relates the induced dipole to external electric fields via electric transition dipole, $s \rightarrow p$, as shown in equation 3.1.5. Similarly, if the LPM basis set included $d$ orbitals then the field gradient induced transitions, $s \rightarrow d$, will be allowed\(^{30}\). This implies that the $d$ orbitals are directly

\(^{30}\) For instance, for field gradient along the $z$-axis, there will be non-zero transitions for $s \rightarrow d$, e.g.

$$
\langle s|xz|d_{xz} \rangle = \langle dxz|xz|s \rangle = \langle s|yz|dy_2 \rangle = \langle dy_2|yz|s \rangle = \sqrt{3}/2 \langle s|z^2|d_{xz} \rangle = \sqrt{3}/2 \langle d_{xz}|z^2|s \rangle \neq 0.
$$
associated with the dipolar-quadrupolar and quadrupolar-quadrupolar polarisabilities. Thus, with the current basis set, including the $s$ and the $p$ atomic orbitals, the LPM cannot capture the effect of quadrupole related polarisabilities. Likewise the shell model does not have the quadrupole related effect [33,122], and with the basis we could carry out the fair comparison in applications, using the both LPM and the shell model where the details will be delivered later in chapter 4.

#### 3.1.3 Lone pair charge and onsite energy separation parameters

LPM parameters including the charge of the electron lone pair, $\omega$, and the onsite energy separation between $s$ and $p$ states for an isolated lone pair cation, $\lambda$, are key in determining the lone pair energy, dipole moment and polarisability. So far, we have assumed that the charge, $\omega$, should be in the range of [-2.0, -1.0], implicitly including an electron-electron repulsion within the lone pair; and that the onsite parameter, $\lambda$, would be set by the difference in relevant ionisation potentials, e.g., the energy difference between the second and the third ionisation energies of tin and lead (see section 2.1).

In this section, in order to narrow down the model parametric ranges of $\omega$ and $\lambda$, we have mapped the dipole moment and the polarisabilities on a physically relevant domain of those parameters and compared those with calculated DFT values of the lone pair cations in our interest, namely, Sn$^{2+}$, Pb$^{2+}$ and Bi$^{3+}$.

**Preparation of DFT reference data.** The DFT reference data include the energies$^{31}$ and the dipole moments of lone pair cations, Sn$^{2+}$, Pb$^{2+}$ and Bi$^{3+}$ (in specific oxidation states giving the lone

---

$^{31}$ The energy difference between an isolated cation and the cation to external field, i.e. $E = E(F) - E(0)$. 

96
pair electron configuration, \( s^2p^0 \), under uniform electric field, calculated by the FHI-AIMS code. The PBEsol exchange and the correlation density functional was employed for their DFT calculations [113]. The field values were mapped from 0.0 to 10.0 V/Å with an increment of 0.25 V/Å. Here, the maximum field value was set to 10.0 V/Å, which was chosen considering that a point charge of -2 (for instance, same with the formal charge of an oxygen anion) will generate electric field with magnitudes from 9.40 to 5.69 V/Å for the distances between 1.75 and 2.25 Å where the distance range covers the typical \( M-O \) (\( M = \text{Sn}, \text{Pb} \) and \( \text{Bi} \)) bond lengths in their close-packed oxide materials [2,96,123,124].

**Comparison scheme.** In order to quantitatively characterise the difference between the LPM and the DFT predicted energies and dipole moments for the range of external electric field, we employed the root-mean-square deviations (RMSD) of \( \Delta E \) and \( \Delta \mu \), which are defined as:

\[
\Delta E = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left\{ E_{\text{DFT}}(F_i) - E_{\text{LPM}}(F_i, \omega, \lambda) \right\}^2} \quad \text{and} \quad \Delta \mu = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left\{ \mu_{\text{DFT}}(F_i) - \mu_{\text{LPM}}(F_i, \omega, \lambda) \right\}^2}.
\]

(3.1.33)

In addition, we compared the dipole polarisabilities of the LPM and the DFT at zero-field. Here, the dipole polarisabilities of DFT for different cations were calculated numerically using the method of central differences:

\[
\alpha_{\text{DFT}} = \frac{\mu_{\text{DFT}}(0 + \Delta F) - \mu_{\text{DFT}}(0 - \Delta F)}{2\Delta F}.
\]

(3.1.34)

with \( \Delta F = 0.0005 \) eV/Å. The calculated dipole polarisabilities, \( \alpha_{\text{DFT}} \), of \( \text{Sn}^{2+} \), \( \text{Pb}^{2+} \) and \( \text{Bi}^{3+} \) are 0.14899, 0.13187 and 0.09834 (Å\(^2\)/eV), respectively, where the values are in the order, \( \alpha_{\text{Bi}^{3+}} < \)
\[ \alpha_{\text{Pb}^2+} < \alpha_{\text{Sn}^2+}, \] which matches well with the degree of the lone pair induced structural distortions in their oxide crystal structures, reported by Walsh et al. [15], i.e., more polarisable species have more stereo-chemically active lone pairs.

**Parametric space analysis.** Contour maps of the deviations, \( \Delta_E \) (eV) and \( \Delta_\mu \) (e Å), are presented in figure 3.2 for the three lone pair cations. In each plot, the values of experimentally measured ionisation potential differences and calculated DFT HOMO-LUMO gap are marked as red solid and dashed lines, respectively, the values used can be found in table 3.1, section 3.1.1. These values were used to determine the scanning range of 5 to 32 eV for \( \lambda \), including the ranges marked by the lines, which we believe the domains are physically relevant. Parameter \( \omega \) represents the charge of electron lone pair, varying from -2.5 to -0.5, which includes -2, the formal charge of two electrons.

In each plot of figure 3.2, one can find the valleys showing relatively low RMSD (\( \Delta_E \) and \( \Delta_\mu \)) values. In common to all plots, the parameter windows of \( \lambda \) (marked by red solid and dashed lines) cross such valleys, where the conditions of \( \Delta_E < 0.4 \) and \( \Delta_\mu < 0.16 \) are satisfied; conversely, if \( \lambda < 10 \) eV, \( \Delta_E \) and \( \Delta_\mu \) of \( \text{Sn}^{2+} \), for instance, cannot satisfy such inequalities no matter what value of \( \omega \) is chosen. We conclude that in practice we can choose the LPM parameter \( \lambda \) sensibly, to match experimentally observed or accurately calculated characteristics of lone pair cations.

In practice, the LPM parameter \( \lambda \) for each cation was chosen based on the species ionisation potentials. Whilst \( \lambda \) is fixed, parameter \( \omega \) was tuned to reproduce the polarisability, \( \alpha^{\text{DFT}} \), calculated using the DFT method or at least the parameter pair was chosen that within the sensible range as given in figure 3.2. Assuming the polarisability of a model cation is reliable, a lone pair cation polarisation, or a dipole induced by coordinating anions will also be predicted correctly for
Different environments and systems of different complexity, including multiple lone pair cations and anions – see the model applications in chapter 4. As set forth the LPM description is within the ionic picture, tuning the lone pair charge also affects the core charge of a lone pair cation subject to a constraint on the sum of the two charges to be equal to the formal charge of the cation.
Since the core charge has a positive value it will have a strong effect on a cation-anion bond formation, viz., for instance, how far the negatively charged shell of an anion is displaced under the Coulomb interaction with the cation core, which in turn determines the ionic dipole moments of both the cation and anion. Note that the LPM parameters were found with the criterion for the value of $\alpha^{\text{LPM}}$ to deviate from the target $\alpha^{\text{DFT}}$ for each cation by less than 0.1%, which means that the found solutions do not coincide with the local minima on the RMSD surfaces in figure 3.2.

<table>
<thead>
<tr>
<th>Species</th>
<th>$\lambda$ /eV</th>
<th>$\omega$ /e</th>
<th>$\alpha^{\text{LPM}}$ (err % w.r.t $\alpha^{\text{DFT}}$) (Å$^2$/eV)</th>
<th>$\Delta_E$</th>
<th>$\Delta_\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sn$^{2+}$</td>
<td>15.87</td>
<td>-1.567</td>
<td>0.1503 (0.06)</td>
<td>0.838</td>
<td>0.321</td>
</tr>
<tr>
<td>Pb$^{2+}$</td>
<td>16.53</td>
<td>-1.424</td>
<td>0.1326 (0.04)</td>
<td>0.778</td>
<td>0.309</td>
</tr>
<tr>
<td>Bi$^{3+}$</td>
<td>19.74</td>
<td>-1.343</td>
<td>0.0987 (0.04)</td>
<td>0.327</td>
<td>0.134</td>
</tr>
</tbody>
</table>

Table 3.2 The LPM parameters of $\{\omega, \lambda\}$ for three different lone pair cations fitted to the dipole polarisabilities of the DFT-PBEsol calculated results where $\Delta_E$ and $\Delta_\mu$ are the root-mean-square deviations of the LPM energy and the dipole moment for each cation measured with respect to the DFT references – details are summarised in figure 3.2.

The particularly notable point here is that if the formal lone pair charge $\omega = -2$ was employed, the required values of the parameter $\lambda$ would be found outside of the windows marked (ranges within solid and dashed red lines) in the plots of figure 3.2. Specifically, 25.87, 32.64 and 43.80 eV are required for $\lambda_{\text{Sn}}$, $\lambda_{\text{Pb}}$ and $\lambda_{\text{Bi}}$, respectively, to keep the value of $\alpha^{\text{LPM}}$ within the error with respect to $\alpha^{\text{DFT}}$ as in table 3.2; and those values of $\lambda$ are significantly off from their physically relevant windows determined by the experimentally observed and the DFT calculated ionisation potentials. Consequently, the optimal values of the formal lone pair charges, found within the windows, allow to account for the electron-electron repulsion within the lone pair explicitly absent in the model, which will be further discussed in section 3.2.1 by comparing with the Hartree’s method.
3.2 The effective valence state of a lone pair cation

Next, we investigated the interaction of the model lone pair cation with an external point charge of $q = -1$. Recall that the lone pair model we develop employs a minimum basis set consisting only of $s$ and $p$ atomic orbitals. In previous section 3.1.2, we have shown that the integration of the $d$ orbitals to the model may be necessary to introduce the effect of higher-order electric multipoles and related polarisabilities where especially the properties are related to the electric quadrupolar properties. Such an expansion of the basis set would make the LPM be more general, however, it might be too excessive. Therefore, here, in order to verify the necessity of including $d$ orbitals, the valence state of the Pb$^{2+}$ lone pair cation was investigated with a higher-level DFT theory, and compared with the model.

3.2.1 Effective valence state of the lone pair ion model in comparison with DFT

For the comparison with the DFT results, the LPM calculations were carried out and the relevant theory is provided in section 2.3. Note that the LPM calculations were based on the model parameters, $\{\lambda, \omega\} = \{16.53, -1.424\}$, taken from table 3.2, which are for modelling the LPM Pb$^{2+}$ cation, and the cation was fixed at the origin interacting with the point charge on the $z$-axis at various distances. The results of the LPM results as before were compared with the DFT-PBEsol level of theory.

Orbital occupancy of the valence state of Pb(II) cation. In figure 3.3 (a), we compare $6p_z$ and $5d_{x^2}$ orbital occupancies of the Pb$^{2+}$ HOMO, calculated by using DFT; and the $p_z$ occupancy predicted by the LPM and its scaled LPM $p_z$ occupancy. The position of the point charge on the $z$-
Figure 3.3 (a) Atomic orbital contributions to the HOMO of Pb\(^{2+}\) cation interacting with a point charge \(q = -1\) as a function of the separation distance. Comparison is made between the PBEsol DFT 6\(p_z\) and 5\(d_{z^2}\) orbitals and the LPM (black dashed line) and its uniformly scaled LPM (red dashed line) \(p_z\) orbital occupancies. LPM parameters \(\{\lambda, \omega\} = \{16.53, -1.424\}\) are used. (b) \(p_z\) orbital occupancy ratio between DFT and LPM, calculated as a function of \((-1/\omega)\) whilst using fixed \(\lambda = 16.53\) eV. The red circled data-points with \(\{a, b, c\}\) correspond to \(\omega \{-1.640, -1.532, -1.424\}\); the leftmost and the rightmost red circles mark -2.0 and -1.0 values of \(\omega\). The dashed line represents a linear fit to the given five points whose equation is shown in the plot.

\(^a\) For instance, the datapoint c is the average of \(\{p_z^{\text{DFT}}/p_z^{\text{LPM}}\}\) from fig. (a) with \(\omega = -1.424\), and the rest of datapoints with different lone pair charge \(\omega\) were measured by the same manner. Accordingly, the error bars are shown with the confidence interval of 95\% for each datapoint.
axis is varied in the range of [2,0.5,0] – in Å unit – including the characteristic interionic distance between Pb and F of 2.57 Å in the closed-packed (fluorite) phase of PbF₂ where F ions are in the oxidation state of -1 [125]. Note that the squared sum of DFT calculated orbital contributions – by 6s, 6pₓ and 5dₓ² – accounts for $>99.5\%$ occupancy of the HOMO, and in LPM, the squared sum of pₓ and s is exactly equal to 1.

At the first glance of figure 3.3 (a), the occupancy of the 5dₓ² orbital may look substantial. The ratio of the orbital occupancies, $6p_{\text{DFT}}^z / \{(5d_{\text{DFT}}^{z^2})^2 + (6p_{\text{DFT}}^z)^2\}^{1/2}$, however, is 0.95, which implies the effect of $6p_{\text{DFT}}^z$ is much more significant than that of $5d_{\text{DFT}}^{z^2}$ state\(^{32}\). Although, such a high ratio is only valid for the given range, i.e. below 2.0 Å, the effect of $5d_{\text{DFT}}^{z^2}$ becomes significant, such a close distance is unlikely to be observed except for extreme conditions as realised for example under very high pressures or high energy ionic collisions [126–129]. Accordingly, for typically accessible interionic distances above 2.0 Å was explored (for the materials in interest containing lone pair cations) with an example of the point charge, and it is valid to neglect the effect of d states.

**Scaling the polarisation of valence state, bridging the LPM with DFT.** The occupancies of $p_{\text{LPM}}^z$ and $6p_{\text{DFT}}^z$ are in qualitative agreement as seen in figure 3.3 (a), where both are monotonically decreasing as the point charge goes away from the lone pair cation, and, thus the lone pair polarisation becomes weaker. However, the magnitude of $p_{\text{LPM}}^z$ occupancy is evidently higher than

---

\(^{32}\) The occupancies of the rest of DFT 5d orbitals, 5dₓᵧ, 5dₓz, 5dᵧz and 5dₓ²−y², are negligibly small with the magnitudes below $10^{-28}$. 
that of $6p_{z}^{DFT}$, e.g., at the distance of 2.0 Å, the amplitude of $p_{z}^{LPM}$ as above 0.16 is higher than $6p_{z}^{DFT}$ as 0.11, and also for the rest of the distances.

In theoretical aspect, such discrepancy shown in figure 3.3 (a) is not that surprising. Since the electron lone pair of the LPM is modelled as a single electron molecular orbital with the model effective charge, $\omega$, which partially occupies $p_{z}^{LPM}$; on the other hand, DFT depicts individual electrons explicitly and the occupancy of $6p_{z}^{DFT}$, shown in the figure, corresponds to the partial occupancies of the one of two electrons forming the HOMO of Pb$^{2+}$ cation, i.e., the HOMO is occupied by two electrons and $6p_{z}^{DFT}$ shows the partial occupancy of each electron. Although the two approaches describe the HOMO of Pb$^{2+}$ cation differently, we have still found a useful scaling factor between the $p_{z}^{LPM}$ and $6p_{z}^{DFT}$, that is:

$$6p_{z}^{DFT} = \frac{0.869}{|\omega|} p_{z}^{LPM} \quad (3.2.1)$$

where the result of scaling can be examined in figure 3.3 (a) by the scaled $p_{z}^{LPM}$ shown as the red dashed line in the plot that gives good matching with $6p_{z}^{DFT}$. Going further, figure 3.3 (b) shows the ratio, $6p_{z}^{DFT}/p_{z}^{LPM}$, as a function of $-1/\omega$ where the data points are given as red circles, and the dashed line is a linear fit whose slope is equal to the 0.869 constant in equation 3.2.1.

In figure 3.3 (b), the physically instructive feature can be found at the right extreme where the extrapolated line reaches $6p_{z}^{DFT}/p_{z}^{LPM} = 1$ at $|\omega| = 0.869e$ (marked by the breaking arrow ends on the x-axis). The result shows that to reproduce the occupancy of $6p_{z}^{DFT}$ with using LPM, the required effective charge on the lone pair should be $-0.869e$ rather than $-1.0e$, or $-1.738e (-0.869e \times 2)$ to fully reproduce the effect of two electrons occupying the HOMO of Pb$^{2+}$ cation, calculated by using DFT. Even if we include the contribution from the $5d^{DFT}$ orbitals, the ratio of
\{(d_z^{\text{DFT}})^2 + (p_z^{\text{DFT}})^2\}^{1/2}/p_z^{\text{LPM}} \text{ gives the slope of 0.879, which is only 0.01 higher}\footnote{Such difference cannot be fully explained by the presence of 5d_z^{\text{DFT}} orbital contributions, since the amplitude of 5d_z^{\text{DFT}} is still too small to compensate the discrepancy between } p_z^{\text{LPM}} \text{ and } 6p_z^{\text{DFT}} \text{ that shown in figure 3.3 (a)}. \text{ Thus, we concluded that the missing charge contribution of -0.131e within the LPM is mainly due to the absence of the electron-electron repulsion interaction within the lone pair model, which has been briefly discussed in section 3.1.3.}

**Validation of the effective charge employed in LPM.** Instead of a single electron molecular orbital, to explicitly represent the two electrons within the lone pair electron density, in the Hartree approximation, the two-electron wavefunction can be written as \(\Psi(\mathbf{r}_1, \mathbf{r}_2) = \psi(\mathbf{r}_1)\psi(\mathbf{r}_2)\). Note that the anti-symmetrisation constraint on the orbitals does not apply, as the two electrons of opposite spin are always paired in the lone pair and occupy the ground state, and the spin effect also neglected in the LPM. To determine the energy of electrons interacting with a point charge \(q\), we should solve the following equations:

\[
\left[\hat{h}^0 + j - \frac{q}{r}\right] \psi_1 = \varepsilon_1 \psi_1 \tag{3.2.2a}
\]

\[
\left[\hat{h}^0 + j - \frac{q}{r}\right] \psi_2 = \varepsilon_2 \psi_2 \tag{3.2.2b}
\]

where operator \(\hat{h}^0\) accounts for the electron’s kinetic energy and potential energy in the field of the nucleus, which is equivalent to the onsite Hamiltonian in the LPM; \(j\) represents the electron-electron repulsion term, which is the Coulomb potential experienced by one electron due to the average field generated by the other electron in the counterpart molecular orbital.
Here, if the electron pair is in the lowest energy state – could be determined by the SCF method, analogous to the case of the He atom [130] – both electrons will occupy the same ground state, i.e., $\psi_1 = \psi_2$, and the electronic energy can be written as:

$$\varepsilon = \varepsilon_1 + \varepsilon_2 - \langle j \rangle = 2\langle h^0 \rangle - 2\langle q/r \rangle + \langle j \rangle$$

where $\varepsilon_1 = \varepsilon_2 = \langle h^0 \rangle - \langle q/r \rangle + \langle j \rangle$. \hfill (3.2.3)

In equation 3.2.3, as molecular orbitals are equal, so are energies of each of the electrons, and the expectation value $\langle j \rangle$ is subtracted to remove the doubly counted electron-electron repulsion energy contribution.

In the LPM, on the other hand, the model describes the electron lone pair as a single electron molecular orbital, $\psi$, with the effective lone pair charge, $\omega$, and thus an analogous lone pair equation to equation 3.2.2 can be written as:

$$\left[ \hat{H}^0 - \frac{q}{r} |\omega| \right] \psi = E\psi.$$ \hfill (3.2.4)

One notable difference in equation 3.2.4 is that there is no explicit electron-electron repulsion term $\langle j \rangle$ in equation 3.2.2, and instead there appears a scaling factor $|\omega|$ in front of the term, $q/r$, representing the Coulomb interaction with the point charge. Thus, the electronic energy is:

$$E = \langle H^0 \rangle - |\omega|\langle q/r \rangle.$$ \hfill (3.2.5)

We focus here on the interaction of a lone pair with external environment. We note an inequality,

$$-2\langle q/r \rangle + \langle j \rangle > -|\omega|\langle q/r \rangle.$$ \hfill (3.2.6)
which strictly holds if \( \omega = -2 \) and \( \hat{\mathbf{H}}^0 = \hat{\mathbf{H}}^0 \), due to the positive contribution of the electron-electron repulsion, \( \langle j \rangle \), in the explicit Hartree model. The excessive positive contribution by \( \langle j \rangle \) could be separately absorbed into: (i) an onsite contribution (which would be equal to \( \langle j \rangle \) in the absence of the point charge); and (ii) a polarisation related term. The former contribution can be shifted to the LPM onsite Hamiltonian and the latter remaining polarisation term assumed to be proportional to the average potential of the point charge (the first order perturbation theory). The resultant factor summed with the formal charge of -2 will constitute the effective charge \( \omega \geq -2.0 \) of the LPM, which thus includes implicitly the polarisation electron-electron repulsion effect.

### 3.2.2 Dipole moments and the relationship between the valence states of LPM and DFT

Equation 3.2.1 can be used to establish a relationship between the LPM and the DFT calculated dipole moments of Pb\(^{2+}\). Again, for the lead cation interacting with a point charge on the \( z \)-axis, the dipolar moment of LPM can be calculated as:

\[
\mu_{z}^{\text{LPM}} = \langle \psi | \omega z | \psi \rangle = 2\omega c_s c_{p_z} \langle s | z | p_z \rangle = 2\omega u c_s c_{p_z} = 2\omega u \left( 1 - c_{p_z}^2 \right)^{1/2} c_{p_z} \tag{3.2.7}
\]

where \( u = \langle s | z | p_z \rangle \) is the electric dipole transition moment taken from equation 3.1.5, and the normalisation of the molecular orbital coefficients \( \left( c_{s}^2 + c_{p_z}^2 \right)^{1/2} = 1 \) is used to exclude \( c_s \).

By substituting equation 3.2.1 into 3.2.7, one can derive a relationship between the LPM and DFT dipole moments. First, we obtain dipoles using the \( c_{p_z}^{\text{DFT}} \) as an independent variable:

\[
\mu_{z}^{\text{LPM}} = 2\omega u \left( 1 - \left( 0.869^{-1} |\omega| c_{p_z}^{\text{DFT}} \right)^2 \right)^{1/2} \left( 0.869^{-1} |\omega| c_{p_z}^{\text{DFT}} \right) \tag{3.2.8}
\]

and
\[
\mu_z^{\text{DFT}} = -2 \times \langle \psi | z | \psi \rangle \approx -2 \times 2 c_s c_{p_z} \langle s | z | p_z \rangle = -2 \times 2 u' \left\{1 - \left( c_{p_z}^{\text{DFT}} \right)^2 \right\}^{1/2} \frac{c_{p_z}^{\text{DFT}}}{c_{p_z}}
\]

(3.2.9a)

or conversely as a function of \( c_{p_z}^{\text{LPM}} \):

\[
\mu_z^{\text{DFT}} \approx -2 \times 2 u' \left\{1 - \left( 0.869 |\omega|^{-1} c_{p_z}^{\text{LPM}} \right)^2 \right\}^{1/2} \left( 0.869 |\omega|^{-1} c_{p_z}^{\text{LPM}} \right)
\]

by \( c_{p_z}^{\text{DFT}} = 0.869 |\omega|^{-1} c_{p_z}^{\text{LPM}} \).

Note that instead of equality an approximate equality sign is used in 3.2.9b, as the formula is only valid when the HOMO, or the valence state of the DFT consists mainly of \( s \) and \( p_z \) states with negligibly small \( d \) or higher angular moment function contributions; the multiplier -2 in front of each leading term represents there are two electrons forming the lone pair; and the \( u' \) is the DFT dipole transition moment.

To calculate the dipole moments, \( \mu_z^{\text{LPM}} \) and \( \mu_z^{\text{DFT}} \), the main lack were the transition moments, \( u \) and \( u' \). In practice, it is straightforward to get the LPM transition moment \( u \), and for the DFT \( u' \), it requires a numerical method to obtain the value. Therefore, in figure 3.4, we plot the ratio of (dipole moment = d.m.) / (transition moments = t.m.) as a function of dipole moments using LPM and DFT\(^{34}\). Then the values of the transition moment for the \( \text{Pb}^{2+} \) cation were obtained for each model by taking the reciprocal of the slopes, given in the figure, which are \( u = 0.7350 \) and \( u' = 0.8865 \) in Å unit.

\(^{34}\) The dipole moments in the figure has been already probed when the interaction of a single \( \text{Pb}^{2+} \) cation and a point charge was investigated in the right previous section.
With those obtained transition moments, the LPM and DFT dipole moments as a function of $c_{p_z}^{\text{DFT}}$ are plotted in figure 3.5. In the figure, plots, using differing candidates of LPM effective charge, are compared with the DFT data. According to the comparison, the LPM charge of $\omega = -1.424$ shows the best agreement with the DFT dipole moment in the range of $c_{p_z}^{\text{DFT}} \in [0, 0.1]$, which covers the practical values of $c_{p_z}^{\text{DFT}}$ according to figure 3.3 (a) – which was written as $6p_{p_z}^{\text{DFT}}$.

Recall that the employed parameter set of LPM, $\{\omega, \lambda\} = \{-1.424, 16.53\}$, taken from table 3.2 for the Pb$^{2+}$ cation (see section 3.1.3), shows the least error with respect to DFT, and thus such agreement is not really surprising.

Note that the DFT plot in figure 3.5 is an approximation (or scaled), calculated using equation 3.2.9a, the approximation is reliable only for modelling the polarisation of Pb$^{2+}$ cation when $c_{p_z}^{\text{DFT}}$ is not too great. For instance, a reliable upper limit for $c_{p_z}^{\text{DFT}}$ would be 0.10, considering the $c_{p_z}^{\text{DFT}}$ graph in figure 3.3 (a) where $c_{p_z}^{\text{DFT}}$ reaches its highest value of 0.113 at the separation distance of 2.0 Å. For the validation of the upper limit, we have compared the dipole moments, calculated by
using equation (a) 3.2.9a; (b) equation 3.2.8; and (c) the dipole moments directly obtained from the DFT calculations where the comparison results are summarised in table 3.3.

According to table 3.3, for the first two smaller values of $c_{pz}^{DFT}$ (0.049 and 0.070) the error in the calculated dipole moment using both DFT based approximation (b) and LPM (c) is less than 2.0%. However, as the value of $c_{pz}^{DFT}$ rises above 0.10, the error becomes significant, for instance with the greatest value of $c_{pz}^{DFT} = 0.157$, the error of the DFT based approximation is near -8.9% and the corresponding LPM result is even worse as -13.9%. Such a discrepancy is, however, expected, since, the value of $c_{pz}^{DFT} = 0.157$ is from the Pb$^{2+}$ cation interacting with the point charge of $q = -1$ at the distance of 1.50 Å, which is below the left boundary in figure 3.3 (a). In the circumstance, the underestimated dipole moment calculated by the approximate methods are clearly related to the missed contribution from $d_{z^2}$, which is absent in the approximation forms.

---

**Figure 3.5** Dipole moments (d.m.) as function of the molecular orbital coefficient $c_{pz}^{DFT}$ where its scaling to the LPM counterpart is provided in the parentheses. The red line shows the results of a DFT approximation given in equation 3.2.9a, the black and green lines show LPM results based on equation 3.2.8 with different charge values of $\omega$, distinguished by different line styles.
Our investigation of a single Pb\(^{2+}\) cation interacting with a unit charge has shown the validity of the minimal basis set employed in the LPM – one s and three p atomic orbitals are practically sufficient to describe the effective valence state of electron lone pair or its polarisation – since the inner d orbitals of the cation are relatively inert, which was demonstrated by analysing the HOMO of the Pb\(^{2+}\) cation, calculated by using DFT level of theory. Comparative analysis of the orbital occupancy in LPM and DFT allowed us to find a scaling equation (see equation 3.2.1), between the p orbitals of the LPM and DFT, responsible for the \(sp\)-hybridisation. Furthermore, a simple analysis of a two-electron system by using the Hartree approximation, we have shown that how the LPM could implicitly model electron-electron repulsion.

Particularly, by using the the scaling equation, we have shown that the LPM could be potentially used to predict or approximate the DFT dipole moment for an isolated lone pair cation, interacting with a point charge, which was verified that the approximation shows errors less than 3.0\% (see the first two cases in table 3.3). The demonstration was carried out by using known DFT answers to validate the derived relationship (equations 3.2.8 and 3.2.9a); therefore, there is a limitation of scaling model for predicting the DFT dipole moment of which the method relies on a few

<table>
<thead>
<tr>
<th></th>
<th>(a) d.m. / (c_{p_z}^{DFT})</th>
<th>(b) d.m. (err (%))</th>
<th>(c) d.m. (err (%))</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFT-PBEsol</td>
<td>0.169 / 0.049</td>
<td>0.172 (1.8)</td>
<td>0.166 (-1.8)</td>
</tr>
<tr>
<td></td>
<td>0.245 / 0.070</td>
<td>0.248 (1.2)</td>
<td>0.239 (-1.4)</td>
</tr>
<tr>
<td></td>
<td>0.408 / 0.113</td>
<td>0.398 (-2.5)</td>
<td>0.381 (-6.6)</td>
</tr>
<tr>
<td></td>
<td>0.604 / 0.157</td>
<td>0.550 (-8.9)</td>
<td>0.520 (-13.9)</td>
</tr>
</tbody>
</table>

Table 3.3 Dipole moment (d.m.) (e Å), of Pb\(^{2+}\) cation interacting with a point charge at representative distances. (a) DFT values along with the \(p_z\) (or \(c_{p_z}^{DFT}\)) occupancy of the HOMO, (b) and (c) show approximate values calculated using respective equations and error percentages relative to the DFT values in column (a). Note that the effective charge \(\omega = -1.424\) was used to calculate values reported in (c).

3.2.3 Summary
parameters must be prepared in advance: (i) the constant term of 0.869 in equation 3.2.1 and (ii) the transition moments, \(u'\) and \(u''\), for both LPM and DFT where their derivations were demonstrated in figure 3.4.

### 3.3 The effective potentials in the lone pair model

In chapter 2, we have introduced two types of effective potentials employed in the LPM to reproduce the size effect of a lone pair interacting with other species. The first type is for non-lone-pair species, which are described in the level of molecular mechanics, and the second type is for external lone pairs (i.e., other lone pair cations). These two effective potentials can be found in equations 2.4.4 and 2.4.5, respectively, and they require species dependent parameters. Due to the requirement for those parameters, it is worthwhile to investigate the properties of those two potentials before the start of studying more complicated systems, including many numbers of lone pair cations and anions in chapter 4.

#### 3.3.1 Pb(II) cation interacting with a point charge

As the first start of the investigation, the simplest system would be again a single lone pair cation interacting with a point charge – this time we focus on the model energetics – which would be the instructive basis to find out the effect of the LPM effective potential. In figure 3.6, the plot shows the energy profiles of a Pb\(^{2+}\) cation interacting with a point charge \((q = -1)\) as a function of separation distance, calculated using LPM\(^{35}\) and DFT.

\(^{35}\) The LPM is modelling a Pb\(^{2+}\) cation; therefore the parameters, \(\{\lambda, \omega\} = \{16.53, -1.424\}\), were employed (see table 3.2); and yet there is no LPM effective potential used between the electron lone pair and the point charge.
**Electrostatic interactions in LPM interacting with a point charge.** Here, all LPM interactions are purely electrostatic, i.e., no effective potential between the lone pair and the point charge is yet used. In the plot, the energy profile of the LPM is decomposed into the electron lone pair and the core contributions, \( E_{LP} \) and \( E_{core} \), respectively. Note that \( E_{core} \) is simply a Coulomb potential energy for the core point with charge \( Z \) interacting with the point charge \( q \). Furthermore, the lone pair and the core charges sum to the formal charge of the lead cation, \((Z + \omega) = +2\), i.e., \( Z = +3.424 \) since \( \omega = -1.424 \); hence, \( E_{LP} \) and \( E_{core} \) show positive and negative contributions in the plot, respectively.

In figure 3.6, the energy of DFT (\( E_{DFT} \), with red cross markers) and LPM (\( E_{LP} + E_{core} \), with green solid line), show good agreement in the practical distance range above 2.0 Å.\(^{36}\) In this region,

![Graph](image.png)

**Figure 3.6** Energy of a \( \text{Pb}^{2+} \) cation interacting with a point charge \((q = -1)\) as a function of the separation distance, calculated by using the LPM and the DFT-PBEsol (using FHI-AIMS code); \( E_{Coulomb} \) shows the Coulomb energy of two point charges, \( q = +2 \) (formal charge of \( \text{Pb}^{2+} \) cation) and \( q = -1 \). Note that the total LPM energy (green line) is decomposed into, \( E_{LP} \) and \( E_{core} \), which are energy contributions by the electron lone pair and the core part of the \( \text{Pb}^{2+} \) cation, respectively.

\(^{36}\) The deviation below the distance is due to the effect of \( d \) states, which is absent in the LPM. However, such distance range is unlikely observed, and the relevant discussion can be found in section 3.2.
especially at the right end, their energy profiles also match $E_{\text{Coulomb}}$, which confirms that the LPM energy shows correct asymptotically behaviour in that the polarisation energy of the lead cation decays with the distance to the point charge much faster than the leading Coulomb energy.

At the other extreme is the distance region close to 1.0 Å where both energies of the LPM and DFT diverge to negative infinity as the point charge gets closer to the lone pair cation. This result exhibited by both methods is natural as the electrostatic interaction between the positively charged Pb nucleus and a negative point charge will dominate over other terms. Even though both energies are qualitatively in agreement in this region – separation distance below 2.0 Å – where the LPM energy starts visibly deviating out from that of DFT. In this region, the DFT energy also shows more stabilisation, which is explained by the additional polarisation that involves the higher energy $d$ states on the Pb$^{2+}$ cation, i.e., $d$ states hybridise with $s$ and $p$ states as discussed in section 3.2. In contrast, the LPM basis set does not include atomic $d$ orbitals, and, therefore, such additional energy stabilisation cannot be observed.

However, the key separation distances are in the range of 2.0 Å – 2.5 Å where the chemical bonds are formed between Pb$^{2+}$ and typical anions, e.g., O$^{2-}$ and F$, which we modelled in chapter 4. In this distance region, the energetics of LPM, $E_{\text{LP}} + E_{\text{core}}$, overlaps well with that of DFT, $E_{\text{DFT}}$, since the effect of $d$ state becomes negligible – discussed in section 3.2 – and the lone pair polarisation or the $sp$-hybridisation become dominant, evidenced by the both energies of LPM and DFT are lower than $E_{\text{Coulomb}}$. In detail, at the distance of 2.0 Å, the LPM energy differs from the DFT by 0.21 eV and from $E_{\text{Coulomb}}$ by 0.90 eV. The smaller energy difference of 0.21 eV between the LPM and the DFT could be modelled by a classical short-range potential, which would describe the interaction between the core part of LPM and the point charge, including the polarisation of the core, and provide additional energy stabilisation. However, the current version of LPM treats
the core part as a rigid-ion, which cannot be polarised; and such small energy correction is neglected in order to keep this prototype model as simple as possible.

**The role of LPM effective potential.** Figure 3.7 demonstrates the effect of the LPM effective potential on the lone pair localised on a \( \text{Pb}^{2+} \) cation interacting with a point charge \( q = -1 \), including the lone pair energy in panel (a) and the dipole moment in panel (b).

As seen from the plots in panel (a), the employment of the LPM effective potential does not result in markable changes in the interaction energy, especially in the region of interest, above 2.0 Å, it improves, however, the dipole moment as evidenced by the plots in panel (b) where the plots of the LPM effective potential parameters, \( D = 2500 \) and \( 5000 \) eV, closely bracket the dipole moments of the DFT. Here, the repulsive character of the model effective potential results in additional polarisation of the lone pair, i.e., the lone pair is pushed away from the approaching point charge; and thus the dipole moments of the model \( \text{Pb}^{2+} \) cation is increased. Furthermore, even though the energy plots of the different effective potentials visibly merge at 2.0 Å in panel (a), the energy by the parameter of \( D = 5000 \) eV is 0.09 eV higher than that of \( D = 0 \) eV, due to the repulsive nature of the potential.

The key point we have learnt from this investigation is that the LPM effective potential by itself alone may not be sufficient – considering there is no repulsive interaction between the core part of \( \text{Pb}^{2+} \) cation and the point charge – to counterbalance the negative attractive interaction between the LPM core and the point charge (compare the value of \( E_{\text{LP}} \), in figure 3.7 (a); and \( E_{\text{core}} \), which is attractive, in figure 3.6), the overall potential remains attractive throughout the range of separation distances, and on standard energy minimisation the point charge would collapse onto the LPM core.
On the other hand, at the short separation distances in figure 3.7 the repulsive behaviour of the lone pair unexpectedly stabilises, which would imply a few features of the LPM. Foremost, for distances below 0.6 Å, as can be seen from the plots in panel (b), the dipoles are dropped suddenly to zero, and this is because the lone pair states lose their characters of sp-hybridisation and turned into pure p states. Such transition can be understood as term crossing upon which by adopting a p-shaped states that the lone pair exposes an electronically deficient region around the node of its

Figure 3.7 Effect of the effective potential on observed properties of a lone pair. (a) Energy of the Pb(II) lone pair interacting with a point charge \( q = -1 \) as a function of separation distance. The graphs were obtained by using different parameter \( D \) (eV) of the LPM effective potential, and they are shown with different line styles; note that parameter \( \rho = 0.09 \) Å was used throughout. (b) Corresponding dipole moment (d.m.). DFT-PBEsol (using FHI-AIMS code) calculated references are marked with red crosses.
dumbbell-shaped orbital, aligned normally to the approaching point charge, which is energetically more favoured. This behaviour of the LPM lone pair can be verified by tracing the eigenvalues of the lone pair as a function of the separation distance, which are shown in figure 3.8. In the figure, we find a crossing point between terms \( A \) and \( B \) where a transition from the polarised \( sp \)-hybridised lone pair state to the degenerate pair of pure \( p \) states occur where the latter state is without the \( s \) contribution and has a zero dipolar moment. The ground state \( A \) in the figure is clearly separated from the three excited states – two-fold degenerate \( B \) and one \( C \) states – for the distances above 2.0 Å where most interatomic interactions occur, with the energy difference exceeding 14 eV, roughly close to the employed model onsite potential parameter, \( \lambda = 16.53 \) eV.

![Figure 3.8](image)

Figure 3.8 Eigenvalues of the lone pair electron density calculated by the LPM where the ground state (\( sp \)-hybridisation) \( A \) corresponds to the energy of \( D = 10000 \) (eV) in figure 3.7 (a); \( B \) is the eigenvalue for the degenerate \( p \) states, and \( C \) is the anti-state of \( A \).

### 3.3.2 Pb(II) cation interacting with a fluoride anion

Our next target system is a lone pair cation, \( \text{Pb}^{2+} \), interacting with an ion. In the previous section, we demonstrated that the LPM effective potential can be employed to induce an additional
polarisation of the lone pair under interacting with a point charge, evidenced by delicately changing the model dipole moment, see figure 3.7 (b), especially above the distances of 2.0 Å. However, at much closer separation distances, since the point charge does not have its own size, the lone pair cation and the point charge will collapse onto each other if a standard energy minimisation is carried out. Here, to ameliorate the size effect of an LPM cation, we introduced a classical interatomic potential between the core of the cation and external ion as discussed in section 1.3. This choice is justified since the outer core d states remain relatively inert to polarisation, as demonstrated in section 3.2, and also by the consideration of the rest of inner electrons, forming the core part of the model lone pair cation. In this section, moving from a model cation, Pb$^{2+}$, interacting with a point charge to more realistic system, we considered the interaction of Pb$^{2+}$ with F$^-$. 

**The role of rigid ion core of a LPM lone pair cation.** The same Pb$^{2+}$ LPM potential parameters were employed, $\{\lambda, \omega\} = \{16.53, -1.424\}$ along with an LPM effective potential using $\{D, \rho\} = \{5000, 0.09\}$, as employed in the previous section. Here, we modelled a fluoride ion as the rigid ion model; therefore the Buckingham Pb-F potential was employed with parameters of $\{D, \rho, C\} = \{4494, 0.264, 0.\}$ (see equation 1.11)$^{37}$. Note that the LPM effective potential and the interatomic potential have the same decaying exponential form, using the same notation for the pre-exponential and exponent parameters $D$ and $\rho$ but their usage differs by context. In figure 3.9, the energy of the LPM Pb$^{2+}$ cation interacting with a fluoride anion is shown.

$^{37}$ The parameter set $\{D, \rho, C\}$, was taken from the application of LPM in section 4.4 where the details are in Appendix D. The main difference of usage is that the shell model was employed to describe the fluoride ion in the application, however, in this experiment, the shell model is instead replaced with the rigid ion model.
The marked difference with the case of model interaction with a point charge – see figure 3.6 – is that there appears an energy equilibrium (or minimum) point at the distance of 2.03 Å (2.00 Å by DFT) on the green curve, representing $E_{LP} + E_{core}$. The difference is due to the repulsive energy contribution, $E_{core}$, by the interaction between the core of Pb and the F (see the dashed black curve) where its profile sharply rises as the distance decreasing below the equilibrium point. The equilibrium distance is much shorter than what is experimentally observed in PbF$_2$ fluorite structure, 2.57 Å for Pb-F [125], where the lead cation is in a symmetric environment, at the centre of symmetry in a body-centred cubic lattice with coordination number of 8; therefore it does not show a stereoactive lone pair density [14]. As a dimer or a molecule, PbF$^+$, on the other hand, unlike the condensed fluorite environment, there are no constraints for the lone pair to be formed, and thus the ligand, fluoride anion, can approach the lead cation closer, resulting in the lone pair density

![Energy profiles of Pb(II) cation interacting with a fluoride anion as a function of the separation distance, calculated by the DFT-PBEsol (using FHI-AIMS code) and LPM; $E_{Coulomb}$ shows the Coulomb energy of two point charges, $q = +2$ and $q = -1$, i.e., the formal charges of Pb$^{2+}$ and F$^-$. The LPM energy is decomposed into $E_{LP}$ and $E_{core}$, which are the energy contributions by the lone pair and the core part of the Pb$^{2+}$ cation. The difference between $E_{DFT}$ and $E_{C,DFT}$ is that the latter energy was calculated under the constraint of prohibited charge transfer from F$^-$ to Pb$^{2+}$.](image)

**Figure 3.9** Energy profiles of Pb(II) cation interacting with a fluoride anion as a function of the separation distance, calculated by the DFT-PBEsol (using FHI-AIMS code) and LPM; $E_{Coulomb}$ shows the Coulomb energy of two point charges, $q = +2$ and $q = -1$, i.e., the formal charges of Pb$^{2+}$ and F$^-$. The LPM energy is decomposed into $E_{LP}$ and $E_{core}$, which are the energy contributions by the lone pair and the core part of the Pb$^{2+}$ cation. The difference between $E_{DFT}$ and $E_{C,DFT}$ is that the latter energy was calculated under the constraint of prohibited charge transfer from F$^-$ to Pb$^{2+}$. 

119
polarises away from the anion. Here the distance of 2.0 Å is also close to the sum of the Pb\(^{4+}\)(0.77 Å) and F\(^-\)(1.33 Å) radii [131], 2.10 Å, which implies that Pb\(^{4+}\) acts as the least compressible sphere, consisting of a lone pair cation unlike its valence state or lone pair density, which readily changes its shape.

The DFT energy \(E_{\text{DFT}}\), marked as red circles in figure 3.9, shows a minimum at 2.00 Å, which is quite close to the LPM result. However, its asymptotic (above 4.0 Å) clearly diverges from the LPM graph. In contrast, the good matching between \(E_{\text{Coulomb}}\) and the LPM energy, \(E_{\text{LP}} + E_{\text{core}}\), profile at large distances. This implies that there was the interionic charge transfer (see the discussion in section 1.3) in the DFT calculations, from the electron rich to electron poor species, i.e., F\(^-\) → Pb\(^{2+}\), which was also verified by using the Mulliken analysis [132–134].

The deviation of energies between \(E_{\text{DFT}}\) and \(E_{\text{LP}} + E_{\text{core}}\) is particularly significant at large separation distances, and this is because the charge transfer or delocalisation occurs in such a long-range – since DFT treats the molecular, PbF\(^+\), as a single system – is absent in the LPM. To verify the results we, therefore, have used the constrained DFT approach [135–137] that does not allow for charges to transfer between different fragments, i.e., partitioning Pb\(^{2+}\) and F\(^-\), and its energy, \(E_{\text{C.DFT}}\), is plotted in figure 3.9, marked as blue circles. The profile of \(E_{\text{C.DFT}}\) is seen to match with both \(E_{\text{LP}} + E_{\text{core}}\) and \(E_{\text{Coulomb}}\) graphs, asymptotically. Note that the constrained DFT is used here only for the validation of LPM result, but such scheme of DFT calculation was not employed in our further application work, simulating the nanoclusters consisting of multiple number of lone pair cations and anions, reported in chapter 4. The energy difference between \(E_{\text{C.DFT}}\) and \(E_{\text{LP}} + E_{\text{core}}\) near their equilibrium distance could be explained by the rigid ion model used to describe F\(^-\) for the LPM calculations, which does not have the polarisability. For instance, if the shell model is employed instead, the additional polarisation of F\(^-\) – the effect was still in the C.DFT
calculations, since only the electrons of fluoride are confined to the ion centre, but they can still deform – could provide additional energy stabilisation for the LPM calculated dimer by the displacement of F\(^-\) shell to Pb\(^{2+}\) centre, i.e., the model charge transfer occurs; and this is evidenced by the overlapping profiles of \(E_{\text{DFT}}, E_{\text{C,DFT}}\) and \(E_{\text{LP}} + E_{\text{core}}\) at short distances, below 1.7 Å.

**Conclusion.** In figure 3.10, we summarised the behaviour of model lone pair cation, which is interacting with a point charge (A); and a lone pair cation interacting with a fluoride anion (B), having been discussed so far in this section 3.3.

As illustrated in diagram (A), the lone pair electron density prefers to keep itself away from an approaching negatively charged ligand. Furthermore, according to the dipole moment analysis in section 3.3.1, including the model effective repulsive potential between the electron lone pair and the ligand would eventually – this is an extreme case but for instructive model description – leads to a transition of the ground state from an \(sp\)-hybridised to a pure \(p\) character. Consequently, in the total energy \((E_{\text{LPM}} + E_{\text{core}})\) plot, there is no minimum, or an equilibrium point, and at the distance \(x_3\), the transition occurs where the ligand moves into the node area of a \(p\) orbital (only one of the two degenerate orbitals is shown in the diagram whereas the other presumably oriented normal to the plane of the diagram is omitted for clarity). In diagram (B), on the other hand, the short-range Buckingham repulsive potential is employed between the ligand and the core of an LPM cation, which accounts for their mutual size effect or electron-electron interaction. Consequently, on the energy plot, one finds a minimum at equilibrium point \(x_{eq}\), which is formed by the positive energy contribution from the repulsive potential – see black dashed curve\(^{38}\).

\(^{38}\) The transition from \(sp\)-hybridised to a pure \(p\) character will occur as the former case at \(x_3\). The distance, however, more difficult to be reached due to the core repulsive interaction and therefore omitted for clarity.
In conclusion, the repulsive short-range potential, acting between the core part of LPM cation and anion ligands, is essential to model interactions at relevant interionic distances between a lone pair cation and coordinating anions. The repulsive potential accounts for the mutual size effect of...
the core part of the LPM cation, including the core electrons and nucleus, and the ligand, as discussed earlier in section 3.3.2. Here, for its demonstration, the molecular cluster, PbF\(^+\), was chosen where the both LPM and DFT energy functions, include equilibrium points\(^{39}\), see figure 3.9.

### 3.3.3 Pb(II) cation interacting with another Pb(II) cation

The last validation system of interest consists of two lone pair cations. Unlike a typical atomistic description, which uses a point charge to represent valence electrons on an ion, LPM employs a molecular orbital occupied by a lone pair, continuously distributed in space. For two Pb\(^{2+}\) cations investigated in this section, there are acting: (i) the Coulomb interaction between the two cores, (ii) the Coulomb interaction between each lone pair and the core of the other cation, and (iii) the Coulomb and the effective potential (see equation 2.4.7) interactions between the two lone pairs. The effective potential interaction is repulsive that captures the size effect or their overlaps for a pair of lone pair densities at close distances. Usually in atomistic models, such short-range cation-cation interatomic potentials are omitted [138], however, allowing for much larger effective radii of lone pair cations, we have added the corresponding effective potentials in our implementation.

**Energetics of two interacting Pb(II) cations.** For the chosen pair of Pb\(^{2+}\) cations we employed the same LPM parameter set, \(\{\lambda, \omega\} = \{16.53, -1.424\}\), that has been used in the two previous sections so far, 3.3.1 and 3.3.2, which was tuned to reproduce the energy and dipole moment of an isolated Pb\(^{2+}\) cation under electric field, calculated by using DFT in section 3.1.3. Furthermore,

\(^{39}\) It was not presented in this main text, however, even for the case of neutral molecular cluster, PbO, the repulsive short-range potential was required for the LPM to reproduce an equilibrium point or interionic distance, observed in the DFT calculation of PbO.
the system behaviour has been investigated with and without the effective potential acting between the lone pair densities where the results are summarised in figure 3.11.

As seen from the figure, the LPM energetics with using the effective potential (black curve) shows better match with the DFT reference (red circles) than the energetics with using the Coulomb potential only (green curve); although the green curve starts deviating out from both the black curve and red circles below the distance of 2.8 Å. In fact, such distances are unlikely to be observed between Pb²⁺ cations, for instance, even in the dense crystalline environment, the shortest Pb²⁺-Pb²⁺ distances in α-PbO and fluorite PbF₂ are 3.847 Å and 4.194 Å, respectively. According to the insert in figure 3.11, showing the distance range of 3.5-4.5 Å, the both LPM
calculated energies with green and black curves practically overlap and are close to the red circles, calculated by DFT.

Considering the typical cation-cation interionic distances, such result implies that it might not be necessary to employ the effective potential for the lone pair – lone pair interaction. However, the potential could become useful in stochastic simulations, e.g., structure prediction of nanoclusters including the lone pair cations. In detail, a randomly generated initial configuration might include nearest neighbouring lone pair cations, in which case reproducing the energy landscape similar to DFT will improve the ability of the LPM to capture correct atomic structures.

At short distances, the Coulomb energy of two interacting point charges, represented by the dashed purple line, deviates from both DFT and LPM energy graphs. In contrast at large distances, the energy profiles of Coulomb, LPM and DFT converge showing the same asymptotic behaviour, i.e., in the far separation regime, it could be interpreted the lead cations behave as point charges. The Coulomb energy also proves to be the $\lambda \to \infty$ limit of the LPM, as shown by the purple circles overlapping the purple line (in practice, a large value of $\lambda = 99999$ eV was used). In this limit, the lead cations remain in pure $s$ states, as the $s \to p$ transitions become impossible, i.e., no $sp$-hybridisation or polarisation could be realised, which also implies that the energy difference between the Coulomb interaction of two point charges and the DFT is due to polarisation of the lead cations, realised by the DFT calculation.

**Formation of lone pair electron density and energy stabilisation.** To verify the role of the polarisation effect we have considered the ratios of DFT and LPM energies to the Coulomb energy – plotted in figure 3.11 – and their ratios are summarised in the next figure, 3.12. In the figure, for
In comparison, we have added the corresponding ratio graph for a Pb$^{4+}$-Pb$^{4+}$ dimer (shown as blue circles), calculated by using DFT.

As should be expected, the energy ratio for Pb$^{2+}$ is appreciably lower than for Pb$^{4+}$, which is much less polarisable, evidenced by the DFT calculated results in figure 3.11, i.e., if the value of a ratio is equal to 1, this means the pair of cations are not polarised at the distance but act like point charges or vice-versa. On the other hand, although the both DFT and LPM values for Pb$^{2+}$-Pb$^{2+}$ are closer to each other, nevertheless there is of about 0.01 point difference between them in the given range of distances, that are of main interest, which implies that there is an additional polarisation in the DFT description. Indeed, the LPM modelled Pb$^{2+}$ cations faithfully reproduce the sp-hybridisations, and the examination of the both LPM and DFT eigenvectors shows that the geometric orientations of lone pairs are facing each other due to the net positive charge of Pb$^{2+}$.
cations. To check possible reasons for such discrepancy, we have analysed the DFT valence state and found that unlike the case of isolated single Pb$^{2+}$, discussed in section 3.2.1, the lead cations of Pb$^{2+}$-Pb$^{2+}$ dimer show admixture of the core 3s states to the valence 6s and 6p states (3s is a part of the default basis set employed in the FHI-AIMS code for Pb atom), which may cause the difference with the LPM result. Recall that the radius of the 3s atomic orbital is in lower energy state and much smaller than that of 6s, if there were additional such s orbital introduced to the LPM, the lone pair densities of Pb$^{2+}$-Pb$^{2+}$ could lower the energy by additional polarisation. Furthermore, the admixture of smaller 3s orbital would make the LPM represented cation become compressible by changing respective contributions of two s orbitals of different size adaptively to the environment. Such functionality is similar to that of the breathing shell and compressible ion models [139], but the LPM yet does not have such capability, and it would be an interesting topic for the future research.
Chapter 4

Application of the New Model

The lone pair ion model (LPM) has been developed. The theory of the LPM was provided in the previous chapters, which included the mathematical descriptions of the energy and physical properties obtained from the derivatives – high order electric moments, response under electric fields, and the influence of LPM employed parameters. Here the LPM will be applied to modelling nanoclusters. The measure of success will be judged on how well the match is between the local minima atomic configurations on the LPM and the DFT-PBEsol\textsuperscript{40} energy landscapes, where the latter are computed using full electronic structure theory using the FHI-AIMS code. With an energy expression that is computationally cheaper to evaluate, then investigations requiring many configurations to be assessed becomes more feasible. Thus, in this chapter, structure prediction results for nanoclusters of different stoichiometry, lead monoxide (PbO) and lead fluoride (PbF\textsubscript{2}), are presented. We begin with a brief description of the global optimisation scheme used in

\textsuperscript{40} The PBEsol DFT exchange-correlation functional was employed primarily to refine the LPM obtained nanocluster configurations, in order to keep the consistency of its use throughout this thesis; however, we also have carried out some benchmark calculations that allowed us to compare the main PBEsol results those obtained using different functionals: B3LYP and PBEsol0, which will be discussed later in sections 4.4.2 – 3.
predicting the atomic structure of local minima (LM) nanoclusters. Then, the results of the LPM LM are compared with DFT LM, including a closer look at the structural distortions caused by the presence of the lone pair electrons.

4.1 Structure prediction of nanoclusters

An important concept in nanocluster structure prediction is assumption that the atomic structures adopted in nature corresponds to the atomic configurations of the local minima (LM) on the potential energy landscape. Assuming the energy of a nanocluster can be written as a function of atomic coordinates (see equation 1.1), then the energy landscape can be mapped out as a function of atomic coordinates. The dimensionality of the potential energy landscape (or hypersurface) matches the total number of degrees of freedom for the configuration of interest. For a non-linear cluster composed of $N$ atoms, this is $3N - 6$ as rotational and translational motions do not change the potential energy. A schematic diagram of a one-dimensional energy landscape is shown in figure 4.1. This landscape is already quite complicated, with many LM to find (at least the lower energy ones). The number of LM and the complexity of the landscape is expected to increase with an increase in dimensionality of the landscape: both visualising higher dimensional landscapes as well as searching these for low energy LM becomes extremely difficult and the problem of minimising the computational cost of finding low energy LM includes minimising the number of sample points to evaluate before finding the LM, knowing when the search is complete, and minimising the computational cost of each evaluation.

At the bottom of each basin there is an energy minimum corresponding to a metastable nanocluster, i.e. an atomic configuration with no resultant atomic forces (see equation 1.2). The landscape shown in figure 4.1, can also be separated into two regions (A and B), typically referred to as
super-basins. The lowest energy, or global minimum (GM), which is within Region A, corresponds to the ground state configuration of the nanocluster and the configurations of the nearby LM within region A likely to be of importance when the effect of temperature is considered. In this respect, it is important to not only find the GM but also all other low energy LM. An indication of stability can also be gained from the curvature of the bottom of basin the LM occupies, so we will also calculate vibrational frequencies, which may be compared to future spectroscopy measurements.

Finding the GM and other low energy LM becomes increasing difficult with cluster size (number of atoms). As already stated, the complexity in energy landscapes will increase, likewise the number of LM will typically exponentially increase with the number of atoms/ions. Whereas local optimisation schemes are efficient at finding a LM, once a nanocluster structure is determined by energy minimisation, it is not straightforward to guide the search to escape the current basin with the aim of finding LM in region A, as opposed to those in region B. Also, how to avoid falling back into the same energy basin or revisiting energy basins already found.

There has been a plethora of global optimisation algorithms that have been designed and applied to predicting low energy structures. Monte Carlo basin hopping [140,141], evolutionary algorithms [142–144], simulated annealing [145,146] and stochastic quenching [147] have been successfully
applied to efficiently explore energy landscapes for energetically low-lying nanocluster structures. One common approach is to employ the global optimisation scheme to locate LM on the energy landscape defined using interatomic potentials and then, as a second stage, refine the better configurations using an electronic structure approach [148–152].

In this study, we employed the stochastic quenching method, which typically has a workflow of optimising randomised structures first using an interatomic potential approach and the low-energy candidates are finally refined using an electronic structure method. Here, we added an additional step, instead of employing LPM directly characterising LM from randomised structures, those structures were initially relaxed using the rigid-ion model (RIM), and LPM was employed for subsequent refinement before the final step using an electronic structure method. The insertion of initial RIM relaxations was made due to the instability of LPM that frequently fails when randomised structures are in non-sensible configurations, e.g., including too close interatomic distance. Consequently, the RIM step acts as a pre-screener to increase the stability of LPM refinement process, and overall, this modification was effective to reduce the computational cost locating LM on the LPM energy landscape. More details will be delivered in section 4.2.3.

4.2 Methodology

To determine nanocluster structures, the most essential two elements are the energy of a nanocluster and geometric derivatives of the energy. The geometric derivatives are the least requirement to carry out standard energy minimisation to characterise an energetically stable structure, i.e., LM. Starting with a discussion about those two key elements, we will deliver how LPM relaxes the energy of nanoclusters using a local minimisation algorithm; and its association with the global optimisation scheme, which is for investigating the energy landscape of LPM. In this study, to
carry out LPM simulations, we have implemented a software package, *Sp-Lone pair integrated Atomistic Model* (SLAM), which is used throughout, where the details can be found in Appendix E.

### 4.2.1 Energy and Derivatives of the LPM

In the LPM, we recall that a lone pair cation is treated as the core (representing the nucleus and core electrons), which is treated like a rigid-ion, and the lone pair electron density of the molecular orbital. In chapter 2, we established an expression for $E_{LP}$, the electronic energy for lone pair electron densities with its environment (electric field generated from the surrounding ions). For nanoclusters of PbO and PbF$_2$, interactions involving the lone pair densities include: (i) that between lone pair densities of Pb cations; (ii) that between the lone pair densities of Pb cations and core parts of Pb cations; and (iii) that between the lone pair densities of Pb cations and the oxygen or fluoride anions. The total potential energy,

$$E = E_{MM} + E_{LP},$$

also includes $E_{MM}$, the energy contribution from the two-body Buckingham and Coulomb interactions between all anions, and the interactions between the anions and the core part of lone pair cations (see equation 1.11). The anions are modelled using the shell model and as aforementioned the core part of lone pair cations are modelled using a rigid ion. The Coulomb-Buckingham potential is used to replicate the electrostatic interactions and the size effect of the species. The shell model is used to enable the polarisation of the anions. In contrast to the application of the shell model, the Buckingham potential is always applied between the core parts of lone pair cations and anions. The rationale is that this interaction represents the size effect of a lone pair core by fully occupied $d^{10}$ and the core electrons of the lone pair cations; see section 3.3.2 for more details.
4.2.2 Local optimisation

With mathematical expressions for the energy of a nanocluster and its derivatives with respect to atomic displacements, deterministic based algorithms can be employed, which are far more efficient than a Monte Carlo stochastic approach, for relaxing the atomic structure to obtain an energy local minimum (LM), i.e. a metastable configuration of atoms. The relevant theories for energy and derivatives of LPM can be found in section 2.4 and 2.5.

Algorithm III. The backtracking line-search (LS) method implemented in SLAM, which searches for a local minimum of $E(\mathbf{x})$ via a sequence of $k$ line searches, where the $3N$ coordinates after each line search is $\mathbf{x}_k$.

| initialise | $\rho \in [0, 1/2]$ and $\beta \in (0,1)$ | (define internal parameters) |
| $\mathbf{x}_{k=0}$ | (load initial atomic coordinates) |
| $E(\mathbf{x}_{k=0}), \mathbf{p} = -\nabla E(\mathbf{x}_{k=0})$ | (evaluate initial energy/direction) |

outer loop while $|\mathbf{p}|/(3N) > \tau$ (keep going until forces $<$ tolerance $\tau$)

| inner loop while $E(\mathbf{x}_k + \alpha \mathbf{p}) > E(\mathbf{x}_k) - \rho \alpha |\mathbf{p}|^2$ (LS subproblem) |
| $\alpha = \beta \alpha$ | (reduce step size $\alpha$) |
| $k = k + 1$ | (outside inner loop) |
| $\mathbf{x}_k = \mathbf{x}_{k-1} + \alpha \mathbf{p}$ | (update configuration) |
| $\mathbf{p} = -\nabla E(\mathbf{x}_k)$ | (update direction) |

end

return $\mathbf{x}_k, E(\mathbf{x}_k)$ (final values after exit of outer loop)

There are several algorithms based on evaluating the energy, gradient and/or higher derivatives, which are routinely used in the field of computational chemistry to relax atomic configurations to a local energy minimum. For example, the Newton-Raphson, BFGS, Conjugate Gradient and Line Search methods [153]. We employed the backtracking line-search method [154], which is one of the simplest algorithms that requires only the energy and its first derivatives. The backtracking algorithm, described in Algorithm III, was implemented in the model software, SLAM, mentioned
earlier. The algorithm requires the functions $E$ and $\nabla E$, two parameters, $\rho$ and $\beta$, which we set to 0.5, 0.84, respectively, and the initial atomic configuration, $x_0$. The algorithm starts by evaluating the energy of a nanocluster, $E(x_0)$, and the atomic forces, $p = -\nabla E$. $p$ is also the steepest direction on the energy landscape and it is along this direction that the line search is performed. If the average component of the atomic forces is not less than a set tolerance, $\tau$, then it begins a search for a local minimum (LM) along the direction of $p$. The step size $\alpha p$, which is controlled by the parameter $\alpha$, needs to be sufficiently small to ensure not to exit the basin containing the local minimum (LM) along the chosen direction $p$, but not too small otherwise progress downhill to LM will be too slow. The initial step implemented for each line search is therefore based on the current local gradient, i.e., $\Delta x = \alpha p$. If this results in a significant downhill step, i.e., the change in the total energy drops more than $\rho \alpha |p|^2$ then the step is accepted and $p$ is calculated for the new configuration, which is labelled as $k + 1$, and the condition for the outer loop is checked to see if another step is required along the new direction. If the energy drop is less than required, then it is assumed that the step is too large and has effectively step over the LM and the trial point is further up the other side of the energy basin. Thus, the step size is reduced by a factor of $\beta$ and the new trial point is tested but with the target drop adjusted to $\rho (\beta \alpha) |p|^2$ (a smaller drop is expected for a smaller step size, $\Delta x = \beta \alpha p$). This inner loop continues, with the step size reduced each time, until an appropriate step, $\alpha p$, is found.\footnote{The condition used for the inner loop is based on linear approximation of $E(x_k + \alpha p)$ by the Taylor expansion about the current point along the direction $p$, where $E(x_k + \alpha p) \approx E(x_k) + \rho \alpha |p|^2$. With smaller step, the approximation is expected to become more accurate, and once the agreement is within the chosen tolerance then the step is accepted.} Furthermore, the inner loop is repeated until the average atomic forces become less than set tolerance, $\tau = 5. \times 10^{-5}$ (eV/Å) – used in this study.
4.2.3 Global optimisation

A simple global optimisation scheme was employed in this study: a sequence of local optimisations applied to random atomic configurations. The chosen global optimisation method does not require any technical algorithms to be developed and implemented, e.g., for performing a mutation or a crossover in an evolutionary algorithm and, therefore, was straightforward to implement the required interface for the software packages that implement LPM and DFT. The employed sequence of local optimisations is summarised in figure 4.2, which consists of three steps: (1) sample preparation of refining the random configuration to a rigid-ion model (RIM) local minimum (LM), (2) refining the RIM LM to a LPM LM, and (3) refining the better LPM LM ranked nanoclusters to a DFT LM.

1) Sample preparation

In the first step, a finite number of randomised atomic configurations of \((\text{PbO})_N\) and \((\text{PbF}_2)_N\) for a fixed cluster size of \(N\) formula units, where the ions are represented using the RIM. The randomised configurations of rigid ions were generated using the Monte Carlo method [31,155] and relaxed until the average force on ions becomes less than \(5 \times 10^{-5}\) eV/Å. The employed Monte Carlo method is unbiased, based on sampling random coordinates, i.e., the effect of temperature was not included. For this scheme, we used our in-house software, which can be found in the reference [156]. In more detail, all atoms were generated within a predefined cubic cell with volume of \(8 \times 8 \times 8\) Å\(^3\) for smaller sizes \((N \leq 6)\) and \(12 \times 12 \times 12\) Å\(^3\) for the other larger sized clusters; the Buckingham potential was employed between cations and anions, where parameters \(\{D(\text{eV}), \rho(\text{Å}), C(\text{eV}\text{Å}^6)\}\) – see equation 1.4 – for the \(\text{Pb}^{2+}-\text{O}^{2-}\) and \(\text{Pb}^{2+}-\text{F}^{-}\) interactions were \(\{72276.42, 0.2223, 32.32\}\) and \(\{22940, 0.2303, 32.32\}\), respectively.
We found that this step is sometimes required as the LPM is not always robust enough during local refinement, i.e., the LPM can fail during the local optimisation processes when initial randomised configurations do not resemble chemically reasonable structures. For example, when initial inter-atomic distances between cations are too short. A similar problem is found for the shell model and, likewise, a similar three-step optimisation process applied in previous work [157]. If, however, initial nearest neighbour interatomic distances are too long, the computational cost increases as more local optimisation steps are required to locate a local minimum configuration. Therefore, the RIM was used as the initial pre-screener, and any duplicate structure, as determined by their total potential energy, were removed.

**Figure 4.2** (a) The workflow of optimising random configurations employed to predict nanocluster structures. Stages (1), (2) and (3) are energy minimisations on landscapes defined using three different energy definitions, RIM, LPM, and DFT. (b) Schematic diagram showing random points on energy landscape A moving to the bottom of the black basin it sits within, then, after transferring vertically down onto energy surface B, sliding down to the bottom of the red basin it sits within, which, after transferring vertically down onto energy surface C, sliding down to the bottom of the green basin it sits within. The colours match three energy definitions associated with the workflow: landscape A (black curve) is the energy function of rigid ion model, B (red curve) and C (green curve) are the energies from the LPM and DFT approach. The DFT LM are labelled (i), (ii), (iii) and (iv).
2) LPM refinement

In the second step, all unique RIM LM were refined to a LPM LM with tolerance of averaged atomic forces less than $5 \times 10^{-5}$ eV/Å as in the (1) sample preparation. As will be discussed shortly in section 4.2.4, the LPM potential parameters used in the simulations of $(\text{PbO})_{N=1-12}$ and $(\text{PbF}_2)_{N=1-8}$ were tuned to reproduce the atomic DFT LM configurations of the smaller size clusters, i.e. $(\text{PbO})_{N=1-5}$ and $(\text{PbF}_2)_{N=1-2}$, where the employed LPM parameters can be found in Appendix C, see table C.2 for $(\text{PbO})_N$, and Appendix D, see table D.2 for $(\text{PbF}_2)_N$. The switch between energy landscapes then the refinement from the LM configuration found in the first step is relaxed to the bottom of an energy basin of the LPM as schematically shown in figure 4.2 (b), where the black circles (on black curve) represent sample structures on the RIM energy landscape, and their structures were transformed into a new set of structures after energy relaxations on the LPM energy landscape (red curve).

3) DFT refinement

In the final step, the better candidates (red circles) after the second step, i.e., the lower energy LPM LM were refined on the DFT energy landscape to the DFT LM. Now consider what happens to the four LPM LM configurations indicated as a red circle in figure 4.2(b). The first three cases of (i)-(iii) have a nice one-to-one relationship between the LPM LM and the DFT LM. On the other hand, there are two red circles that relax to DFT LM (iv), indicating that at least one LPM LM is not stable on the DFT energy landscape. The former cases are ideal, which we hope that LPM would achieve, and more detailed analysis method we employed in this study will be discussed.

---

42 In figure 4.2(b), it is described that all LPM LM are transformed into the DFT LM, however, in practice, the lower energy LPM LM were chosen, e.g., (i) and (ii) in the diagram, to the DFT refinement.
shortly in section 4.2.4. As specification of the DFT refinement, the FHI-AIMS code was employed using all-electron method [71]. In the calculations, the package default “Tier-2” basis and the PBEsol exchange-correlation functional [113] were employed with scalar-relativistic effect of each element treated at the ZORA level [158]. As the local minimisation algorithm, the BFGS method was used with the atomic force tolerance less than $1 \times 10^{-4}$ eV/Å. The calculations were carried out using 18 CPU cores, whose specification is “Intel Xeon(R) Gold 6140 CPU @ 2.30GHz”.

### 4.2.4 How we assessed the LPM when used to predict nanoclusters

As we do not have suitable experimental data, our aim is to reproduce DFT LM, i.e., an assessment of the LPM can be achieved by comparing the local minima (LM) nanocluster configurations obtained by the LPM with the equivalent DFT LM. Moreover, it will be useful for the application of global optimisation algorithms used in structure prediction that there is a one-to-one relationship between LPM and DFT LM structures, or at least that each LPM LM resides in the energy basin containing the matching DFT LM; see, for example, the different scenarios described in figure 4.3. Furthermore, even if a set of nanoclusters obtained by the LPM are in good agreement with their DFT refinements in terms of structure (figure 4.3(a) and (b)), it is also desirable that the LM energy ranking (from lowest to highest energy) match as shown in example figure 4.3(a). In this respect, examples shown in figure 4.3(c) and (e) are still useful if the procedure for finding the DFT global minimum (or more generally the lower energy LM) is to only relax the lowest (lower) energy LPM LM.

Plotting DFT LM rank ($y$) as a function of LPM LM rank ($x$) and drawing a line between points will generate a straight line, $y = x$, when there is a perfect match. Small changes in LM ranking
(small noise in the graph) is anticipated if the energy differences between the LM is small (this also applies to comparing DFT LM from calculations using a different basis set or functional). If the criteria for a good match between energy landscapes is matching LM atomic structures and matching LM energy ranking, then for the latter it is probably better to plot the DFT LM energies as a function of the LPM LM energies to assess the quality of the LPM. For the former, we need a quantitative measure. It is also important to be able to measure the improvement the LPM provides when compared to what previously would typically be employed, i.e. the shell model.

![Figure 4.3](image)

**Figure 4.3** Example matches between two energy landscapes. The arrows show the results of relaxing the local minima (LM) structures obtained from the red energy landscape to a LM on the green energy landscape. If the criteria for a good match between energy landscapes is matching LM atomic structures and matching LM energy ranking, then: (a) is the most ideal for both conditions; (b) structures match, but incorrect ranking of LM energies; (c) LM ranking correct, but metastable structure heavily distorted; (d) LM ranking incorrect and metastable structure heavily distorted; (e) both LM heavily distorted but energy ranking correct. There is still the question as to whether the heavily distorted LM is a match to the initial LM structure.
1) Quantitative measure of the differences between two atomic structures

The aim is to develop (and fit the parameters of) the LPM so that the LPM LM atomic configurations match the DFT LM atomic structures. A better match will ensure a reduced computational cost to refine the LPM structures on the DFT energy landscape to find the DFT LM, i.e., fewer steps required (fewer DFT energy evaluations) on the DFT energy landscape during the structural relaxation. Thus, we have written software to calculate a similarity index (SI), which provides a measure of how different two atomic configurations are (after translated so that their respective centre of mass is at the origin and axes of their principal moments of inertia aligned by rotation), with a score of zero for a perfect match. The SI is a simple root-mean-square-deviation of two different atomic configurations. The details of the implementation can be found in Appendix B. In practice, the target is to have an SI value of below 0.05 in Å unit. For brevity, if it is not particularly necessary, the unit will be not specified in the text throughout.

2) Comparison with shell model

To get a better idea of how well (or not) the LPM is at simulating (PbO)_N and (PbF_2)_N nanoclusters, we also employed the shell model [32] to describe the lone pair cations. More precisely, the same global optimisation scheme, as described in figure 4.2, was used but with the LPM replaced with the shell model in stage 2.

For the results from the shell model that are presented in this thesis, we used the potential parameters shown in Appendix C, see table C.1. Both the lead (II) cations and oxygen anions are represented using the shell model. The parameters for the Buckingham interatomic potentials (IP) for Pb^{2+}-O^{2-} and O^{2-}-O^{2-} as well as the shell model are taken from reference [54], which were refined by Pirovano et al to reproduce experimentally observed lattice parameters for α-PbO [159].
Likewise, both the lead (II) cations and fluoride anions are represented using the shell model. No suitable IP were found for modelling PbF$_2$ and so we decided to use the shell model and Buckingham IP parameters for F$^-$-F$^-$ refined by Valerio et al, which is taken from reference [160], but we still needed to develop an IP for Pb$^{2+}$-F$^-$ interactions. While keeping the shell model for Pb$^{2+}$ to investigate transferability as used for the shell model PbO, the required IP parameters were refined to reproduce experimentally observed lattice parameters and static dielectric constants of PbF$_2$ fluorite (experimental data taken from [161]) by using the GULP software [162] where the IP parameters can be found in Appendix D, see table D.1. Note that the calculation using the shell model were carried out by using the GULP software package in this study.

4.3 Nanoclusters of lead monoxide

LPM Parameters: Before conducting a search for low energy LM configurations for (PbO)$_N$, the parameters for the LPM were refined to reproduce the atomic structures and energy rankings of the DFT LM for (PbO)$_{N=1-5}$ clusters. The LPM parameters are provided in Appendix C (see table C.2). Note that we chose to constrain the IP parameters for anion – anion interactions and shell model parameters for the anions to be the same in both the LPM calculations and the Shell Model (SM) calculations. Thus, differences between LPM and SM results will not be dependent upon how the anions were treated.

4.3.1 Structures of (PbO)$_N$ clusters (N = 6 - 12)

In tables 4.1 and 4.2, we present the lower energy, DFT LM structures for nanoclusters composed of $N$ formula units of PbO. The energies shown, $E_p (N)$, for the ground states of (PbO)$_N$ nanoclusters are all relative to $N$ infinitely separated $N=1$ ground state cluster, i.e.
\[ E_p(N) = \frac{[E(N) - N \times E(1)]}{N}, \]  
(4.3.1)

where \( E(N) \) and \( E(1) \) are the total energies of the nanocluster shown and the nanocluster (linear stick) of ground state for \((\text{PbO})_1\), respectively. For each LM configuration we also calculated the vibrational frequencies to check that each corresponds to a true minimum (local minimisation schemes can erroneously lead to a saddle point structure). The relative energies are reported with respect to the ground state of each size, \( E_p^{GM}(N) \), when employing DFT and the LPM (values in parentheses). For each LM configuration shown, we also include the value of similarity index (SI), which provides a quantitative measure of how much the atomic configurations of the LPM LM distorted during the relaxation to the DFT LM.

Generally, the LM configurations of \((\text{PbO})_N\) clusters are more amorphous in comparison with alkaline earth oxide clusters, \((\text{BaO})_N\) and \((\text{SrO})_N\), which do not contain any lone pair electrons on those cations and whose configurations resemble cuts from the rock salt lattice structure [163,164]. Consider \((\text{PbO})_6\), for example, the ground state atomic structure has a shape that resembles an hexagonal prism, whereas the LM cuboid formed of \(3 \times 2 \times 2\) atoms, is ranked fourth for \((\text{PbO})_6\) and first for \((\text{SrO})_6\) and \((\text{BaO})_6\). This difference in structural motif extends to larger sized clusters: ground states for these alkaline earth oxide clusters are cuboids formed of \(4 \times 2 \times 2\) \((N = 8)\), \(3 \times 3 \times 2\) \((N = 9)\) and \(5 \times 2 \times 2\) \((N = 10)\) atoms – where cations and anions alternating on those Cartesian grids. These cuboid structures are found to be metastable for \((\text{PbO})_N\), however as they are not ranked within the top 5 for their respective size, they are not shown in either table 4.1 or table 4.2.
Table 4.1 Ball-and-stick models for the 5 lowest energy, PBEsol (as implemented within FHI-aims) LM structures of \((\text{PbO})_N\) for sizes \(N=6-9\), where grey and red spheres represent Pb and O ions, respectively.

<table>
<thead>
<tr>
<th>(N = 6)</th>
<th>(E_p^{GM} = -2.5287\ \text{(-8.2018) eV})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Delta E_p/eV)</td>
<td>(D_{3d})</td>
</tr>
<tr>
<td>SI /Å</td>
<td>0.0055</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(N = 7)</th>
<th>(E_p^{GM} = -2.5571\ \text{(-8.1198) eV})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Delta E_p/eV)</td>
<td>(C_{3v})</td>
</tr>
<tr>
<td>SI /Å</td>
<td>0.0078</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(N = 8)</th>
<th>(E_p^{GM} = -2.5939\ \text{(-8.1940) eV})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Delta E_p/eV)</td>
<td>(S_d)</td>
</tr>
<tr>
<td>SI /Å</td>
<td>0.0108</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(N = 9)</th>
<th>(E_p^{GM} = -2.6010\ \text{(-8.1994) eV})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Delta E_p/eV)</td>
<td>(C_{5v})</td>
</tr>
<tr>
<td>SI /Å</td>
<td>0.0052</td>
</tr>
</tbody>
</table>
Considering the ionic radius of Pb$^{2+}$ is between those of Sr$^{2+}$ and Ba$^{2+}$\textsuperscript{43} [131] and the oxides of those alkaline earth metals share the same ground state structure, it may be reasonable to expect the same ground state nanocluster structures for (PbO)$_N$. Such an expectation is only valid without the knowledge of the existence of the lone pair electrons on the Pb$^{2+}$ cation, which can drive

\textsuperscript{43} Crystal ionic radii of elements (6-coordinate): Pb$^{2+}$ (133), Sr$^{2+}$ (132) and Ba$^{2+}$ (149) in pm.
distortions of the atomic structure. Related to this lone pair effect, there is the preference for a lower coordination number of the Pb\(^{2+}\) cations, especially for the ground state structures, which tends to be as low as three. For example, consider the Pb\(^{2+}\) cations at the edges of fourth LM structure for \(N = 6\). It is energetically favourable to have a coordination number of three to secure enough space for lone pair polarisation, which was verified with both LPM and DFT results confirming that the rocksalt-like (PbO)_\(N\) structures are higher in energy. Exceptions only include the largest size clusters investigated, \(N = 11\) and 12, where one Pb\(^{2+}\) cation has a higher coordination, and the lone pair is not stereo-active, i.e., the contribution of model \(p\) orbitals to the valence state is negligible.

The reduction in the coordination number could be explained by the formation of a lone pair lobe next to Pb\(^{2+}\) cations, which replaces the bond to a fourth oxygen anion, as is seen in the LPM results. In contrast, the coordination number of four is more preferred in BaO and SrO nanocluster structures – and the LPM results also verify that the empty coordination space is occupied by lone pair electrons. Furthermore, the preference for adopting a lower coordination number is consistent with the trend observed in the crystal structure of \(\alpha\)-PbO where the coordination number of Pb is four, which is also lower than the typical value of six for lone pair free cations in other closed-packed oxides, e.g., SrO and BaO. Furthermore, the above observation is also supported by crystallographic and \textit{ab initio} studies [22,165] on Pb\(^{2+}\) complexes, showing that the low coordinated cation with a stereo-active lone pair is energetically more preferred.

For \(N = 8\), consider the fourth and ground state LM structures for PbO. The atomic structure of the former is essentially a hexagonal prism face sharing with a cuboid formed of 2 \(\times\) 2 \(\times\) 2 atoms and ranked second for (BaO)_8, whereas the latter can be constructed by inflating the former after breaking two bonds and is ranked third for (BaO)_8. The breaking of two bonds leads to the highest
two coordinated cations changing from 4 to 3. Similarly, the third LM for \((\text{PbO})_{10}\) can be constructed by inflating two face-sharing hexagonal prism (third LM for \((\text{BaO})_{10}\)), again by reducing the 4 coordinated cations to 3. For \((\text{PbO})_N\) and \((\text{MO})_N\) \(\{M = \text{Ba, Sr}\}\) with odd values of \(N\), there are examples of matching ground state structures. The ground state LM structure for \((\text{PbO})_7\) matches the ground state LM for \((\text{SrO})_7\), although ranked second for \((\text{BaO})_7\). The fifth LM structure for \((\text{PbO})_{11}\) matches the configuration for the second LM of \((\text{BaO})_{11}\).

### 4.3.2 \((\text{PbO})_N\) LM clusters: comparison of SM and LPM with DFT

#### 1) Computational Cost of Refining LM

One approach to gaining an insight into the benefit of employing the LPM, as opposed to the SM, is to consider the computational cost of refining their respective LM to the target DFT LM. We report the required wall-time of the FHI-AIMS runs to obtain the PBEsol LM starting from the LM atomic configurations for both LPM and SM in figure 4.4. Note that for each size of \(N\) for \((\text{PbO})_N\), 50 lowest energy LM obtained from LPM and SM were refined using DFT, and the 5 lowest energy LM from each interface were shown.

![Scaled DFT-PBEsol refinement wall-times](image)

**Figure 4.4** Scaled DFT-PBEsol refinement wall-times on an Intel(R) Xeon(R) Gold 6140 CPU @ 2.30GHz with 18 cores, as a function of the similarity index (SI) for the five lowest energy LM(PBEsol) configurations of \((\text{PbO})_N\), for \(N = 6\) to 12, before-and-after refinement. The markers • and × indicate LPM to PBEsol LM and SM to PBEsol LM, respectively. The black and red error ellipses, with a confidence interval of \(\sigma = 1\), are for the circle and cross markers, respectively.
For both sets of relaxation, LPM→DFT and SM→DFT, the wall-time shown in figure 4.4 is scaled with the number of atoms in the cluster,

\[
\text{scaled wall time} = \{\text{(raw wall-time)}/(\text{number of atoms})^2\}.\]

The horizontal axis indicates how much the atoms move during the relaxation, as measured by the similarity index (SI). As expected, there is a linear correlation between SI and the scaled wall-time, i.e., the closer a given configuration is to the DFT LM, the smaller the required computational cost to locate the DFT LM; see error ellipses in figure 4.4.

![Figure 4.5](image)

**Figure 4.5** The PBEsol LM energies, \(E_P(N)\), of the best 5 \((\text{PbO})_N\) nanocluster configurations found from performing a finite search on the LPM (●) and SM (x) energy landscapes for each size. Panel (b) shows a magnified section of panel (a) for sizes \(N = 10-12\).

2) Energetics

The energies, \(E_P(N)\), for LM of \((\text{PbO})_N\) nanoclusters shown in figure 4.5, are all relative to \(N\) infinitely separated \(N=1\), ground state clusters as defined in equation 4.3.1. From the SM and LPM, the 5 lowest energy LM found by DFT refinement were shown (i.e. the same data set of LM relaxation used in figure 4.4).
From figure 4.5 it can be seen that the lowest 5 DFT LM for $N = 6$ and 7 are found from relaxing the finite LPM and SM LM configurations on DFT energy landscape. On the other hand, for sizes $N = 8$-12, at least one DFT LM structure is missed from SM→DFT relaxations for each size. For example, the first and third DFT LM was not found for $N = 8$, and the first and second for $N = 9$.

In summary, employing the LPM, as opposed to the SM, in the global optimisation scheme that was introduced in figure 4.2 proves more success at finding lower-lying DFT LM nanoclusters. The data in figure 4.5 came from a finite number of global optimisation steps so does not imply that the missing DFT LM do not have a matching LM on the SM energy landscape. Instead, the width of the SM energy basin containing LM that match the DFT LM are either much smaller than the equivalent on the LPM energy landscape or we were very unlucky/lucky in the search on the SM/LPM landscapes. Ideally, we should perform a saturated search of each energy landscape and calculate the probabilities of finding each LM, including the missing nanoclusters.

Alternatively, we can focus on whether the SM landscape has more missing DFT LM than the LPM landscape by relaxing all DFT LM configurations on each energy landscape, i.e. DFT LM configurations on the LPM energy landscape (referred in this thesis as “LM transition DFT → LPM”) and DFT LM configurations on the SM energy landscape (referred in this thesis as “LM transition DFT → SM”). Thus, we can determine whether the missing DFT LM from relaxing the SM LM on the DFT energy landscape (LM transition SM → DFT) exist on the SM energy landscape without the need to perform an exhaustive global optimisation search. The results of this type of investigation are reported in figure 4.6.

Consider the case of $N = 6$ in figure 4.6. The second and third LM have SI values above 0.10 for the LM transition LPM → DFT (black circles in figure 4.6), which implies a poor match between the LPM and DFT LM structures and it is likely that the LPM LM is unstable on the DFT
energy landscape and then it relaxed to a different configuration. Relaxing these DFT LM on the LPM landscape (LM transition DFT → LPM, which is labelled “LPM → DFT → LPM” and marked as black plus signs in figure 4.6), the SI value of second LM structure is reduced (0.2231→0.0476) indicating that the originally LPM LM is indeed unstable on the DFT landscape and the new LPM LM is a better match with the DFT LM. In contrast, for the third LM structure there is an increase in the value of SI (0.1419→0.2389) and, unfortunately, is much higher than all LM transitions involving the SM, which are stacked at the same low value suggesting a better mapping for this particular DFT LM with the SM. Whereas, a similar mapping is found for the second DFT LM with both the newly found LPM and SM LM (similar SI value of 0.0511), which are themselves the ground state LM for both SM and LPM energy landscapes.

Figure 4.6 The similarity index (SI) as a function of the DFT energy, $E_P$, for the best 5 PBEsol LM nanoclusters of $(\text{PbO})_N$ with $N = 6-12$. Note that the SI values measured the differences between the DFT LM atomic structures and the respective LPM or SM LM structures.
For \( N = 8 \), where the first and third LM are missed for LM transition SM→DFT (missing red circle in figure 4.6), the DFT LM is shown to be unstable (does not exist as a LM) on the SM energy landscape as verified by the high SI values for the LM transitions DFT→SM (red crosses in figure 4.6) of 0.0764 and 0.1236, respectively. In contrast, the LPM LM matches with the fourth LM structure, which was missed from relaxing the non-complete set of LPM LM, was found, as indicated by the drop in SI values of 0.1090 to 0.0278.

For \( N = 9 \), a SM LM match was also found with the first and second DFT LM that were missed when the non-complete set of SM LM were relaxed on the DFT energy landscape. The reason why they were initially missed was because the SM energies of those structures are much higher in energy and above a threshold used when selecting SM LM candidates for DFT refinement.

For \( N = 10 \), the red circle and missing black circle for the second LM structure indicate it was only initially found from relaxing the non-exhaustive set of SM LM. However, this DFT LM configuration does not have a match on SM energy landscape (high SI value of 0.1805, after relaxation failed due to an unphysical shell-core separation of Pb\(^{2+}\) cations), whereas there is in fact a match between the DFT LM and a LPM LM that had an energy higher than the threshold used when selecting LPM LM candidates for DFT refinement.

In general, the LPM landscape shows a better mapping than the SM landscape compared to the DFT energy landscape, which has been demonstrated by matching low-lying LM nanocluster structures of \((\text{PbO})_N\). Whereas the SM was partially able to approximate some of the DFT LM structures, many of the DFT LM were found by relaxing non-matching SM LM configurations that were unstable on the DFT energy landscape, e.g., the second, fourth and fifth LM of \( N = 11 \) collapse to different configurations on SM energy landscape, according to the measured SI values.
(see figure 4.6). It is still debatable whether the employed interatomic potential parameters for the SM calculations were suitable for modelling PbO structures.

### 4.3.3 \((\text{PbO})_N\) GM clusters: second-order energies and their dipole moments

For the ground state structures, the second-order energies were calculated using the formula,

\[
E_{\text{SOE}}(N) = E(N) - \frac{1}{2} [E(N + 1) + E(N - 1)].
\]  

The second-order energy, \(E_{\text{SOE}}\), provides a measure of the relative stability of neighbouring cluster sizes. A negative value of \(E_{\text{SOE}}\) implies that the cluster is more stable, i.e. ground state clusters of the neighbouring sizes, 1 more and 1 less, will become more stable if the larger cluster transfers one formula unit to the smaller cluster.

**Figure 4.7** Second-order energies, \(E_{\text{SOE}}\), of the global minima \((\text{PbO})_N\) nanoclusters calculated by DFT and LPM. Ball-and-stick models (Pb black and O red) of the nanocluster structures shown above and below the data points are DFT LM and LPM LM, respectively.
In figure 4.7, the $E_{\text{SOE}}$ results for DFT and LPM are presented for ground state nanocluster structures of $(\text{PbO})_N$. It can be seen that the behaviour of $E_{\text{SOE}}$ as calculated using the LPM is in good agreement with that calculated using DFT, even though the energy ranking of LPM LM within each cluster size does not match perfectly with the ranking of DFT LM. According to the LPM relative energies of $N = 10, 11$ and 12 given in table 4.2, the global minima (GM), or ground state structures on DFT energy landscape are LM on the LPM energy landscape. Furthermore, the scale of $E_{\text{SOE}}$ is also different with that of LPM roughly four times larger than that of DFT.

A similar good match is found for the calculated dipole moment, using either DFT or LPM, of the ground state $(\text{PbO})_N$ nanoclusters; see figure 4.8. For $N = 5, 7, 10, 11$ and 12, the dipole moment of the $(\text{PbO})_N$ nanoclusters calculated by DFT are between ~ 0.1 and 3 D, whereas the dipole moment on the other sized ground state $(\text{PbO})_N$ nanoclusters are negligible. Curiously, nanoclusters with a high dipole moment show relatively higher value for $E_{\text{SOE}}$.

![Figure 4.8 Dipole moments of global minima $(\text{PbO})_N$ nanoclusters, calculated by DFT and LPM where the structures by each level of theory can be found in figure 4.7.](image)

**Figure 4.8** Dipole moments of global minima $(\text{PbO})_N$ nanoclusters, calculated by DFT and LPM where the structures by each level of theory can be found in figure 4.7.
For correlation between nanoclusters with negative $E_{SOE}$ and negligible dipole moments agrees with the behaviour reported for $(\text{BaO})_N$ [163], even though the ground state structures may be quite different from those of $(\text{PbO})_N$. For example, the $3 \times 3 \times 2$ cuboid structure is the $N = 9$ global minimum nanocluster for barium oxide, has a positive $E_{SOE}$ and strong dipole moments of $\sim 5$ D (as calculated using PBEsol). Whereas, the $(\text{PbO})_9$ ground state structure has 3-fold symmetry ($C_{3h}$ point group, see table 4.1), which contributes to reducing dipole moment and has a negative $E_{SOE}$.

On the other hand, the dipole moments calculated by LPM are of the order of $\sim 0.5 – 10$ D which is roughly 3 times larger for the $N = 5, 7, 10, 11$ and 12 $(\text{PbO})_N$ ground state nanoclusters in comparison with dipole moments calculated using DFT results. Here, LPM captures the trend of dipole moments qualitatively by reproducing the DFT nanocluster structures, however the disparity in the magnitudes of dipole moments was expected, since the charges of ions, not only for the electron lone pair but also the oxygen anions, are fixed on each localised sites, e.g., the model charge of LPM molecular orbital and the fixed shell and core charges of anions. Therefore, the model cannot readily redistribute or delocalise the charges around the system. In contrast, the DFT approach also has a larger set of basis functions and are on all ions, which gives more freedom for redistributing or delocalising electrons around the entire system and reduces the dipole moment on the cluster.

### 4.3.4 Effect of stereo-active lone pair in nanocluster structures

Low energy LM nanoclusters of alkaline earth oxides (MgO, CaO, SrO and BaO), which have the same cation to anion stoichiometry, typically adopt rocksalt lattice resembling atomic configurations, which are low in energy [166]. In this study, however, such configurations for $(\text{PbO})_N$ were rarely found as low energy local minimum on either DFT or LPM energy landscapes (see table 4.1...
and 4.2). The only examples found within the five most stable LM configurations for each size, are for the smaller size clusters of \( N = 4 \) and 6. Such results are in contrast to the low energy minima configurations of \((\text{SrO})_N\) and \((\text{BaO})_N\) with \( N = 8, 9, 10 \) and 12 cluster as mentioned in preceding section 4.3.1.

For an octahedral coordination, the Shannon ionic radius of \( \text{Pb}^{2+}, \text{Sr}^{2+} \) and \( \text{Ba}^{2+} \) is 133, 132 and 149 pm, respectively [131]. Thus, \( \text{Pb}^{2+} \) is very similar in size to \( \text{Sr}^{2+} \) and smaller than \( \text{Ba}^{2+} \). Assuming each species can be modelled as simple incompressible non-polarisable spheres, one can expect the configurations of \((\text{PbO})_N\) clusters should resemble those of strontium oxide. However, as previously stated, such an assumption is not valid as \( \text{Pb}^{2+} \) cations have residual valence electrons, which can form lone pairs that will influence the atomic configurations. In an attempt to gain an insight as to the effect of the creation of lone pairs on the atomic configurations and their stabilities, we now take a closer look at several LPM and DFT LM clusters.

1) Slab and cage configurations of \((\text{PbO})_{12}\)

For \( N = 12 \), it is possible to construct a rocksalt lattice configuration, or slab, formed of 4 \( \times \) 3 \( \times \) 2 (cations and anions alternating on this Cartesian grid) as shown in figure 4.9. By using LPM and DFT, the \((\text{PbO})_{12}\) slab was relaxed leading to structures labelled A1’ and A1, respectively, in figure 4.9 (throughout this section, the prime symbol ‘ indicates a LPM LM structure as opposed to DFT). On the DFT energy landscape, there was no change in the structural motif, i.e. the slab configuration was preserved, whereas for LPM, two Pb atoms (marked by blue arrows in figure 4.9) moved outwards and effectively reduce their coordination number from four to three. The reduction in coordination number is more likely during structural relaxations to minimise the LPM energy (corresponding to distortions by stereo-active lone pairs). On the other hand,
presumably the DFT model was able to distribute the charges of lone pairs around the cluster, which the LPM is not capable of, and thus retained the slab configuration. After relaxing the structure $A_1'$ on the DFT landscape, the structure labelled $A_2$ in figure 4.9 was obtained. Moreover, the similar configuration labelled $A_2'$ is also stable on LPM energy landscape, which is more stable than $A_1'$. The energy reduction from $A_1'$ to $A_2'$ can be explained by the changes in two Pb atoms (marked by green arrows) whose coordination numbers decreased from four to three.

The $T_h$ sodalite cage configuration for $(\text{PbO})_{12}$ is stable on both energy landscapes and are labelled as $B_1$ and $B_1'$ in figure 4.9. The noticeable structural feature of $(\text{PbO})_{12}$ in the cage configuration can be found where the Pb$^{2+}$ cations are sticking outward in order to secure the space to create stereo-active lone pairs. In contrast, it is the anions that are more polarisable than the cations and so stick out for SrO, BaO, BeO, and ZnO [167–170].
2) Tetromino configurations of (PbO)$_{10}$ and (PbO)$_{9}$

For these specific cluster sizes, a set of configurations so called the tetromino shapes – classified as Alphabet symbols of \{L, S, T, I\} for $N = 10$, \{O\} for $N = 9$ – are frequently found as low energy clusters of $(MO)_{N=9,10}$ with $M = Mg^{2+}$, $Ca^{2+}$, $Sr^{2+}$ and $Ba^{2+}$, which resemble the rocksalt lattice structures (see figure 4.10 and 4.11). These tetromino configurations were employed as initial structures, which were relaxed on the LPM and DFT energy landscapes to a LM to help gain insight of the atomic structural distortions caused by lone pair electrons in $(PbO)_{N=9,10}$.

In figure 4.10, the relaxation results of \{S, T, I\} are summarised. The relaxation result of $S$-mino gave the distorted LM structures of $S2$ and $S2'$, with the distortion driven in-part by the formation of lone pairs in order to secure space for the lone pair electrons to occupy. Here, the centred rock-salt $(PbO)_6$ subunit (which is itself a low energy cluster of $N = 6$, see table 4.1) is deformed into the quasi-hexagonal configuration for both cases. This results in bond breaking and some higher coordinated atoms reducing to 3.

If all higher coordinated atoms in $S$ ($S2$ or $S2'$) are reduced to 3 then the LM structure of $(PbO)_{10}$ becomes $S3$ or $S3'$ as shown in figure 4.10. Note going from $S2$ and $S2'$ to $S3$ and $S3'$, respectively, breaks the two Pb-O bonds indicated by the blue arrows. The energy of $S3$ ($S3'$) is 0.037 (0.100) eV less than $S2$ ($S2'$). Note that there are no imaginary frequencies for $S2$ ($S2'$) so are metastable and there would be at least one energy barrier to between $S2$ ($S2'$) and $S3$ ($S3'$) – in terms of approximating the energy landscape of DFT by using LPM, this result would correspond to what is described in figure 4.3(a).

Upon relaxing the $T$-mino configuration for $(PbO)_{10}$, which is reported to be the ground state for $(MgO)_{10}$ [171], the $T2$ ($T2'$) configuration was obtained where structural distortion occurred at the joining hexagonal subunit (see the top view of $T$-mino where the Pb atoms are marked by
The distortion is possibly caused by the Pb\(^{2+}\) cations whose coordination number is four, adjoining with two (PbO)\(_2\) subunits, and the configuration is unfavourable in terms of forming stereo-active lone pairs. This was verified by T3 (T3') cluster whose configuration is similar with T2 (T2') and 0.041 (0.032) eV less in energy. According to the top view of T3 (T3'), one can find that the coordination number of those adjoining Pb\(^{2+}\) cations are reduced to three, and the blue arrows). The distortion is possibly caused by the Pb\(^{2+}\) cations whose coordination number is four, adjoining with two (PbO)\(_2\) subunits, and the configuration is unfavourable in terms of forming stereo-active lone pairs. This was verified by T3 (T3') cluster whose configuration is similar with T2 (T2') and 0.041 (0.032) eV less in energy. According to the top view of T3 (T3'), one can find that the coordination number of those adjoining Pb\(^{2+}\) cations are reduced to three, and the
configuration becomes more advantageous to maximise lone pair formation where the preference is as one on discussed for the case of S3 (S3’).

The I-mino configuration for (PbO)$_{10}$, which is reported as the global minimum cluster of (MO)$_{10}$ with $M = \text{Mg}^{2+}, \text{Ca}^{2+}, \text{Sr}^{2+}$ and Ba$^{2+}$, relaxes to the I2 (I2’) configurations shown in figure 4.10, i.e. the cuboid structure was maintained. However, the energy of the I2 (I2’) LM cluster is 0.079 (0.185) eV higher than the ground state of (PbO)$_{10}$ – see table 4.2 – perhaps as a result of the Pb$^{2+}$ cations whose coordination numbers are four (as opposed to 3), pronounced stereo-active lone pairs causing unfavoured structural strains.

So far, the relaxation results of S-, T- and I-minoes has shown good agreement where both LPM and DFT gave the same configurations. On the other hand, the LPM and DFT results of L- and O-minoes disagreed, which are summarised in figure 4.11. The configuration of L-mino is reported as a low energy cluster of (MO)$_{10}$ with $M = \text{Mg}^{2+}, \text{Ca}^{2+}, \text{Sr}^{2+}$ and Ba$^{2+}$, however for (PbO)$_{10}$ it relaxes to structures L2 and L3’. Although they are in different configurations, each deformation was probably caused by the Pb$^{2+}$ cation with coordination number of five in the starting L-mino configuration where the environment is unfavorable to form pronounced stereo-active lone pair. Accordingly, the relaxation of this structure led to the reduction of the coordination number of these lead cations to four.

Interestingly the configurations of L2 and L3’ can be relaxed to LM structures L2’ and L3 i.e., both structures exist on the energy landscapes of LPM and DFT. However, the match in the energy ranking is not good: the relative energy of L2 (L2’) with respect to L3 (L3’) is -0.026 (0.081) eV and the LPM to DFT mapping corresponds to what is described in figure 4.3(b).

The configuration of O-mino is reported as the global minimum cluster of (MO)$_{9}$ with $M = \text{Ca}^{2+}, \text{Sr}^{2+}$ and Ba$^{2+}$. For (PbO)$_{9}$, this structure relaxed to O2 (O2’), but upon closer inspection
O₂′ was found to be unstable (some vibrational modes had imaginary frequencies) and can be subsequently refined to O₃′, which has a matching DFT LM structure O₃. We also found the cluster O₄ (O₄′) – obtained from the global optimisation for (PbO)₉, which is the sixth LM and thus omitted in table 4.1 – whose structure is similar to that of O₂ (O₂′). Here, even though LPM LM configurations exist for O₃ and O₄, the energetics is other way round as the preceding case of L₂ and L₃, where the relative energy of O₃ (O₃′) with respect to O₄ (O₄′) is 0.012 (-0.152) eV, i.e., according to the DFT calculations O₄ is more stable than O₃.

**Figure 4.11** Ball-and-stick models of LPM and DFT LM configurations for (PbO)₉ and (PbO)₁₀ clusters. Black and dark green spheres represent Pb atoms for DFT and LPM respectively, whereas red spheres represent O atoms. Lone pair electron densities shown for LPM clusters as semi-transparent dark green surfaces. Black and green arrows indicate the flow of relaxation carried out by using DFT and LPM, respectively.
4.3.5 Concluding remarks

The atomic structures of lead monoxide nanoclusters based on LPM and DFT of sizes up to (PbO)\textsubscript{12} were investigated. In general, the application of the newly developed LPM turned out effective in terms of approximating the DFT energy landscape, which was demonstrated by matching DFT LM with LPM LM. Furthermore, the trends predicted by LPM for second-order energies and correlated dipole moments showed good agreement with that predicted by DFT. The LPM was also compared with the shell model, and in general LPM calculated clusters provided better accuracy in predicting atomic configurations and energetics. The structures of ground state (PbO)\textsubscript{N} nanoclusters are typically different to those of (MO)\textsubscript{N} where \( M = \text{Mg}^{2+}, \text{Ca}^{2+}, \text{Sr}^{2+} \) and \( \text{Ba}^{2+} \), which prefer cuboid structures. In contrast, the cuboid structures for (PbO)\textsubscript{N} were energetically unfavoured and the ground state configurations of lead monoxide were typically more amorphous and less symmetric due to the presence of stereo-active lone pairs. Relaxing LM configurations of (MO)\textsubscript{N} where \( M = \text{Mg}^{2+}, \text{Ca}^{2+}, \text{Sr}^{2+} \) and \( \text{Ba}^{2+} \), using both LPM and DFT methods typically gave similar resulting structures (good agreement between LPM and DFT LM).

4.4 Nanoclusters of lead fluoride

(PbF\textsubscript{2})\textsubscript{N} nanoclusters have been chosen as our next target for LPM applications. As lead fluoride has a different cation-anion stoichiometric ratio from the lead monoxide, structure prediction for (PbF\textsubscript{2})\textsubscript{N} clusters is not only an interesting challenge but also a good opportunity to test parameter transferability of our model.

This study is structured similar to the previous application, using the same global optimisation technique, described in figure 4.2. We will carry out quality assessment of (PbF\textsubscript{2})\textsubscript{N} nanoclusters, predicted using LPM, judged by reproduction of DFT refined atomic configurations, energetics
and dipole moments. The shell model (SM) was employed alongside as an additional control group to test the efficiency of LPM in capturing DFT-like structures. Lastly, we will also discuss the effect of lone pairs on nanocluster structures in comparison to other fluorides including lone pair free cations.

**LPM parameters:** To verify the transferability of LPM parameters, we have kept the changes in model parameters to a minimum. The Pb-Pb interaction parameters were retained from the previous application, and only the Pb-F parameters were fitted using atomic configurations and energy ranking of small-sized \((\text{PbF}_2)_{N=1-2}\) clusters, see table D.2 in Appendix D. Note that – as in the previous application to lead monoxide – for LPM calculations, a published interatomic potential was employed where the onsite spring and Coulomb-Buckingham potentials were used to describe the \(\text{F}^-\) anion and the \(\text{F}^-\text{F}^-\) interaction, respectively [160].

The same interatomic potentials were employed for \(\text{F}^-\) anion in the SM calculations and the other parameters of the \(\text{Pb}^{2+}\text{F}^-\) interaction were fitted to reproduce lattice parameters and static dielectric constant of fluorite \(\text{PbF}_2\), which was mentioned earlier in section 4.2.4(2) where the parameters are in table D.1, see Appendix D.

**4.4.1 Structures of \((\text{PbF}_2)_N\) clusters \((N = 3 - 8)\)**

In tables 4.3 and 4.4, we present the low energy DFT LM structures for nanoclusters composed of \(N\) formula unit. In the tables, the structure and energy of all structures have been refined using DFT, and vibrational analyses confirmed that the structures are stable minima. The reported relative energies are with respect to the ground state of each size, \(E_{p}^{GM}(N)\) (the energy of the global minimum structure of each size), and the values in parentheses are obtained from the LPM.
Table 4.3 Ball-and-stick models for the 5 lowest energy, PBEsol (as implemented within FHI-aims) LM structures of \((\text{PbF}_2)_N\) for sizes \(N=3-6\), where black and white spheres represent Pb and F ions, respectively.

<table>
<thead>
<tr>
<th>(N)</th>
<th>(E_p^{GM})</th>
<th>(\Delta E_p/\text{eV})</th>
<th>(\text{Sl./\AA})</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>-1.2732 eV</td>
<td>0.0 (0.0)</td>
<td>0.0045</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(D_{3h})</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_s)</td>
<td>0.0164</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_2)</td>
<td>0.2388</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_s)</td>
<td>0.0180</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_{2v})</td>
<td>0.0048</td>
</tr>
<tr>
<td>4</td>
<td>-1.4414 eV</td>
<td>0.0 (0.0012)</td>
<td>0.0095</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_{2v})</td>
<td>0.0163 (0.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_{3v})</td>
<td>0.0037</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_2)</td>
<td>0.0585 (0.1693)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(D_{2d})</td>
<td>0.0075</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_s)</td>
<td>0.0125</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_s)</td>
<td>0.0882 (0.2856)</td>
</tr>
<tr>
<td>5</td>
<td>-1.5299 eV</td>
<td>0.0 (0.0218)</td>
<td>0.0075</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(D_{2d})</td>
<td>0.0012 (0.0050)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_s)</td>
<td>0.0085</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_s)</td>
<td>0.0092 (0.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_s)</td>
<td>0.0199 (0.0160)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_1)</td>
<td>0.0204 (-)</td>
</tr>
<tr>
<td>6</td>
<td>-1.5828 eV</td>
<td>0.0 (0.0146)</td>
<td>0.0056</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_{2v})</td>
<td>0.0016 (0.0154)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_1)</td>
<td>0.0071</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_s)</td>
<td>0.0101</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_s)</td>
<td>0.0240 (0.0504)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_s)</td>
<td>0.0176</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(C_s)</td>
<td>0.0260 (0.0)</td>
</tr>
</tbody>
</table>

Table 4.3 Ball-and-stick models for the 5 lowest energy, PBEsol (as implemented within FHI-aims) LM structures of \((\text{PbF}_2)_N\) for sizes \(N=3-6\), where black and white spheres represent Pb and F ions, respectively.
calculations. Below each structure, a similarity index (SI) is shown, which is used as in table 4.1 and table 4.2.

Examination of the configurations of (PbF$_2$)$_N$ clusters shows a pronounced structural distortion. Although in the thermodynamically stable crystal structure of PbF$_2$ (cubic fluorite-type), the lone pairs of Pb$^{2+}$ cations are non-stereo-active, which means that the statistical distribution (including effects of thermal lattice vibrations) of the valence electrons around lone pair cations is symmetrical [14]. In contrast, in a molecular cluster of PbF$_2$, the Pb$^{2+}$ cations become stereo-active, by exposing their lone pairs on the surface of the cluster as shown in tables 4.3 and 4.4. It is still possible to recognise the motif (rhombic dodecahedron) of crystalline PbF$_2$ lattice in its nanoclusters, for instance, in the first and fourth ranked minima of (PbF$_2$)$_6$, see table 4.3, however the
presence of extra anions (e.g., F⁻ just outside and inside of the cluster) significantly distorts the structures.

4.4.2 (PbF₂)₅ LM clusters: comparison of SM and LPM with DFT

1) Computational Cost of Refining LM

As it was carried previously for lead monoxide clusters, we again report the required computational wall-time for DFT (FHI-AIMS) calculations to refine LM atomic configurations for both LPM and SM, which is summarised in figure 4.12.

In the figure, the scaled wall-time was measured same as in the previous analysis in section 4.3.2(1). For the given sets of relaxation, LPM→DFT and SM→DFT, the refinement results starting with LMP structures congregate more than staring with SM where this time the error ellipses are omitted due to the obvious difference between the two data sets.

![Figure 4.12 Scaled DFT-PBEsol refinement wall-times as a function of the similarity index (SI) for the five lowest energy LM(PBEsol) configurations of (PbF₂)₅, for N = 3 to 8, before-and-after refinement. The markers • and x indicate LPM to PBEsol LM and SM to PBEsol LM, respectively – the same computer specification was used as in figure 4.4.](image)

Considering that the same finite number (low energy 50 as the same number of samples used in the previous study in section 4.3) of LPM and SM LM was used for each set, this finding clearly demonstrates that compared to SM using LPM as starting points for DFT refinement is a more successful strategy for finding DFT atomic configurations, and at the same computational cost.
2) Quality of LPM and SM in approximating DFT energy landscape

Since the DFT refinement of LPM and SM LM structures was carried out on a finite number of samples, it requires the same type verification as the one performed on the clusters of lead monoxide in the previous study. We should answer, for instance, the question whether a nanocluster found on the transition LPM→DFT map truly exists on the transition SM→DFT map and vice versa. If a LM of LPM or SM structure is merely not included within the samples for DFT refinements, there is still possibility to find the low energy DFT cluster by increasing the number of sample precursors, given that a plenty of computational resources is available. Using the same procedure as in the previous work on lead monoxide, the data are summarised in figure 4.13.

**Figure 4.13** The similarity index (SI) as a function of the DFT energy, $E_p$, for the best 5 PBEsol LM nanoclusters of (PbF$_2$)$_N$ with $N = 3$-8. Note that the SI values measured the differences between the DFT LM atomic structures and the respective LPM or SM LM structures.
In figure 4.13, for the case of \( N = 3 \), out of the 5 low energy LM, the third LM is missed for LM transition LPM \( \rightarrow \) DFT, which shows SI value of 0.2388 marked as black circle, and the rests are with overall SI values less than 0.02. The missing third LPM LM gives SI value of 0.0746 for LM transition DFT \( \rightarrow \) LPM (labelled “LPM \( \rightarrow \) DFT \( \rightarrow \) LPM” and marked as black plus signs in the figure) whose SI value is reduced significantly, 0.2388 \( \rightarrow \) 0.0746.

In contrast, only the first and third LM transitions of SM \( \rightarrow \) DFT were found marked as red circles, and the rests are omitted since they are not found, i.e. they were missed from the initial DFT refinement for SM samples. The missing SM LM are turned out to show high SI value, above 1.5, found in LM transition DFT \( \rightarrow \) SM, which is labelled “SM \( \rightarrow \) DFT \( \rightarrow \) SM” (marked as red plus). For the LM transitions to DFT of both LPM and SM, they failed to reproduce the structure of third DFT LM. The difference in configurations can be investigated in more details by the total correlation function, given in figure 4.14 (d). For this correlation plot, the third DFT LM has a bent structure, see table 4.3, however the structure obtained by LM transition DFT \( \rightarrow \) LPM is linear. Such difference can be verified by the plot (d) in figure 4.14, where the curve of LPM (dashed red) shows sparser distribution than that of DFT (solid black).

In figure 4.13, for \( N = 5 \), the DFT GM structure was initially found in the LM transition LPM \( \rightarrow \) DFT map with a low SI value of 0.0075. However, its inverse LM transition, DFT \( \rightarrow \) LPM, gave 0.1716. This result is due to the transition of LPM structure that turned into the second LM structure, and the reason was turned out that the LPM structure (used for the LM transition LPM \( \rightarrow \) DFT) has a few imaginary vibrational frequencies. Consequently, the small displacements, caused by refining DFT GM structure on the LPM energy landscape, i.e. LM transition DFT \( \rightarrow \) LPM, dropped out the LPM structure from the saddle point to other low energy LM, which can be found.
in table 4.3 that the energy ranking of LPM is other way round with that of DFT for the first two LM.

In the same size cluster, LPM failed to capture the fifth DFT LM where there is no black circle mark indicating LM transition LPM → DFT. In its inverse, LM transition DFT → LPM (marked by black cross) shows SI value of 0.1321, which does not show good structural agreement with DFT LM, and this implies the structures is not stable on the LPM energy landscape.

For $N = 6$, the second DFT LM was not initially captured where LPM → DFT (marked as black circle) shows SI value of 0.0477, however later the inverse LM transition DFT → LPM show 0.0102 (marked as black plus). Such difference in SI values also can be verified by the changes in total correlation function, see plot (a) in figure 4.14 where the red curve represents the latter transition agrees more with DFT (solid curve) than the former (black dashed curve).

**Figure 4.14** Total pair correlation function for a few cases of $(\text{PbF}_2)_N$ structures where the different type of curves represent which type of energy definition is used, which is distinguished by the legend provided within the plot (d). Note that the Gaussian broadening factor of 0.05 Å is used.
For the third DFT LM, it was not found from LM transition SM → DFT but verified by its inverse, DFT → SM (marked by red plus), with SI value of 0.0286, and whose SI value of LPM → DFT is 0.0101. Since both cases show low SI values, and the total correlation function of LPM and SM structures show good agreement with that of DFT, see plot (b) in figure 4.14.

For \( N = 7 \), in figure 4.13, none of LM transitions SM → DFT succeeded to capture DFT LM (where there are no red markers in the plot) except the fourth LM with SI value of 0.0529, which is in contrast 0.0175 for LPM → DFT → LPM (where their total correlation functions are given in plot (c) in figure 4.14). Furthermore, for their LM transitions DFT → SM show high SI values above 0.1, which implies that those SM structures are not stable on the DFT energy landscape. In contrast, the overall SI values, measured from LPM → DFT, are low as below 0.01 (except the fourth LPM LM), and the similar trend is found in the cases of \( N = 8 \) in figure 4.13.

Overall, again as in the previous study for lead monoxide, the LPM energy landscape shows better ability in terms of approximating the DFT energy landscape than the SM landscape. We have investigated a few cases where LPM initially failed to find DFT resembling structures, however, using an inverse DFT→LPM map, similar configurations can be found on the LPM energy landscape. The SM was able to approximate some of the DFT LM structures, however, many of the DFT LM were found by SM structures that were unstable on the DFT energy landscape.

3) Energetics of LPM comparison with different types of DFT functionals

As discussed above and also in section 4.3, studying nanoclusters of lead monoxide, the problem of LPM is in the energy ranking of different LM on the energy landscape compared to DFT, we recall that in these tests we have used the PBEsol exchange and correlation density functional. For instance, in the PBEsol energy ranking for \( N = 6 \), the first and second ranked LM structures come
up only as third and fourth ranked on the LPM energy landscape. The PBEsol functional, however, is still only an approximation. Consequently, we have tested the effect of the functional choice on the structure ranking taking the size of \( N = 6 \) as an example and refining the PBEsol LM structures using popular hybrid PBEsol0 (hybrid functional in analogy to PBE0 \[172\])\(^{44}\) and B3LYP \[173\] exchange-correlation functionals (the same basis set was employed as in PBEsol calculations implemented in the FHI-AIMS code, see section 4.2.3(3)), which are considered to be more accurate, results are summarised in table 4.5. Note that the atomic configurations of clusters were preserved on refinement with both functionals, verified by the SI values provided alongside in the table.

![Table 4.5](image)

Table 4.5 Relative polymerisation energy of \((\text{PbF}_2)_6\) nanoclusters – taken from table 4.3 – calculated by using different DFT functionals where the numbers in parentheses indicate energy ranking for each functional. Note that the similarity index (SI) was measured with respect to the PBEsol nanocluster structures.

\( a \) Relatively high SI value was measured due to the displacement of a single \( F^- \) anion at the centre of the configuration, whose position is slightly off centre (see table 4.3), however, the position reserves to the centre according to the PBEsol0 and B3LYP results.

\( b \) SI value from the DFT(PBEsol)→LPM map was taken (see figure 4.13) instead of using the value given in the table 4.3.

\( ^44 \) Except that the PBEsol GGA functionals are used, mixing 75\% GGA exchange with 25\% Hartree-Fock exchange.
On PBEsol0 refinement, the ranking of the GM structure is preserved while there are a few changes for the higher energy structures. The energy ranking by B3LYP is quite different from that of PBEsol where the third structure becomes the GM, and the first ranked structure – the GM by PBEsol – becomes the fourth. The ranking by B3LYP proves to be closer to LPM, even though the rankings are still reversed for (i) the first and second structures, and (ii) the third and fifth structures. Furthermore, the ranking of the fifth LM structure agrees in both energy definitions. As the comparison of different DFT functionals shows, the energy ranking of different structures is not set in stone, LPM rankings proved to be similarly scattered to DFT. Since the DFT energetics may vary from functional to functional [174], we can reassign the primary objective for LPM used for structure prediction to narrow down the number of candidates, which lie plausibly within low energy basins, instead of pursuing prediction of “correct” energy rankings.

4.4.3 (PbF₂)_N GM clusters: second-order energies and their dipole moments

As shown in the previous study, another useful measure of structure stability is the second-order energy difference (see equation 4.3.2) of global minima nanoclusters, a plot of which with cluster size is shown for (PbF₂)_N in figure 4.15. Note that for the second order energy differences calculated by LPM, the nanoclusters used may not be the global minima structures on their own energy landscape (see table 4.3 and 4.4).

According to the figure, the overall trend of \( E_{SOE} \), calculated by LPM, agrees with that of DFT. The most problematic case is at \( N = 4 \) where the \( E_{SOE} \) of LPM shows a sudden drop down away from the DFT data, which is due to the relative stability of the \((\text{PbF}_2)_4\) nanocluster calculated by LPM with respect to its neighbouring sizes, i.e., \((\text{PbF}_2)_3\) and \((\text{PbF}_2)_5\).
The most distinct feature of $E_{SOE}$ profiles for $(\text{PbF}_2)_N$ is that there is no alternating pattern instead showing consistent negative values approaching zero. In contrast, the $(\text{PbO})_N$ nanoclusters in the previous study showed alternating positive and negative values of $E_{SOE}$, depending upon the dipole moments of the clusters. When the dipole moment of a cluster was getting close to zero, the $E_{SOE}$ of $(\text{PbO})_N$ tended to show negative value and *vice versa* (see figure 4.7). The difference in the pattern of $E_{SOE}$ for the two different compounds, $(\text{PbF}_2)_N$ and $(\text{PbO})_N$, should relate to their different cation-anion stoichiometric ratio.

We refer to results shown in figure 4.7 for $(\text{PbO})_N$, a compound with a 1:1 cation-anion ratio. For instance, it was relatively easy for the clusters with $N = 4, 6$ and $8$ to form atomic configurations of high symmetry that minimise their dipole moments unlike clusters with odd $N = 3, 5$ and
7. For \((\text{PbF}_2)_N\) clusters with a 1:2 cation-anion ratio it is comparatively more difficult to form symmetric atomic configurations. In figure 4.16, for instance, one can see that the \((\text{PbF}_2)_N\) clusters, whose atomic configurations are shown in figure 4.15, consistently exhibit non-zero dipoles, with appreciable values of \(1.2 - 5.2\) D for \(N = 4-8\). The only exceptional case is \((\text{PbF}_2)_3\) with negligible dipole moment due to its highly symmetric configuration of the \(D_{3h}\) point group.

![Figure 4.16 Dipole moments of global minima (PbO)_N nanoclusters, calculated by DFT and LPM where the structures by each level of theory can be found in figure 4.15.](image)

**4.4.4 Transferability of LPM for simulating lone pair free cations**

1) **Effect of lone pairs on nanocluster structures**

As a part of LPM validation in chapter 3, we have investigated the effect of model parameters: (i) charge of lone pair \(\omega\) and (ii) the onsite energy term \(\lambda\), responsible for separating the energy levels of \(s\) and \(p\) states, where the details can be found in section 3.1.1. Here, especially the term \(\lambda\) is intuitive to understand the LPM. For instance, if the value of \(\lambda\) becomes large, the LPM cation will hardly be polarised due to the large energy separation between the \(s\) and \(p\) state,
Figure 4.17 $(MF_2)_N$ nanoclusters with $N = 2$ to 7 where $M =$ Sr and Ba, optimised using DFT PBEsol. In column A and B, the structures of $(MF_2)_N$ are shown overlapped on top of the LPM optimised $(PbF_2)_N$ structures with artificially adjusted model parameter $\lambda = 30.09$ and $32.09$ eV used for respective compounds (cf. the value of $\lambda$ employed for lead fluoride is $14.09$ eV). In the fifth column, the structures of $(PbF_2)_N$ are shown, also optimized using DFT(PBEsol) for comparison. The ball-and-stick model is used with Ba (green), Sr (blue), Pb (black) and F (white). The red lines on the structures of $N = 2$ are drawn to indicate the bond angles $\angle F-M-M (M =$ Ba, Sr and Pb).
i.e., the hybridisation of $s$ and $p$ states becomes blocked analogous to the effect of large spring constant within the shell model ion. As related validation, we have shown the response of LPM cation as function of $\lambda$ in figure 3.2 and the discussion can be found in section 3.1.3. Thus, by tuning parameter $\lambda$, it is possible to force LPM cations to behave like lone-pair-free cations. Although this procedure is somewhat artificial, it could still be useful to demonstrate the relationship between lone pairs and structural deformations.

2) LPM reproduced nanocluster structures of ($\text{SrF}_2)_N$ and ($\text{BaF}_2)_N$

In figure 4.17, we presented the nanocluster structures of ($\text{SrF}_2)_N$ and ($\text{BaF}_2)_N$ with $N = 2$ - 7, which are shown in the first two columns, calculated using DFT (PBEsol is used as for lead monoxide/fluoride). Examination of their atomic configurations shows that their structures are similar to those of ($\text{PbF}_2)_N$ low energy nanoclusters provided in the fifth column, taken from tables 4.3 and 4.4. In the figure, for $N = 2$, a clear difference between the structures of ($\text{PbF}_2)_N$ and the other compounds, ($\text{SrF}_2)_N$ and ($\text{BaF}_2)_N$, with lone-pair-free cations is the bond angle $\angle \text{F-M-M}$ (marked with red lines in the figure), including terminal F$^-$ anion, that changes from 92.0° for $M=\text{Pb}^{2+}$ to 132.3° for $M=\text{Sr}^{2+}$ and to 120.1° for $M=\text{Ba}^{2+}$. Here, the strong bending occurs in $\angle \text{F-Pb-Pb}$ is due to the presence of stereo-active lone pairs on the Pb$^{2+}$ cations that are repelled by the F$^-$ anions.

In the figure, the columns A and B show the LPM optimised ($\text{PbF}_2)_N$ nanoclusters with artificially adjusted parameter $\lambda$ whose structures are overlapped with DFT calculated ($\text{SrF}_2)_N$ and ($\text{BaF}_2)_N$. In order to reproduce the structures of ($\text{SrF}_2)_N$ and ($\text{BaF}_2)_N$, the parameter $\lambda$ was gradually increased to lower the polarisabilities of the two lone-pair-free cations, Sr$^{2+}$ and Ba$^{2+}$. Thus, it was possible to keep the LPM cations from excessive $sp$-hybridisation and make them behave
like Sr\(^{2+}\) and Ba\(^{2+}\) cations. This effect of reduced polarisation can be found as the configurations of LPM calculated \((\text{PbF}_2)_N\) clusters with constrained \(\lambda\) parameter fall in good coincidence with the nanoclusters of the lone-pair-free compounds.
To carry out a more detailed analysis of the effect of lone pair stabilisation, we have employed various $\lambda$ parameters in the LPM calculations of the $(\text{PbF}_2)_N$ nanoclusters given in figure 4.17, and the resultant atomic configurations were compared with those of $(\text{BaF}_2)_N$ and $(\text{SrF}_2)_N$. The results are summarised in figure 4.18.

In panels 1a and 1b of the figure, the measured similarity indices (SI) first show a decreasing trend, i.e., the atomic configurations of $(\text{PbF}_2)_N$ calculated using a “tweaked” LPM start agreeing with those of $(\text{BaF}_2)_N$ and $(\text{SrF}_2)_N$ with parameter $\lambda$. The agreement is valid, however, only up to a certain value of $\lambda$; for instance, as one can expect, at the right-hand extreme of $\lambda \sim 46$ eV, the atomic configuration of $(\text{PbF}_2)_2$ becomes linear, in this region the LPM cation starts behaving like a rigid ion. For these SI values (shown with circle markers), one finds minimum points along each line. In particular, in the case of $N = 2$ in panels 1a and 1b, the minimum points are at $\lambda = 30.09$ and $32.09$ eV, which were taken to optimise $(\text{PbF}_2)_N$ clusters presented in figure 4.17 A and B.

Overall, with these artificially adjusted parameters $\lambda$, for the different cluster sizes of $N = 3-7$ we observe smaller SI values. Considering that the original value of parameter $\lambda$, used to reproduce the DFT structures of $(\text{PbF}_2)_N$, was 14.09 eV, this experiment demonstrates the effect of lone pair polarisation on the nanocluster scale, and furthermore provides additional possibility for LPM to represent lone-pair-free cations using an adaptively tuned model parameter.

It is useful for this analysis to define the expansion factor $\varepsilon_f$, which minimises the similarity index (SI):

$$
\varepsilon_f = \arg \min_{\varepsilon_f} \text{SI}(r_A - \varepsilon_f r_B)
$$

(4.4.1)

where $r_A$ and $r_B$ are vectors, representing the coordinates of $n$ number of atoms in nanoclusters $A$ and $B$. As discussed earlier in section 4.2.4 (1), for a pair of clusters, SI value – simple root-mean-
squared deviation of two different atomic configurations – is measured after those two nanoclusters are translated so that their respective centre of mass is at the origin and axes of their principal moments of inertia aligned by rotation, see Appendix B, it is possible to find the expansion factor $\varepsilon_f$, minimising the SI value. The expansion factor can also be used to show the difference in cluster sizes. For instance, suppose $r_A$ and $r_B$ refer to coordinates of $(\text{PbF}_2)_N$ and $(\text{BaF}_2)_N$, the expansion factor $\varepsilon_f < 1.0$ would imply that the overall size of $(\text{BaF}_2)_N$ is greater than that of $(\text{PbF}_2)_N$.

In panels 1a and 2b of figure 4.18, alongside the “raw” (circle) values of the SI shown with circle markers we have plotted with cross makers where the corresponding SI values are after applying an expansion factor to atomic configurations. For the case of $(\text{SrF}_2)_N$ in panel 1b, there was no significant reduction in SI values after application of the expansion factor. In contrast in panel 1a, the reduction in SI values is evident. Considering that the ionic radius of $\text{Ba}^{2+}$ is greater than that of $\text{Sr}^{2+}$, and the latter one is similar to radius of $\text{Pb}^{2+}$, the larger disparities in SI values in figure 1a, before-and-after applying the expansion factor, are clearly due to the size difference between different cations. This conclusion is supported by the calculated $\varepsilon_f$ values shown in figure 2a and 2b where their averages are 0.94 ($\lambda = 30.09$) and 0.99 ($\lambda = 32.09$), respectively, which also agrees well with the bond-length ratio of $r_{\text{PbF}}/r_{\text{MF}}$: 0.93 ($M = \text{Ba}$) and 0.99 ($M = \text{Sr}$) for the $N = 2$-$7$ clusters (DFT calculated) used in this experiment. Thus, a simple analysis of expansion factors highlights the need for additional parameter tuning in LPM, for example if we wish to simulate $(\text{BaF}_2)_N$, LPM parameters related to the cation-anion interaction will be necessary to adjust to reproduce correct Ba-F bond lengths.
4.4.5 Concluding remarks

Following application of our new model to lead monoxide, we have investigated nanoclusters of \((\text{PbF}_2)_N\), employing the same global optimisation scheme, and found that LPM was effective in predicting atomic configurations of nanoclusters found on the DFT energy landscape. Here, we mainly focused on the LPM parameter transferability, and therefore we kept the model parameters related to the lone pair cation unchanged from what we have used in the previous study (nanoclusters of lead monoxide in section 4.3). Compared with the polarisable shell model, whose parameters were also similarly retained, the LPM shows a better accuracy in predicting DFT resembling structures.

Considering the energetics of LPM, the ranking of nanoclusters was found to be not fully matching the DFT-PBEsol results. Further DFT calculations for low energy structures of \((\text{PbF}_2)_6\), have shown that the energy ranking is sensitive to the choice of the exchange-correlation functional employed as confirmed by comparison of our original PBEsol data with results obtained using popular PBEsol0 and B3LYP functionals. This observation implies that instead of targeting the energy ranking, the objective of using semi-empirical potentials (i.e., including LPM and the polarisable shell model) for structure prediction may be reproduction of correct atomic configurations, which would allow one to reduce the computational cost, especially required for the following DFT refinement process.

For the low energy nanoclusters of \((\text{PbF}_2)_N\), we found that they tend to show non-negligible dipole moments, unlike the lead monoxide clusters from the previous study. The main source of the dipole moment may relate to the cation-anion stoichiometric ratio of \((\text{PbF}_2)_N\), which hinders stabilisation of symmetric atomic configurations that could effectively suppress the dipole creation. The LPM calculated dipole moments are also found in good agreement with DFT, which highlights
the importance of the atomic configuration for reproducing accurate physical properties. Furthermore, we have compared the \((\text{PbF}_2)_N\) with other lone-pair-free compounds, \((\text{SrF}_2)_N\) and \((\text{BaF}_2)_N\). By manipulating the LPM parameter \(\lambda\), we have demonstrated that LPM can also reproduce their structure. This investigation shows a potential applicability of LPM to simulating compounds with non-lone-pair species.

For the overall efficiency of global search, using both LPM and SM, in figure 4.12, for the case of \((\text{PbF}_2)_N\) nanoclusters, the averaged scaled wall-time for refining LPM and SM clusters on the DFT energy landscape are 2.18 and 5.79, respectively, where the refinement cost of SM clusters is 2.66 times more than those of LPM. On the other hand, in figure 4.4, for the case of \((\text{PbO})_N\), both LPM and SM show similar cost for the DFT refinements with 3.84 and 3.76, respectively. However, one could give points to LPM since it requires a smaller number of samples to characterise the energetically low-lying nanoclusters on the DFT energy landscape. This is evidenced by figure 4.5 and 4.6, recall that same number of samples were taken from LPM and SM for their refinements, some of the low-lying nanoclusters on the DFT energy landscape were missed from the refinement process of SM. Consequently, in terms of computational cost, would LPM then show better performance for global optimisation of nanoclusters than SM? The answer depends on how vastly searching the energy landscape. Since the computational cost of LPM is still an order of magnitude more expensive, practically taking tens of seconds to optimise a nanocluster, c.f., SM, taking a few seconds\(^45\). This means that SM could collect more samples than LPM for a given amount of time. However, every SM samples cannot go through the DFT refinement process,

\(^{45}\) The refinement wall clock time for DFT calculations, in figure 4.4 and 4.12, ranged from a few hundreds to thousands of seconds, depending on the size of nanoclusters, i.e., which are, even disregarding the number of used CPUs – 18 and 2 cores for DFT and LPM, respectively – an order of magnitude more expensive than LPM.
therefore, the practical way of using LPM in future studies would be refining the samples by SM and subsequently using the LPM before the final step of DFT calculations.
Chapter 5

Conclusions and Future Work

5.1 Conclusions and summary of completed work

In this thesis, we have presented development of a physically transparent and effective model for simulating sp-lone pair cations in material systems starting from the simplest isolated lone pair cation and scaling up to nanoparticles consisting of multiple interacting lone pair cations and anions. The validity and usefulness of the model were demonstrated on nanocluster structures of PbO and PbF$_2$ compounds that have been thoroughly verified against DFT data collected in this work, as discussed in chapter 4.

To provide the full description required for the systems of interest, it was necessary to develop a model for interactions that occur between one lone pair cation and other lone pair cations, as well as other species, including non-lone-pair cations and anions. As discussed in the introduction, see section 1.2 and 1.3, we have chosen the ionic picture within our atomistic approach, which underpines the method of interatomic potentials, to formulate and build the lone pair model (LPM) \[57,175\]. Accordingly, non-lone-pair cations and anions in our work are described using standard
rigid ion and polarisable shell models. Hence, our development work focused on the behaviour of a lone electron pair, described using a single atomic orbital, in an embedding potential due to environment. Thus, our material model has combined a quantum mechanical representation of the lone pair and a molecular-mechanical account for interactions between the non-lone pair ions.

As the model’s single atomic orbital, or the QM part, is localised on each model cation, interactions between lone pairs leading to mutual polarisation prove to be impossible to solve analytically. In a practical implementation of the model for systems that contain interacting lone pair cations, we find two model’s distinctive features: (i) calculation of the energy and wavefunctions requires a self-consistent field procedure, whereas (ii) for calculating model derivatives a fixed-point iterative method proves to be necessary. These model features have analogues in the ab initio methods of electronic structure calculations, e.g., absolutely localized molecular orbital (ALMO) [83,84,100,102], as well as semi-empirical methods, e.g., multistate empirical valence bond (MS-EVB) [88,89], see section 2.4.1 and 2.4.4. Therefore, we have developed the model adapted algorithms, which were tested on sample systems to demonstrate their effectiveness. As the model atomic orbital, we chose a linear combination of the minimum basis set consisting of one $s$ and three $p$ orbitals, taken from FHI-AIMS DFT code [71]. We have investigated the electronic configuration of a lone pair cation using DFT, see section 3.2, and verified that the polarisation of a lone pair proceeds by the onsite hybridisation of the occupied $s^2$ and empty $p^0$ states.

The model behaviour is determined by two key parameters, $\omega$ and $\lambda$, representing the effective charge of the model electron lone pair and onsite energy separation, respectively. In particular, the onsite parameter, $\lambda$, was introduced to absorb the complications of interactions, accounting for the kinetic energy, interaction with the nucleus and core electrons, and electron-electron repulsion within the model valence state. To determine suitable values for model parameters of Sn$^{2+}$, Pb$^{2+}$
and Bi$^{3+}$ cations, we have mapped their parameter space onto the DFT calculated data, see section 3.1.3 and table 3.1, results of which will be useful for future applications of the model. Furthermore, by comparison with the Hartree method, in section 3.2.1, we have shown that the model lone pair charge, $\omega$, should deviate from its nominal value of -2, accounting for the electron-electron repulsion within the valence state. These parameters are also directly related with the dipole moment and dipole-polarisability of a model cation. We derived their expressions as a function of static electric field, in section 3.1. The model dipole moment and polarisability were also compared with those predicted by the conventional harmonic and anharmonic shell models. Within the domain of electric field values experienced by ions in material systems, our atomic orbital approach has proved to provide a more realistic control over ion’s polarisability.

The core part of a model cation was modelled as a rigid ion to represent the nucleus and core electrons of a lone pair cation. The treatment was validated using DFT, which showed that the inner electrons, particularly fully occupied $d$ orbitals with energy levels right below the valence state of $s^2p^6$ character, are relatively inert to the polarisation of the lone pair, see section 3.2.1. Consequently, we modelled the interaction between the rigid core and anions using the Buckingham potential, which was key to the reproduction of relevant cation-anion equilibrium distances, see section 3.3.2. Since the ionic core carries the model charge $Z$, which has a positive value, without such a repulsive short-range interatomic potential, the anion would collapse onto the core of the lone pair cation.

As discussed earlier at the beginning of this chapter, the current version of LPM was able to successfully reproduce the structures of PbO and PbF$_2$ nanoclusters from DFT calculations, see table 4.1-4.4. A possible limitation of the model was also revealed in this work concerning prediction of “correct” energy ranking (referenced to DFT results) of nanoclusters in different
configurations. However, the absolute ranking of local minima on the energy landscape is also known as a challenging task for electronic structure methods [174], and we provided a benchmark with examples of \((\text{PbF}_2)_6\) nanoclusters, summarised in table 4.5, resulting in different energy rankings when using different DFT exchange-correlation functionals. Furthermore, the choice of a fully ionic description in LPM, where model cations and anions carry formal charges, also caused an issue, especially, with reproducing accurate dipole moments of nanoclusters. The dipole magnitudes predicted by the model were roughly three times greater than those found from DFT calculations, see section 4.3.3 and 4.4.4. We have attributed this inconsistency to the charge redistribution, or delocalisation effect, which is absent in the LPM but unavoidable in DFT and other electronic structure approaches where there are evidence of lone pair cations changing their charges dynamically depending upon coordination environment [176–178]. We also verified this effect on the example of \(\text{PbF}^+\) molecule by DFT calculations, see section 3.3.2. Nevertheless, LPM was able to capture qualitatively the right trend in dipole moments, which was made possible by reproducing DFT nanocluster structures. One viable method to mitigate this problem is by employing partial charges on cations and anions, emulating the charge delocalisation effect, as seen in particular in semi-empirical models [179,180]. Adopting such an approach, however, will bring us a new substantial challenge of investigating the mechanism of charge redistribution across the system, and therefore, we chose not to pursue it.

The entire work and, in particular, that reported in the application chapter above, was carried out based on the lone pair model developed here and implemented in the SLAM software as part of this PhD project, for its description see Appendix E. The SLAM software calculates the total energy and atomic forces on a cluster. The code can also: (i) perform local atomic relaxations to an LPM energy minimum (employing the backtracking line-search algorithm, see section 4.2.2),
and includes a separate toolkit, developed to enable calculation of the dipole moment and vibrational frequencies of a cluster; (ii) compare two atomic cluster configurations (using the similarity index that was also designed as part of the thesis, see Appendix B); and (iii) perform a search for nanocluster lowest local minima (LM) on the energy landscape defined by a rigid ion model (RIM), shell model (SM), LPM, and, significantly, DFT (the latter via calls to a third-party code FHI-AIMS), based on a simple global optimisation algorithm, Monte Carlo without temperature bias.

Using the current implementation of LPM, we have also carried out similar investigations in other compounds, including different species of lone pair cations, including Sn$^{2+}$ and Bi$^{3+}$. In the investigation, employing the global optimisation scheme described in section 4.2.3, promising results have been obtained, with a good match between LPM and DFT structures of SnO nanoclusters. Moreover, for Bi$_2$O$_3$, the atomic configurations of nanoclusters predicted by LPM matched those reported in the literature characterised by ab initio molecular dynamics [181]. The data for these compounds is still under review and require further analysis, and therefore were not included in this thesis. We note, however, that the results are promising, and will be submitted for publication in the near future.

5.2 Suggested future work

Extension of single electron molecular orbital approach. As presented in chapter 2, the LPM employs a single electron molecular orbital, based on the assumption that the two stereo-chemically active electrons are always paired in their ground state, including $sp$-hybridisation induced by external fields. Here, the total electron charge of a lone pair should be -2. However, it turned out that using only a fraction of the charge rather than the full formal charge on the lone pair significantly improves the match of LPM and DFT data. We have traced the origin of this
inconsistency to the approximation in that the LPM does not explicitly include the electron-electron repulsion within the lone pair. Accordingly, a partial charge is employed to represent the effect implicitly, and this approach was validated in sections 3.1.3 and 3.2.1. Adding an explicit account of the electron-electron repulsion to the model is an interesting extension to pursue in the future based on a simple Hartree approximation applied to a He atom in vacuum and in external field [130].

Furthermore, a common feature of lone pair cations is their electron configuration of \(d^{10}s^2\). In the LPM, we assumed that the fully occupied \(d\) states are relatively inert and thus negligible for the stabilisation of an \(sp\)-lone pair. Using the DFT data as reference, we validated that the inclusion of \(d\) atomic orbitals would be excessive for the chosen test system and typical bonding distances, while the current LPM basis set of single \(s\) and \(p\) type shells of atomic orbitals is sufficient to represent the effective valence state of lone pair cations where the details can be found in section 3.2.2. It is however another possible extension to include the \(d\) function in the basis set which would allow to explore not only the quadrupolar polarisation effects for the lone pair cations but also other chemical elements, where the outer core \(d\) states might play an important role in the ionic bonding, e.g., for Cu\(^+\) and Ag\(^+\) cations.

**Dynamic size effect on lone pairs.** While keeping the general framework of LPM, another possible variation would be introducing in the basis set an additional atomic \(s\) orbital. In brief, the revised version of LPM molecular orbital will take the form:

\[
|\psi\rangle = c_{s'} |s'\rangle + c_s |s\rangle + c_{p_x} |p_x\rangle + c_{p_y} |p_y\rangle + c_{p_z} |p_z\rangle
\] (5.2.1)
where the newly added atomic orbital is \( s' \) with its molecular orbital coefficient \( c_{s'} \). Considering the onsite nature of LPM addition of just one orbital to the basis set would only mildly increase the computational cost of the model, i.e., the size of model Hamiltonian matrix will be increased from \( 4 \times 4 \) to \( 5 \times 5 \).

The expected role of the \( s' \) atomic orbital would be in providing a dynamic size effect on the lone pair. To introduce the effect, the \( s' \) orbital must at least have a different effective radius from the original \( s \) orbital and will require a further investigation to make an appropriate choice. The aim will be to allow the lone pair to change its size adaptively to the local atomic environment by adjusting the relative orbital contributions between \( s' \) and \( s \) states, which could be achieved by introducing a second energy separation parameter \( \lambda' \), separating \( s' \) and \( s \) orbitals. Tuning effective radii of the two \( s \) states and bracketing the \( p \) level between them (\( \lambda' > \lambda \)) would keep the model physically transparent and control the extent of hybridisation in the external potential provided by the environment.

The expected effect of this extension is similar to that of the breathing shell model (BSM) [139] or the compressible ion model (CIM) [182,183], which are also able to adaptively change the effective radius of an ion depending upon the local environment. To investigate the variable ion size effect, the LPM can be compared with the BSM and CIM as well as DFT for its validation.

**Extension to periodic system.** In this thesis, we have described our work on development of the lone pair model for nanoclusters. As discussed in section 1.4, nanoclusters are systems with many degrees of freedom, huge numbers of possible atomic configurations where atoms appear in diverse local environments including harsh conditions, for example, in the form of strong electric fields at terminating surface sites [119–121]. Hence, the choice of nanoclusters as our first
application field has been dictated to a large extent by the desire to test the method to the maximum and compare with results of higher-level theories such as DFT.

Since the LPM has proved its potential, the next promising extension would be to implement a periodic version of LPM, in order to simulate periodic solids including crystals and amorphous systems modelled using a supercell approach. With the periodic implementation, the LPM can be applied to a number of topical materials, for instance, simulating binary SnO and PbO materials, including lone pair induced distortions. Furthermore, the model can be employed in investigations of halide perovskites with lone pair cations and the lone pair induced distortions [184–189], which recently have been attracting huge attention in the materials community, but there is still a lack of appropriate models in atomistic level [190]. Any future work will, however, require new careful parameterisation of both LPM and appropriate interatomic potentials.

To realise this extension, the LPM requires a further development to work with conditionally convergent series of Coulomb terms, which is a common problem to all atomistic methods including electrostatic interactions. A popular approach to tackle this problem is due to Ewald who proposed to split the series into two absolutely and rapidly convergent series, one in real and another in reciprocal space [191,192]. Implementation of respective terms and summations will still require more work and proceed in two distinct interactions series: (i) lone pair cations in the unit cell interacting with non-lone pair species within the unit cell and its periodic images, and (ii) lone pair cations in the unit cell interacting with lone pair cations within the unit cell and its periodic images.
Appendix A. Dipole-quadrupole and quadrupole-quadrupole polarisabilities of the lone pair model

Continuing from the main passage in page 94, the time-independent perturbation theory can be alternatively employed to investigate the model quadrupole related polarisabilities, $A_{ijk}$ and $C_{ijkl}$. We have already given general perturbation theory energy corrections, $E^{(1)-(4)}$, in equations, 3.1.13 – 3.1.16, which on substitution of relevant LPM energy transitions and the nonvanishing volume integrals from equations, 3.1.5 and 3.1.27, simplify to\(^{46}\):

\[
E^{(0)} = 0, \tag{A.1}
\]

\[
E^{(1)} = -\omega v(F_{11} + F_{22} + F_{33}), \tag{A.2}
\]

\[
E^{(2)} = -\frac{\omega^2 u^2}{\lambda}(F_1^2 + F_2^2 + F_3^2), \tag{A.3}
\]

\[
E^{(3)} = -\frac{\omega^3 u^2}{\lambda^2} \left[ w(F_1^2(9F_{11} + 6F_{22} + 6F_{33}) + F_1F_2(F_{12} + F_{21}) + F_1F_3(F_{13} + F_{31}) ight. \\
+ F_2F_1(F_{21} + F_{12}) + F_2^2(F_{11} + 3F_{22} + F_{33}) + F_2F_3(F_{23} + F_{32}) \\
+ F_3F_1(F_{31} + F_{13}) + F_3F_2(F_{32} + F_{23}) + F_3^2(F_{11} + F_{22} + 3F_{33}) ] \right] \\
+ v((F_{11} + F_{22} + F_{33})(F_1^2 + F_2^2 + F_3^2), \tag{A.4}
\]

and

\[^{46}\text{The higher order corrections, } E^{(3)} \text{ and } E^{(4)} \text{ are required to investigate the polarisabilities, } A_{ijk} \text{ and } C_{ijkl}, \text{ where the terms of interest include the terms proportional to or, more generally, dependent on products of the field and field gradient, i.e. the paired products } \{F_k, F_{ij}\} \text{ and } \{F_{ij}, F_{kl}\}, \text{ see equations 3.1.15 and 3.1.16.} \]
\[
E^{(4)} = -\frac{\omega^4 u^2}{\lambda^3} \left[ 2\nu(F_{11} + F_{22} + F_{33}) \left\{ -\frac{\lambda^2}{\omega^3 u^2} E^{(3)} \right\} \right. \\
- w^2 \left( F_1^2(F_{12} + F_{21}) + F_1^2(F_{13} + F_{31}) + F_1 F_3(F_{12} + F_{21})(F_{23} + F_{32}) \right) \\
+ F_1 F_2(F_{13} + F_{31})(F_{23} + F_{32}) + F_1 F_2(F_{12} + F_{21})(3F_{11} + F_{22} + F_{33}) \\
+ F_1 F_3(F_{13} + F_{31})(3F_{11} + F_{22} + F_{33}) + F_1^2(3F_{11} + F_{22} + F_{33})^2 \\
+ F_1 F_2(F_{12} + F_{21})(F_{11} + 3F_{22} + F_{33}) + F_1 F_3(F_{13} + F_{31})(F_{11} + F_{22} + 3F_{33}) \\
+ F_2(F_{12} + F_{21})(F_2(F_{12} + F_{21}) + F_3(F_{13} + F_{31}) + F_1(3F_{11} + F_{22} + F_{33})) \\
+ F_3(F_{13} + F_{31})(F_2(F_{12} + F_{21}) + F_3(F_{13} + F_{31}) + F_1(3F_{11} + F_{22} + F_{33})) \\
+ F_3(F_{23} + F_{32})(F_1(F_{12} + F_{21}) + F_3(F_{23} + F_{32}) + F_2(F_{11} + 3F_{22} + F_{33})) \\
+ F_2(F_{11} + 3F_{22} + F_{33})(F_1(F_{12} + F_{21}) + F_3(F_{23} + F_{32}) \\
+ F_2(F_{11} + 3F_{22} + F_{33})) \\
+ F_2(F_{23} + F_{32})(F_1(F_{13} + F_{31}) + F_2(F_{23} + F_{32}) + F_3(F_{11} + F_{22} + 3F_{33})) \\
+ F_3(F_{11} + F_{22} + 3F_{33})(F_1(F_{13} + F_{31}) + F_2(F_{23} + F_{32}) \\
+ F_3(F_{11} + F_{22} + 3F_{33})) \right] - v^2(F_1^2 + 2F_3^2)(F_{11} + F_{22} + F_{33})^2 \\
+ u^2(F_1^2 + 2F_3^2)(F_1^2 + F_2^2 + F_3^2) \\
\] 

(A.5)

Using the energy corrections, from equations A.1 to A.5, the LPM quadrupole-polarisabilities at the zero fields are:

\[
A_{kij}(F_k = 0, F_{ij} = 0) = \frac{\partial}{\partial F_k} \left( -3 \frac{\partial E}{\partial F_{ij}} + \delta_{ij} \sum_k \frac{\partial E}{\partial F_{kk}} \right) \bigg|_{F_k=0,F_{ij}=0} = 0, \quad (A.6)
\]

\[
C_{ijkl}(F_{ij} = 0, F_{kl} = 0) = \frac{\partial}{\partial F_{kl}} \left( -3 \frac{\partial E}{\partial F_{ij}} + \delta_{ij} \sum_k \frac{\partial E}{\partial F_{kk}} \right) \bigg|_{F_{ij}=0,F_{kl}=0} = 0 \quad (A.7)
\]
where the energy is truncated at the 4th order, \( E = E^{(0)} + E^{(1)} + E^{(2)} + E^{(3)} + E^{(4)} \) and the derivatives in equations A.6 and A.7 have been verified using the Mathematica software package [75].
Appendix B. Comparing atomic configurations of two different nanoclusters

To carry out quantitative comparison of different atomic configurations of nanoclusters, we have introduced a concept of similarity index (SI), which quantifies the structural similarity between two different atomic configurations. Here, we are going to deliver the background how the quantification works and introduce the software we have implemented.

To quantify structural differences between two clusters A and B of the same chemical composition, number of atoms and atom order, as illustrated by figure B.1(a), one can measure the root-mean-square-deviation (RMSD) between their atomic coordinates. One clear possible issue is that it is hard to justify the reliability of directly measured RMSD value, since those clusters are likely to be arbitrarily positioned and oriented in space, e.g., if they were relaxed on different energy landscapes or starting from different initial points.

To provide more reliable RMSD measures, we applied appropriate translation that brings the centre of masses of those clusters in coincidence at the origin and rotation that aligns axes of their principal moments of inertia.

**Figure B.1** Structure comparison: (a) cluster A and B are arbitrarily placed in space. (b) after translation (their respective centre of mass is at the origin) and rotation (axes of their principal moments of inertia aligned), the atomic configurations of clusters overlap each other.
For a cluster of $N$ atoms, after translation to the centre of mass, the inertia tensor is calculated as:

$$
\mathbf{I} = \begin{pmatrix}
I_{11} & I_{12} & I_{13} \\
I_{21} & I_{22} & I_{23} \\
I_{31} & I_{32} & I_{33}
\end{pmatrix}
$$

(B.1)

where

$$
I_{ij} = \sum_{k}^{N} m_{k} (|r_{k}|^2 \delta_{ij} - r_{k;i}r_{k;j}) .
$$

(B.2)

In equation B.2, $r_{k}$ and $m_{k}$ are the position vector and mass of atom $k$, respectively; $r_{k;i}$ refers to Cartesian coordinates, $x, y, z$ for $i = 1, 2, 3$, of $r_{k}$; and $\delta_{ij}$ is the Kronecker delta. In our implementation, since we are only interested in the atomic configuration of a cluster, we set all atomic masses equal to one.

Diagonalisation of the inertia tensor yields a rotation transformation matrix $\mathbf{C}$ bringing the tensor to its diagonal form:

$$
\mathbf{P} = \mathbf{C}^{T}\mathbf{IC}
$$

(B.3)

where the diagonal elements of matrix $\mathbf{P}$ are the principal moments of inertia.

The transformation can be applied to atomic coordinates of a cluster by:

$$
r'_{k;i} = \sum_{j=1}^{3} C_{ij}r_{k;j}
$$

(B.4)

where $C_{ij}$ is the element of matrix $\mathbf{C}$, and transformation B.4 should be applied to all atoms.
Similarity Index. If two clusters have similar atomic configurations, one can expect that bringing both to their centres of mass followed by aligning their principal axes of inertia will result in an overlap of the two structures, see figure B.1 (b). Next, the RMSD, or the “similarity-index” in main text, can be calculated resulting in a reliable useful measure. The algorithm utilised in our software for comparing such two atomic configurations and calculation of their SI value is summarised below.

Algorithm. Measuring RMSD value of atomic coordinates of two nanoclusters, including translation (their respective centres of mass are at the origin) and rotation (axes of their principal moments of inertia aligned).

**load** nanocluster structures of A and B, whose atomic coordinates are $r^A_k$ and $r^B_k$, respectively.
where $k = 1, 2, \ldots, N$

1. **calculate** each center of coordinates (COC), $r^A_{COC}$ and $r^B_{COC}$, and **translate** structures
   
   $r^A_{COC} = 1/N \sum_k r^A_k$ and $r^B_{COC} = 1/N \sum_k r^B_k$;

   $r^A_k = (r^A_k - r^A_{COC})$ and $r^B_k = (r^B_k - r^B_{COC})$; (translation operation)

2. **calculate** each inertia tensor, $I^A$ and $I^B$, and **diagonalise** to find transformations, $C^A$ and $C^B$

3. **rotate** each structure using $C^A$ and $C^B$
   
   $r^A_k = C^A r^A_k$ and $r^B_k = C^B r^B_k$; for all $k$;

4. **calculate** RMSD value $\Delta_{RMSD}$ for transformed atomic coordinates
   
   $\Delta_{RMSD} = \sqrt{1/(3N) \sum_k \|r^A_k - r^B_k\|^2}$;

**return** $\Delta_{RMSD}$; (or the so-called “SI value” in main text whose unit is Å).

In the algorithm, there is possibility of degeneracy that occurs in principal moments, $I^A$ and $I^B$, in step 2. The software deals with this issue by brute forcing, adding ghost atoms at the extended position of randomly picked atoms (trying one and two for two- and three-fold degeneracy), i.e.,

$r^\text{ghost}_k = \sigma r_k$ where $\sigma$ is an internal parameter set to 7.5, until degeneracy breaks, and when it is removed, calculating the SI value.
The software is readily available to download and use where the source code can be found at: https://github.com/sweetmixture/CF_CLUSTERSpy.git, where the basic instruction manual and a few tutorial examples are included.
Appendix C. Parameters used in the study of lead monoxide nanoclusters

Table C.1 Buckingham potential and the shell model parameters employed for simulating \((\text{PbO})_N\) nanoclusters.

<table>
<thead>
<tr>
<th>Interactions</th>
<th>(D / \text{eV})</th>
<th>(\rho / \text{Å})</th>
<th>(C / \text{eV \ Å}^6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{Pb}<em>{\text{shell}} - \text{Pb}</em>{\text{shell}})</td>
<td>24.98</td>
<td>0.3284</td>
<td></td>
</tr>
<tr>
<td>(\text{Pb}<em>{\text{shell}} - \text{O}</em>{\text{shell}})</td>
<td>72276.42</td>
<td>0.2223</td>
<td></td>
</tr>
<tr>
<td>(\text{O}<em>{\text{shell}} - \text{O}</em>{\text{shell}})</td>
<td>9547.96</td>
<td>0.2192</td>
<td>32.0</td>
</tr>
</tbody>
</table>

Table C.2 Parameters employed for LPM calculations for simulating \((\text{PbO})_N\) nanoclusters.

<table>
<thead>
<tr>
<th>Interactions</th>
<th>(D / \text{eV})</th>
<th>(\rho / \text{Å})</th>
<th>(C / \text{eV \ Å}^6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{Pb}<em>{\text{LP}} - \text{Pb}</em>{\text{LP}})</td>
<td>422.</td>
<td>0.098</td>
<td></td>
</tr>
<tr>
<td>(\text{Pb}<em>{\text{LP}} - \text{O}</em>{\text{shell}})</td>
<td>122.</td>
<td>0.089</td>
<td></td>
</tr>
<tr>
<td>(\text{Pb}<em>{\text{core}} - \text{O}</em>{\text{shell}})</td>
<td>19600.</td>
<td>0.2430</td>
<td>20.4</td>
</tr>
</tbody>
</table>

| Lone pair charge   | \(\omega / \text{e}\) | \(\lambda / \text{eV}\) |
|--------------------|------------------|------------------|--------------------------|
| \(\text{Pb}\)     | -1.64         | 14.09            |                          |

Parameters given in table C.2 were fitted to the atomic configurations and energy rankings of low energy nanoclusters of \((\text{PbO})_{N=1-5}\), calculated by DFT using FHI-AIMS code, details of the DFT calculations can be found in section 4.2.3-(3) – and their structures is shown in figure C.1. In the table, \(\text{Pb}_{\text{LP}}\) and \(\text{Pb}_{\text{core}}\) indicates the lone pair and the core of a lone pair cation. Parameters \(\{D, \rho, C\}\) for \(\text{Pb}_{\text{LP}} - \text{O}_{\text{shell}}\) and \(\text{Pb}_{\text{LP}} - \text{Pb}_{\text{LP}}\) interactions represent respective terms in equations 2.4.4b and 2.4.6 of the main text. The \(\text{Pb}_{\text{core}} - \text{O}_{\text{shell}}\) interaction is described by the usual Buckingham potential (see equation 1.11), where the core \(\text{Pb}_{\text{core}}\), of LPM is represented by a rigid ion.

Note that for both LPM and shell model calculations the same \(\text{O}_{\text{shell}} - \text{O}_{\text{shell}}\) interatomic potential is employed with parameters given in table C.1.
Figure C.1 Structures of \((\text{PbO})_N\) clusters optimised at the DFT level using the FHI-AIMS package shown as ball-and-stick models: the black colour reserved for Pb and the red colour for O. For the global minimum clusters of each size, the polymerization energy is given as \(E_p^{GM}\) (eV) – see equation 4.3.1 – where the LPM values in parentheses, and the relative nanocluster energies (\(\Delta E_p = E_p - E_p^{GM}\)) are given for the higher energy local minima. Below the energy, the similarity index (SI) and the point group symmetry (PGS) are provided.
Appendix D. Parameters used in the study of lead fluoride nanoclusters

Table D.1 Buckingham potential and the shell model parameters employed for simulating \((\text{PbF}_2)_N\) nanoclusters.

<table>
<thead>
<tr>
<th>Interactions</th>
<th>(D) /eV</th>
<th>(\rho) /Å</th>
<th>(C) /eV Å(^6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{Pb}<em>{\text{shell}} - \text{Pb}</em>{\text{shell}})</td>
<td>24.98</td>
<td>0.3284</td>
<td></td>
</tr>
<tr>
<td>(\text{Pb}<em>{\text{shell}} - \text{F}</em>{\text{shell}})</td>
<td>743.7343</td>
<td>0.348506</td>
<td></td>
</tr>
<tr>
<td>(\text{F}<em>{\text{shell}} - \text{F}</em>{\text{shell}})</td>
<td>1127.7</td>
<td>0.2753</td>
<td>15.83</td>
</tr>
</tbody>
</table>

Shell model

<table>
<thead>
<tr>
<th>Interactions</th>
<th>(q) /e</th>
<th>(k) /eV Å(^{-2})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pb</td>
<td>-4</td>
<td>172.7</td>
</tr>
<tr>
<td>F</td>
<td>-1.59</td>
<td>20.77</td>
</tr>
</tbody>
</table>

Table D.2 Parameters employed for LPM calculations for simulating \((\text{PbF}_2)_N\) nanoclusters.

<table>
<thead>
<tr>
<th>Interactions</th>
<th>(D) /eV</th>
<th>(\rho) /Å</th>
<th>(C) /eV Å(^6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{Pb}<em>{\text{LP}} - \text{Pb}</em>{\text{LP}})</td>
<td>422.</td>
<td>0.098</td>
<td></td>
</tr>
<tr>
<td>(\text{Pb}<em>{\text{LP}} - \text{F}</em>{\text{shell}})</td>
<td>122.</td>
<td>0.089</td>
<td></td>
</tr>
<tr>
<td>(\text{Pb}<em>{\text{core}} - \text{F}</em>{\text{shell}})</td>
<td>4494.</td>
<td>0.264</td>
<td>20.4</td>
</tr>
</tbody>
</table>

Lone pair charge

<table>
<thead>
<tr>
<th>Interactions</th>
<th>(\omega) /e</th>
<th>(\lambda) /eV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pb</td>
<td>-1.64</td>
<td>14.09</td>
</tr>
</tbody>
</table>

Parameters given in table D.2 were fitted to the atomic configurations and energy rankings of low energy nanoclusters of \((\text{PbF}_2)_{N=1-2}\), calculated by DFT using FHI-AIMS code, details of calculations can be found in section 4.2.3-(3), and their structures are shown in figure D.1. In the table, \(\text{Pb}_{\text{LP}}\) and \(\text{Pb}_{\text{core}}\) indicate the lone pair and the core of a lone pair cation. Parameters \(\{D, \rho, C\}\) for \(\text{Pb}_{\text{LP}} - \text{F}_{\text{shell}}\) and \(\text{Pb}_{\text{LP}} - \text{Pb}_{\text{LP}}\) represent respective terms in equations 2.4.4b and 2.4.6 of the main text. The \(\text{Pb}_{\text{core}} - \text{F}_{\text{shell}}\) interaction is described with the usual Buckingham potential (see equation 1.11), where the core, \(\text{Pb}_{\text{core}}\), of LPM is represented by a rigid ion.

Note that for both LPM and shell model calculations the same \(\text{F}_{\text{shell}} - \text{F}_{\text{shell}}\) interatomic potential is employed with parameters given in table D.1. Also, the same LPM lead (II) cation is used as in table C.2 (Appendix C) in order to testify the transferability of LPM.
Figure D.1 Structures of \((\text{PbF}_2)_{N=1-2}\) clusters optimised at the DFT level of theory using the FHI-AIMS package, shown as a ball-and-stick model: the black colour is reserved for Pb and the white colour is reserved for F. For the global minimum clusters of each size, the polymerization energy is given as \(E_p^{GM}\) – see equation 4.3.1 – where the LPM values shown in parentheses, and the relative nanocluster energies (\(\Delta E_p = E_p - E_p^{GM}\)) are given for the higher energy local minima. Below the energy, the similarity index (SI) and the point group symmetry (PGS) are provided.
Appendix E. Software for simulating the lone pair model

The in-house software named *Sp-Lone pair integrated Atomistic Model*, or SLAM, was specifically created for this thesis and contains the implementation of the developed lone pair model (LPM). The programming language used to write SLAM is C, and, at the time of writing this thesis, contains 10K lines of C (excluding required pre-calculated integral library). SLAM was used to produce the LPM data presented in the main text. The source code for the SLAM software can be downloaded at

https://github.com/sweetmixture/SLAM.2.2.4.git.

E.1 Compilation

Within the downloaded /src directory, we provide the Makefile for the compilation of SLAM. The bulleted list below explains the required environment to build the executable binary.

- The compilation of source code requires a C compiler, for example, either the GNU C compiler (https://gcc.gnu.org) or Intel OneAPI will work.

- For the implementation of LPM, we employ BLAS to perform matrix/vector operations, thus BLAS support within the GNU Scientific Library (GSL) is required; see details at https://www.gnu.org/software/gsl.

- For calculating interactions involving lone pairs, which is the biggest bottleneck due to the required large number of integrations, we employ MPI libraries to enable parallel execution; see details regarding MPI at https://www.open-mpi.org.
E.2 Running SLAM

Once SLAM is compiled, we recommend running the simple test example input provided in directory /PbF2_example_1 to familiarise yourself with the basic structure of the input and output files. This input will instruct SLAM to relax the defined \((\text{PbF}_2)_5\) nanocluster to the LPM LM. SLAM will output the relaxed structure, LM energy, and the molecular orbital information of \(sp\)-lone pair cations.

To run SLAM type

```bash
$ mpirun -np 2 /PATH/TO/EXECUTABLE out.xyz 1000
```

Here, two CPU cores are requested for the executable binary located within /PATH/TO. The mandatory arguments are (i) `out.xyz` which is the output file of the final atomic configuration in the standard xyz format, and (ii) `1000` which is the maximum number of energy relaxation steps (i.e. total number of line-searches). Note that “out” and “1000” are chosen by the user and that the general SLAM output will be sent to STDOUT, i.e. printed to the screen.

If you want to visualise lone pair electron densities created by the LPM, the add the argument “out.cube” to the execution line and SLAM will generate a standard Gaussian cube file (see https://gaussian.com/cubegen). For example,

```bash
$ mpirun -np 2 /PATH/TO/EXECUTABLE out.xyz 1000 out.cube > out.txt
```

will create `out.cube` before SLAM finishes and, as the command line includes the hairpin, the general SLAM output will be saved as `out.txt`. 
E.3 Input files

SLAM requires several input files, as summarised in the file, /PbF2_example_1/Example-Overview. Note that the name of input files must match exactly that given below as they are constant environment variables.

In the directory /PbF2_example_1/sp_cluster_prameter_src we have provided three regular text files: sp_cluster_knot.txt, sp_cluster_radial_s.txt, sp_cluster_radial_p.txt. These define the radial wavefunctions of the $s$ and $p$ atomic orbitals employed in the LPM, which are splines created from data generated by the DFT code FHI-AIMS. The radial wavefunctions are species dependent and thus the files should be set by which type of lone pair cations is used. We provide radial wavefunctions for Sn$^{2+}$, Pb$^{2+}$ and Bi$^{3+}$ within the directory /species. Note that you can customise these functions, which are normalised, continuous splines that represent the chosen atomic orbitals.

The initial atomic configurations of the nanocluster are defined within the text file geo.txt; see example provided in directory /PbF2_example_1, which begins with the lines

```
#comment
20 5
F  c    2.341841    2.092431    2.542214
F  s    2.341841    2.092431    2.542214
F  c    0.606138    2.515924    4.641656
F  s    0.606138    2.515924    4.641656
```

and ends with the lines

```
Pb      2.123537    3.290852    0.550270
Pb      3.618500    0.274032    1.763194
Pb      4.942059    3.459499    3.239038
Pb      0.129201    1.323455    2.777418
Pb      1.561445    4.477080    4.221179
```
The first line is a reserved space for the user to leave a comment, and the second line specifies number of non-lone pair species, including cores and shells (in this example we use the shell model to represent the ten F\(^-\) ions), and the number of lone pair cations. Each of the following lines begin with name of species and, for non-lone pair species, whether a core or shell by using the character c and s, respectively, and then three numbers for the cartesian atomic coordinate \(\{x, y, z\}\).

The LPM potential parameters are defined within the text file sp_cluster_species.txt; see example provided in directory /PbF2_example_1, which has the following lines

\[
\begin{align*}
\text{QM_POT} & \quad 2 \\
F & \quad 122. \quad 0.089 \\
Pb & \quad 422. \quad 0.098 \\
\text{ATOM_TYPE} & \quad 2 \\
\text{SHELL} & \quad F \quad 0.59 \quad -1.59 \quad 20.77 \quad 0. \\
\text{SP} & \quad Pb \quad 3.64 \quad -1.64 \quad 14.09 \\
\text{MM_POT} & \quad 2 \\
\text{SHELL} & \quad F \quad \text{SHELL} \quad F \quad 1127.7 \quad 0.2753 \quad -15.83 \\
\text{SP} & \quad Pb \quad \text{SHELL} \quad F \quad 4494. \quad 0.264 \quad -20.4 \\
\text{gnorm_tol} & \quad 0.00005
\end{align*}
\]

On the first line, \texttt{QM_POT 2} indicates there are two types of interactions for the lone pairs. In this example, we have (i) the lone pair electrons interacting with the shell of F\(^-\) and (ii) the lone pair electrons interacting with lone pairs on other Pb\(^{2+}\) cations. Each require two potential parameters, \(\{D, \rho\}\), and are defined using the format

\[
\begin{align*}
F & \quad <D> \quad<Rho> \\
Pb & \quad <D> \quad<Rho>
\end{align*}
\]

The expressions for these interactions are given in equations 2.4.4b and 2.4.6 (within main text).

On the fourth line, \texttt{ATOM_TYPE 2}, indicates data will follow on the next two lines, one per species used in this calculation.
On the fifth and sixth lines, SHELL F and SP Pb, indicates that F and Pb are represented by the shell model and LPM, respectively, with the interatomic potential parameters given in the following the format

```
SHELL F <core_charge> <shell_charge> <k2> <k4>
SP Pb <core_charge> <Lone_pair_charge> <LPM_Lambda>
```

The expressions for the information of species: (i) charges of core and shell, and its spring constants, k2 (harmonic) and k4 (anharmonic – if it does not required, set to zero), and (ii) charges of lone pair core and lone pair electrons, and the onsite potential parameter. On can make match with the data given in table D.1 and table D.2 in Appendix D.

On the seventh line, MM_POT 2, indicates that details of two Buckingham potentials follows using the following format

```
SHELL F SHELL F <D> <Rho> <C>
SP Pb SHELL F <D> <Rho> <C>
```

The expressions for these interactions are given in equations 1.5. (within main text). In the example above, interatomic potential parameters are provided for interactions between pairs of F shells, and between F shell and the core part of the Pb lone pair cation (specified by SP). Note that the sign for the dispersion term <C> is inverted in our implementation.

On the last line, gnorm_tol 0.00005, specifies the required tolerance (in eV/Å unit) of the gnorm, i.e. the average component of the atomic forces.

Note that the numbers in sp_cluster_species.txt are taken from the tables in Appendix C.
E.3 Output files and utilities for further analysis

When SLAM completes it generates the file \texttt{geo.txt.next}, which includes the atomic configurations at the final step of energy relaxation, along with \texttt{out.xyz}, \texttt{out.cube} (if requested) and \texttt{out.txt} (if STDOUT is redirect to this file).

Within the \texttt{out.txt} file, the molecular orbital information of the lone pair cations are written out. For the provided example, you will find the following lines

Lone Pair Molecular Orbital ( Lowest Eigenvalue / EigenVector )

<table>
<thead>
<tr>
<th>Species</th>
<th>Energy(eV)</th>
<th>s</th>
<th>px</th>
<th>py</th>
<th>pz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pb</td>
<td>9.804472</td>
<td>0.972770</td>
<td>-0.031725</td>
<td>-0.204770</td>
<td>0.103835</td>
</tr>
<tr>
<td>Pb</td>
<td>10.578634</td>
<td>0.980287</td>
<td>-0.135424</td>
<td>0.140764</td>
<td>0.029724</td>
</tr>
<tr>
<td>Pb</td>
<td>9.873901</td>
<td>0.969382</td>
<td>-0.203052</td>
<td>-0.136218</td>
<td>-0.022638</td>
</tr>
<tr>
<td>Pb</td>
<td>10.577532</td>
<td>0.980288</td>
<td>0.181832</td>
<td>0.045382</td>
<td>-0.062554</td>
</tr>
<tr>
<td>Pb</td>
<td>9.875376</td>
<td>0.969394</td>
<td>0.061448</td>
<td>-0.215803</td>
<td>-0.099647</td>
</tr>
</tbody>
</table>

The data in the columns headed \texttt{Energy(eV)}, \texttt{s}, \texttt{px}, \texttt{py}, and \texttt{pz}, are the ground state energy and the molecular orbital coefficients of each cation, respectively.

E.3.1 Dipole moment calculator

The molecular orbital coefficients are essential to calculate the dipole moment of LPM nanocluster. We provide a supportive utility, which parses the standard output of LPM and calculates the dipole moment of nanocluster. The utility is implemented as a small program, written in Shell/Python, and it is available to download at the link: \url{https://github.com/sweetmixture/slam_dipole_calculator.git}. 
E.3.2 Normal mode analysis

The SLAM standard output file, out.txt, also provides the geometric derivatives (or first derivatives of energy) of the nanocluster after it has been relaxed to a LM configuration. For example, the text would look like:

Geometric Derivatives ( eV / Angstrom )

| Species | x       | y       | z       | |r|       |
|---------|---------|---------|---------|---------|
| Fc      | 1.296164e-04 | 2.833948e-04 | -2.265939e-05 | 3.124523e-04 |
| Fs      | 2.189214e-04 | 5.557974e-04 | -4.106510e-05 | 5.987685e-04 |
| Fc      | 1.918109e-04 | 5.481092e-04 | -4.106510e-05 | 5.816866e-04 |
| Fs      | 3.892978e-04 | 8.895849e-04 | 1.000890e-03  | 9.846485e-04 |
| Fc      | -1.955227e-05| 8.277019e-04 | -1.524218e-04 | 8.418462e-04 |
| Fs      | -2.240928e-05| 7.439877e-04 | -1.481247e-04 | 7.589208e-04 |
| Fc      | 6.090245e-06 | 2.193616e-04 | -2.950600e-04 | 3.677186e-04 |
| Fs      | -4.616535e-04| -6.829850e-04 | -6.34966e-04  | 1.040278e-03 |
| Fs      | 1.970083e-04 | 3.105515e-04 | -4.555008e-04 | 5.854361e-04 |
| Pb      | 4.427975e-05 | 1.385741e-04 | 3.979148e-06  | 1.455311e-04 |

With such information obtained from a sequence single point calculations, numerical second derivatives of energy can be calculated using the finite difference method,

$$\frac{\partial^2 E}{\partial x_i \partial x_j} = \frac{F_i(x_j + \delta) - F_i(x_j - \delta)}{2\delta}, \quad (E.1)$$

where $\delta$ is a small finite displacement; $F_i$ are the components of analytically calculated first derivatives (taken from the standard SLAM output) and $x_i$ are atomic coordinates.
The second derivatives of energy are required to perform normal mode analysis to verify whether a relaxed nanocluster is a true LM (and not a saddle or maximum point on the energy landscape). We provide a script that automates the task of calculating the numerical second derivatives and calls a small Python program, which computes the normal mode frequencies from these second derivatives of energy (algorithm by Wilson [193]). The script can be downloaded from: https://github.com/sweetmixture/slam.vibration_normal_mode.support.git.

Since the second derivatives of energy are based on the numerical calculations, it is important to avoid choosing a sufficiently large displacement $\delta$ to avoid possible numerical noises, yet small enough to avoid errors caused by anharmonicities. In practice, we found that the value of $\delta = 0.025-0.050$ Å was appropriate in general with the setting of `gnorm_tol 0.00005` when the noise on rotational and translational frequencies were less than $\pm 7$ cm$^{-1}$. It is prudent to check results by using a tighter value of `gnorm_tol`. 
References


List of Figures

Figure 1.1 Three coordination polyhedral types of Pb(II) for tetrahedral [T], triangular bipyramidal [TBP] and square pyramidal [SP] where the pale blue circle marked by “E” represents the lone pair electron density, image taken from Galy et al [1]. ................................................................. 10

Figure 1.2 Schematic diagrams of (A) holodirected and (B) hemidirected coordination. With respect to the centred cation $M$, the black circles and blue arrows refer to the coordination sphere and ligands, respectively. ................................................................. 12

Figure 2.1 Energy diagrams of the valence states for an isolated lone pair cation $M$ (e.g., $M = \text{Sn}$ or Pb) in vacuum: (a) shows two electrons occupying the $s$ state is occupied by lone pair electrons and the three $p$ states left empty after the 2nd ionisation occurs; (b) shows one electron occupying the $s$ state after the 3rd ionisation occurs where the energy difference between $s$ and $p$ states is marked with a Greek letter $\lambda$. ................................................................. 30

Figure 2.2 Radial wavefunctions of (a) $s$ and (b) $p$ atomic orbitals that are employed in the lone pair model. Colour is used to distinguish different species (Sn, Pb and Bi). The inserts show larger values of the orbitals for very short distances from the core of the ion. Note that the circles (datapoints) for each atom were sampled from the FHI-AIMS code, using the provided “Tier-2” basis set, where intervals are interpolated by cubic polynomials. ................................................................. 34

Figure 2.3 Schematic diagram of an interpolated radial wavefunction, $R(r)$. The red dots are known values of $R$ at discrete points $r = r_{(k)}$ and the black solid lines between nearest neighbour points are the functions $R_{(k)}(r)$ for the interval $[r_{(k)}, r_{(k+1)}]$. Assuming the functions $R_{(k)}(r)$ are cubic polynomials, the entire function is twice differential at their boundaries................................. 35

Figure 2.4 Schematic illustration of the distance ................................................................. 40
$|\mathbf{r} - \mathbf{R}_4|$ where the vector, $\mathbf{r}$, refers to the locations of the lone pair electrons of the model cation at the origin. ................................................................. 40

**Figure 2.5** Schematic illustration of two Cartesian reference frames, $\{x, y, z\}$ and $\{x', y', z\}$, which share the same origin. The axes for the former and latter are shown using solid and broken lines, respectively. The black and grey balls represent core of lone pair cation and a point charge, respectively. ................................................................. 44

**Figure 2.6** Schematic diagram shows two lone pair cations whose cores are at $\mathbf{r}_1$ and $\mathbf{r}_2$ where $\mathbf{r}_a$ and $\mathbf{r}_\beta$ refer to the positions of their electron lone pairs, respectively; $r_{a\beta}$ is the separation distance between lone pair electrons $r_{a2}$ and $r_{\beta1}$ indicate several distances as in the diagram. ....................... 53

**Figure 2.7** Differences in the electronic energies, $\text{ELP}_{\text{total}};(k + 1) - \text{ELP}_{\text{total}};(k)$, calculated by using the lone pair model, and the results are presented in logarithmic scale as a function of SCF iterations, $k$. The calculations were carried out for $(\text{PbO})N$ nanoclusters with $N = 2, 3, 4, 8$ and $12$, and a few of their structures are presented with the ball-and-stick model, Pb (black) and O (red) atoms. Note that the omitted bigger size nanoclusters are in similar structures where the unit of $(\text{PbO})$ are simply further concatenated as strands. ......................... 59

**Figure 2.8** Structure of a $(\text{PbO})4$ nanocluster, found by the DFT (PBEsol functional [113]) calculation using the FHI-AIMS code with “Tier-2” basis functions for the elements. The ball-and-stick model is used for Pb (black) and O (red) where the numbering labels on the atoms follow the same order as they are given in the first column of table 2.1. ................................................................. 67

**Figure 3.1** Observable properties of a lone pair cation (LPM modelled Sn2 $^+$ is used for demonstration): (a) energy, $E$, (b) dipole moment, $\mu$, and (c) dipolar polarisability, $\alpha$, as a function of uniform electric field, $F$, for three different models: harmonic shell model (SM)$^a$, anharmonic shell model $(k4)^b$ and lone pair model (LPM)$^c$ – see legend. ................................................................. 89
Figure 3.2 Contour plot heatmaps of root-mean-square deviations (RMSD) in energy $\Delta E$ and dipole moment $\Delta \mu$ between the DFT reference and LPM calculated results for the lone pair cations, Sn$^2+$, Pb$^2+$ and Bi$^3+$, over the LPM $\{\omega, \lambda\}$ parametric domain – see equation 3.1.24 for details. Note that the numerical labels on contours are marked for the first three lowest isosurfaces...... 99

Figure 3.3 (a) Atomic orbital contributions to the HOMO of Pb$^2+$ cation interacting with a point charge $q = -1$ as a function of the separation distance. Comparison is made between the PBEsol DFT $6pz$ and $5dz^2$ orbitals and the LPM (black dashed line) and its uniformly scaled LPM (red dashed line) $pz$ orbital occupancies. LPM parameters $\lambda, \omega = 16.53, 1.424$ are used. (b) $pz$ orbital occupancy ratio between DFT and LPM, calculated as a function of $(1/\omega)$ whilst using fixed $\lambda = 16.53$ eV. The red circled datapoints with $\{a, b, c\}$ correspond to $\omega \{1.640, 1.532, 1.424\}$; the leftmost and the rightmost red circles mark -2.0 and -1.0 values of $\omega$. The dashed line represents a linear fit to the given five points whose equation is shown in the plot................. 102

Figure 3.4 LPM and DFT (d.m.)/(t.m.) calculated as function of d.m. for a Pb$^2+$ cation interacting with a point charge, where the green dashed lines show linear fits and the linear equations are in the same colour scheme as the calculated data, black and red, respectively.... 109

Figure 3.5 Dipole moments (d.m.) as function of the molecular orbital coefficient $c_{pz}^{DFT}$ where its scaling to the LPM counterpart is provided in the parentheses. The red line shows the results of a DFT approximation given in equation 3.2.9a, the black and green lines show LPM results based on equation 3.2.8 with different charge values of $\omega$, distinguished by different line styles. .......................................................... 110

Figure 3.6 Energy of a Pb$^2+$ cation interacting with a point charge ($q = -1$) as a function of the separation distance, calculated by using the LPM and the DFT-PBEsol (using FHI-AIMS code); $E_{\text{Coulomb}}$ shows the Coulomb energy of two point charges, $q = +2$ (formal charge of Pb$^2+$
cation) and \( q = -1 \). Note that the total LPM energy (green line) is decomposed into, \( E_{\text{LP}} \) and \( E_{\text{core}} \), which are energy contributions by the electron lone pair and the core part of the Pb\( ^{2+} \) cation, respectively.

**Figure 3.7** Effect of the effective potential on observed properties of a lone pair. (a) Energy of the Pb(II) lone pair interacting with a point charge \( (q = -1) \) as a function of separation distance. The graphs were obtained by using different parameter \( D \) (eV) of the LPM effective potential, and they are shown with different line styles; note that parameter \( \rho = 0.09 \) Å was used throughout. (b) Corresponding dipole moment (d.m.). DFT-PBEsol (using FHI-AIMS code) calculated references are marked with red crosses.

**Figure 3.8** Eigenvalues of the lone pair electron density calculated by the LPM where the ground state \((sp\)-hybridisation\) A corresponds to the energy of \( D = 10000 \) (eV) in figure 3.7 (a); B is the eigenvalue for the degenerate \( p \) states, and C is the anti-state of A.

**Figure 3.9** Energy profiles of Pb(II) cation interacting with a fluoride anion as a function of the separation distance, calculated by the DFT-PBEsol (using FHI-AIMS code) and LPM; \( E_{\text{Coulomb}} \) shows the Coulomb energy of two point charges, \( q = +2 \) and \( q = -1 \), i.e., the formal charges of Pb\( ^{2+} \) and F\( ^{-} \). The LPM energy is decomposed into, \( E_{\text{LP}} \) and \( E_{\text{core}} \), which are the energy contributions by the lone pair and the core part of the Pb\( ^{2+} \) cation. The difference between \( E_{\text{DFT}} \) and \( E_{\text{C.DFT}} \) is that the latter energy was calculated under the constraint of prohibited charge transfer from F\( ^{-} \) to Pb\( ^{2+} \).

**Figure 3.10** Schematic diagrams and the energy plots for a lone pair cation interacting with an anion ligand: (A) including only the Coulomb interactions between charges \( Z \) of the LP core, \( \omega \) of the lone pair and \( q - \) of the anion; and (B) adding a short-range repulsive potential between the
LP core and the anion. Blue ellipses in the diagram represent the lone pair density whilst cores are shown as grey circles. ................................................................. 122

**Figure 3.11** Energy profiles of Pb2 + -Pb2 + interaction as a function of the separation distance where the insert zooms in on the region of typical Pb2 + -Pb2 + distances in the close-packed compounds including such cation. The energy profiles of the LPM were obtained with and without the effective potential (equation 2.4.6) where the green curve for Coulomb only lone pair potential; and the black curve with an addition of the effective short-range repulsive potential where the units are $D$ (eV) and $\rho$ (Å). The DFT-PBEsol calculated energy (red circles; using FHI-AIMS code) is also shown as a reference. As a secondary reference, the dashed purple curve shows the Coulomb interaction energy of two point charges, carrying $q = +2$, which is the formal charge of a lead cation. Lastly the purple circles show the energy of the LPM, calculated at the limit of $\lambda \to \infty$, which prevents the $sp$-hybridisation of the lone pairs. ................................................................. 124

**Figure 3.12** Ratio of total to Coulomb interaction energies for Pb$n$ + -Pbn +, $n = 2$ (red circle) and $n = 4$ (blue circle), calculated by using DFT-PBEsol (using FHI-AIMS code) and two point charges, $q = +2$ and $+4$, respectively, as a function of separation distance. The black and green curves show the same proportions, calculated using the LPM for Pb2 + -Pb2 +, without and with the effective potential between the two lone pairs densities, where the corresponding energetics are shown as same colour scheme in figure 3.11................................................................. 126

**Figure 4.1** Schematic illustration of a one-dimensional energy landscape or a slice from a hypersurface where, for example, the variables are the atomic coordinates of a nanocluster. Regions A and B are super-basins containing many LM. The global minimum is located in region A................................................................. 130
Figure 4.2 (a) The workflow of optimising random configurations employed to predict nanocluster structures. Stages (1), (2) and (3) are energy minimisations on landscapes defined using three different energy definitions, RIM, LPM, and DFT. (b) Schematic diagram showing random points on energy landscape A moving to the bottom of the black basin it sits within, then, after transferring vertically down onto energy surface B, sliding down to the bottom of the red basin it sits within, which, after transferring vertically down onto energy surface C, sliding down to the bottom of the green basin it sits within. The colours match three energy definitions associated with the workflow: landscape A (black curve) is the energy function of rigid ion model, B (red curve) and C (green curve) are the energies from the LPM and DFT approach. The DFT LM are labelled (i), (ii), (iii) and (iv).

Figure 4.3 Example matches between two energy landscapes. The arrows show the results of relaxing the local minima (LM) structures obtained from the red energy landscape to a LM on the green energy landscape. If the criteria for a good match between energy landscapes is matching LM atomic structures and matching LM energy ranking, then: (a) is the most ideal for both conditions; (b) structures match, but incorrect ranking of LM energies; (c) LM ranking correct, but metastable structure heavily distorted; (d) LM ranking incorrect and metastable structure heavily distorted; (e) both LM heavily distorted but energy ranking correct. There is still the question as to whether the heavily distorted LM is a match to the initial LM structure.

Figure 4.4 Scaled DFT-PBEsol refinement wall-times on an Intel(R) Xeon(R) Gold 6140 CPU @ 2.30GHz with 18 cores, as a function of the similarity index (SI) for the five lowest energy LM(PBEsol) configurations of (PbO)$_N$, for $N = 6$ to 12, before-and-after refinement. The markers • and × indicate LPM to PBEsol LM and SM to PBEsol LM, respectively. The black and red error
ellipses, with a confidence interval of $\sigma = 1$, are for the circle and cross markers, respectively.

Figure 4.5 The PBEsol LM energies, $E_p(N)$, of the best 5 $(\text{PbO})N$ nanocluster configurations found from performing a finite search on the LPM (●) and SM (∙) energy landscapes for each size. Panel (b) shows a magnified section of panel (a) for sizes $N = 10$-12.

Figure 4.6 The similarity index (SI) as a function of the DFT energy, $E_p$, for the best 5 PBEsol LM nanoclusters of $(\text{PbO})N$ with $N = 6$-12. Note that the SI values measured the differences between the DFT LM atomic structures and the respective LPM or SM LM structures.

Figure 4.7 Second-order energies, $E_{SOE}$, of the global minima $(\text{PbO})N$ nanoclusters calculated by DFT and LPM. Ball-and-stick models (Pb black and O red) of the nanocluster structures shown above and below the data points are DFT LM and LPM LM, respectively.

Figure 4.8 Dipole moments of global minima $(\text{PbO})N$ nanoclusters, calculated by DFT and LPM where the structures by each level of theory can be found in figure 4.7.

Figure 4.9 Ball-and-stick models of LPM and DFT LM configurations for three $(\text{PbO})_{12}$ clusters. Black and dark green spheres represent Pb atoms for DFT and LPM respectively, whereas red spheres represent O atoms. Lone pair electron densities shown for LPM clusters as semi-transparent dark green surfaces. Blue and green arrows point towards example 3- and 4-coordinated Pb atoms, respectively.

Figure 4.10 Ball-and-stick models of LPM and DFT LM configurations for $(\text{PbO})_{10}$ clusters. Black and dark green spheres represent Pb atoms for DFT and LPM respectively, whereas red spheres represent O atoms. Lone pair electron densities shown for LPM clusters as semi-transparent dark green surfaces. Black and green arrows indicate the flow of relaxation carried out by using DFT and LPM, respectively, and blue arrows point two Pb atoms.
**Figure 4.11** Ball-and-stick models of LPM and DFT LM configurations for (PbO)\textsubscript{9} and (PbO)\textsubscript{10} clusters. Black and dark green spheres represent Pb atoms for DFT and LPM respectively, whereas red spheres represent O atoms. Lone pair electron densities shown for LPM clusters as semi-transparent dark green surfaces. Black and green arrows indicate the flow of relaxation carried out by using DFT and LPM, respectively. 159

**Figure 4.12** Scaled DFT-PBEsol refinement wall-times as a function of the similarity index (SI) for the five lowest energy LM(PBEsol) configurations of (PbF\textsubscript{2})\textsubscript{N}, for \(N = 3\) to 8, before-and-after refinement. The markers • and x indicate LPM to PBEsol LM and SM to PBEsol LM, respectively – the same computer specification was used as in figure 4.4. 164

**Figure 4.13** The similarity index (SI) as a function of the DFT energy, \(E_p\), for the best 5 PBEsol LM nanoclusters of (PbF\textsubscript{2})\textsubscript{N} with \(N = 3\)-8. Note that the SI values measured the differences between the DFT LM atomic structures and the respective LPM or SM LM structures. 165

**Figure 4.14** Total pair correlation function for a few cases of (PbF\textsubscript{2})\textsubscript{N} structures where the different type of curves represent which type of energy definition is used, which is distinguished by the legend provided within the plot (d). Note that the Gaussian broadening factor of 0.05 is used. 167

**Figure 4.15** Second-order energies, \(E_{SOE}\), of the global minima (PbF\textsubscript{2})\textsubscript{N} nanoclusters calculated by DFT and LPM. Ball-and-stick models (Pb black and F white) of the nanocluster structures shown above and below the data points are DFT LM and LPM LM, respectively. 171

**Figure 4.16** Dipole moments of global minima (PbO)\textsubscript{N} nanoclusters, calculated by DFT and LPM where the structures by each level of theory can be found in figure 4.15. 172

**Figure 4.17** \((MF_2)_N\) nanoclusters with \(N = 2\)-7 where \(M = Sr\) and \(Ba\), optimised using DFT PBEsol. In column A and B, the structures of \((MF_2)_N\) are shown overlapped on top of the LPM
optimised \((\text{PbF}_2)_N\) structures with artificially adjusted model parameter \(\lambda = 30.09\) and 32.09 eV used for respective compounds (cf. the value of \(\lambda\) employed for lead fluoride is 14.09 eV). In the fifth column, the structures of \((\text{PbF}_2)_N\) are shown, also optimized using DFT(PBEsol) for comparison. The ball-and-stick model is used with Ba (green), Sr (blue), Pb (black) and F (white). The red lines on the structures of \(N = 2\) are drawn to indicate the bond angles \(\angle F-M-M\) (\(M = \text{Ba, Sr and Pb}\)).

**Figure 4.18** Similarity index (SI) and expansion factor \((\varepsilon_f)\) as a function of the LPM parameter \(\lambda\) for the nanoclusters of \((\text{PbF}_2)_N\) measured with respect to DFT(PBEsol) calculated \((\text{MF}_2)_N\) with \(M = \text{Ba (1a and 2a) and Sr (1b and 2b)}\) given in figure 4.17. In plots 1a and 1b, there are two different types of SI values are presented, which are distinguished by different markers, circle (directly measured SI value of a LPM to a DFT nanocluster) and cross (minimised SI value of a LPM to a DFT nanocluster, measured after applying expansion to the LPM structure – see equation 4.4.1). The missed data points for \(N = 6\) in blue are omitted due to the changes in atomic configuration, during the LPM energy relaxation process with given \(\lambda\).
List of Tables

**Table 2.1** The analytical geometric derivatives of the atoms in the (PbO)₄ nanocluster – see figure 2.8 – calculated by Algorithm II, and the relative errors by the two other controls, (i) and (ii).................................................................................................................................................. 68

**Table 3.1** Ionisation potentials, or energies (IE) of lone pair cations. Experimental values taken from the reference [64], and the relevant HOMO-LUMO gaps were calculated by using the FHI-AIMS code, using provided intermediate basis set, at the DFT-PBEsol level of theory.......... 91

**Table 3.2** The LPM parameters of \( \{ \omega, \lambda \} \) for three different lone pair cations fitted to the dipole polarisabilities of the DFT-PBEsol calculated results where \( \Delta E \) and \( \Delta \mu \) are the root-mean-square deviations of the LPM energy and the dipole moment for each cation measured with respect to the DFT references –details are summarised in figure 3.2. .................................................................................................. 100

**Table 3.3** Dipole moment (d.m.) (e Å), of Pb2+ cation interacting with a point charge at representative distances. (a) DFT values along with the \( \text{pz} \) (or \( \text{cpzDFT} \)) occupancy of the HOMO, (b) and (c) show approximate values calculated using respective equations and error percentages relative to the DFT values in column (a). Note that the effective charge \( \omega = -1.424 \) was used to calculate values reported in (c). .......................................................................................................................... 111

**Table 4.1** Ball-and-stick models for the 5 lowest energy, PBEsol (as implemented within FHI-aims) LM structures of (PbO)ₙ for sizes \( N=6-9 \), where grey and red spheres represent Pb and O ions, respectively.......................................................... 143

**Table 4.2** Ball-and-stick models for the 5 lowest energy, PBEsol (as implemented within FHI-aims) LM structures of (PbO)ₙ for sizes \( N=10-12 \), where grey and red spheres represent Pb and O ions, respectively........................................................................................................................................... 144
Table 4.3 Ball-and-stick models for the 5 lowest energy, PBEsol (as implemented within FHI-aims) LM structures of (PbF₂)ₙ for sizes N=3-6, where black and white spheres represent Pb and F ions, respectively. .................................................................................................................................................................. 162

Table 4.4 Ball-and-stick models for the 5 lowest energy, PBEsol (as implemented within FHI-aims) LM structures of (PbF₂)ₙ for sizes N=3-5, where black and white spheres represent Pb and F ions, respectively. .................................................................................................................................................................. 163

Table 4.5 Relative polymerisation energy of (PbF₂)₆ nanoclusters – taken from table 4.3 – calculated by using different DFT functionals where the numbers in parentheses indicate energy ranking for each functional. Note that the similarity index (SI) was measured with respect to the PBEsol nanocluster structures.................................................................................................................................................................. 169