FIRST-PRINCIPLES STUDY ON THE
STRUCTURAL AND THERMAL PROPERTIES OF
MOLECULAR CRYSTALS AND LIQUIDS

Hsin-Yu Ko

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
CHEMISTRY
ADVISER: ROBERTO CAR

APRIL 2019
© Copyright by Hsin-Yu Ko, 2019.

All rights reserved.
Abstract

Hybrid density functional theory (DFT) is widely used to obtain a semi-quantitative understanding of the electronic structures of isolated molecular clusters; however, it has limited applicability to large molecules and complex condensed-phase materials due to its high computational cost. To overcome this difficulty, a linear-scaling algorithm based on maximally localized Wannier functions (MLWF) can be employed.

In the first part of this thesis, we present a detailed discussion on the theory, real-space implementation, and performance of this algorithm for enabling large-scale condensed-phase hybrid DFT based *ab initio* molecular dynamics (AIMD) simulations under realistic isobaric-isothermal (*NpT*) conditions. For the theory aspect, we discuss how the MLWF-based linear-scaling approach can be integrated into the Car-Parrinello AIMD framework. For our implementation (named *exx* module), we introduced several features to enable massive parallelization using hybrid MPI/OpenMP technologies, including custom data distribution scheme for MLWFs, static load balancing algorithm, and reusable proto-subdomains to exploit the localization in MLWF representation. Based on performance test using realistic condensed-phase liquid water systems (H$_2$O)$_{64}$, (H$_2$O)$_{128}$, and (H$_2$O)$_{256}$, we find the *exx* module to be quite efficient and scalable, hence making a hybrid DFT based AIMD simulation feasible.

In the second part, we apply the *exx* module to study molecular crystals and liquids. The first application is to simulate the ice Ih, II, and III phases using AIMD at the dispersion-inclusive hybrid DFT level at their experimental triple point (238 K, 2.1 kbar). In the second application, we exploit the computational efficiency and scalability of our *exx* module in conjunction with leadership level supercomputers to perform path-integral AIMD (PI-AIMD) simulations on liquid water (300 K) and ice Ih (273 K) both under ambient pressure (1.0 bar) at dispersion-inclusive hybrid DFT level. We identify this level of theory provides a quite accurate description to water.
From the resulting trajectory, we also find that the nuclear quantum fluctuation with autoprotolysis-type distortion promotes orbital localization. In the third application, we explore how anharmonicity, nuclear quantum effects (NQE), many-body dispersion interactions, and Pauli repulsion influence thermal properties of dispersion-bound molecular crystals using pyridine and similar organic molecular crystals.
Acknowledgements

The journey as a graduate student at Princeton is an extraordinary experience in intellectual growth with many challenges. Looking back, I would not have been able to overcome these challenges without the guidance, assistance, and support from a group of amazing people. As such, I would provide my sincere gratitude to these people via the following acknowledgments.

First and foremost, I want to thank my advisor, Professor Roberto Car, for being a passionate, supportive, and insightful mentor. I deeply admire his genuine passion for understanding nature as well as his rigorous and creative scientific approach. His passion also provided me with the opportunity to freely collaborate with and learn from many experts from a wide variety of disciplines. Besides being a superb teacher in science, he is also a role model for pursuing a scientific career. I am incredibly grateful for the opportunity to work with him as well as his patience and time spent on me.

I would also acknowledge members in my advisory committee: Professor Annabella Selloni for kindly serving as the second reader of my thesis, collaborating on several exciting projects, and providing helpful advice throughout the program; Professor Salvatore Torquato for useful discussion in committee meetings and warm lunch-time conversation; Professor Garnet Chan for insightful questions and comments during committee meetings; Professor Haw Yang for chairing my general exam and out-of-field proposal as well as providing kind encouragement.

I am also grateful for the support, assistance, and friendship from current and former members in the Car research groups: Jiming Sun for being a kind senior with helpful advice; Dr. Lukas Muechler for many useful discussions; Yantao Wu and Linfeng Zhang for stimulating scientific collaboration and interactions. Special thanks to my long-term collaborators originated from the Car group (a.k.a the “gang”): Dr. Fausto Martelli for friendship and tutoring in basic Italian conversation which saved
me during my visit to Trieste; Dr. Biswajit Santra for countless encouragements and
Buddha-level patience—particular for his help in deciphering the Quantum Espresso
source code; Professor Robert DiStasio Jr. for helping me set up useful plans to over-
come research problems and immerse into the American culture as well as spending
many early mornings and late evenings training my writing skills—I am deeply in debt
to him. I also appreciate very much the collaborations connected via this “gang”: I
wanted to thank the Professor Alexandre Tkatchenko and his research group (partic-
ular Johannes Hoja, and Dr. Igor Poltavsky) for kindly collaborating with me and
beneficial scientific discussions on molecular crystals; I also wanted to thank Professor
Giancarlo Franzese and Dr. Carles Calero for collaboration to understand the water
confined by phospholipid membranes; Junteng Jia and Zachary Sparrow for useful
discussions on hybrid density functionals and numerical methods.

I am grateful for Professor Xifan Wu’s for kind encouragement and help through-
out the program. I also thank former members in his group—Professor Mohan Chen,
Professor Zhaoru Sun, and Dr. Lixin Zheng—for useful scientific discussions and
collaboration on understanding liquid water and ions in aqueous solutions as well as
many encouragements.

I thank members in the Selloni and Torquato groups for friendship and com-
panionship. In particular, Marcos Calegari Andrade for scientific interaction and
collaboration on water at interfaces and at high temperatures and pressures. Dr. Jia
Chen and Dr. Sencer Selcuk for many useful discussions. Dr. Erdal Oğuz for collab-
oration on a method to quantify local and intermediate range structural order in the
condensed-phase environments. Dr. Ge Zhang for many scientific discussions.

I also thank friends I interacted during this program: Dr. Liang-Yan Hsu for being
a short-term roommate and a long-term mentor; Dr. James Park and Danrui Ni for
friendly and scientific interactions; People in the Taiwanese Association at Princeton
for support and friendship.
I am indebted to the Quantum Espresso community (particularly Professor Paolo Giannozzi, Professor Stefano de Gironcoli, Professor Ari Seitsonen, Professor Anton Kokalj, and Dr. Emine Küçükbenli) for kindly sharing help on my development and providing useful training in the developer workshop during 2015.

I would also thank Meghan Krause and Meredith LaSalle-Tarantin for kind support and encouragement. I am particularly grateful for Meghan’s help on connecting with Professor Andrew Bocarsly, who I thank for solving my two-body problem.

My gratitude also goes to Princeton University and U. S. Department of Energy for funding support. I also thank the following supercomputing facilities for computer resources: (i) National Energy Research Scientific Computing (NERSC) Center, (ii) Argonne Leadership Computing Facility at Argonne National Laboratory, and (iii) Terascale Infrastructure for Groundbreaking Research in Science and Engineering (TIGRESS) High-Performance Computing Center and Visualization Laboratory at Princeton University.

Words cannot describe how grateful I am for the unconditional support from my parents, sister, and other family members in Taiwan as well as my relatives in the East Coast. Finally, I wanted to end my acknowledgment by thanking my partner, Hsin-Ya Kuo, for her continued company and love.
Publications and Presentations

During the Ph.D. program, my research related (both directly and indirectly) to this thesis has lead to the following publications and presentations.

Publications:


Presentations:


To my parents.
## Contents

Abstract ......................................................... iii
Acknowledgements ................................................ v
Publications and Presentations ............................ viii

1 Introduction ................................................. 1

1.1 Motivation ................................................. 1

1.1.1 Molecular Crystal Structure Prediction and Polymorphism .... 2

1.1.2 Molecular Liquids ................................... 3

1.2 Isobaric-Isothermal Molecular Dynamics Simulation .......... 5

1.3 Electronic Structures Based on Density Functional Theory ...... 8

1.4 Thesis Outline ........................................... 10

1 Developing Efficient Hybrid Density Functional Theory ........ 12

2 Theoretical Framework for Large-Scale Condensed-Phase Systems 13

2.1 Computational Challenge ................................ 14

2.2 Theoretical Foundation ................................ 19

2.2.1 Index Conventions ................................... 19

2.2.2 Constant-Pressure Car-Parrinello Molecular Dynamics ...... 20

2.2.3 Linear Scaling via Orbital Localization .................. 25

2.3 Conclusions .............................................. 31

xiii
## 3 Algorithmic Framework: the exx Module

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Prerequisites</td>
<td>35</td>
</tr>
<tr>
<td>3.2 Data Distribution Schemes</td>
<td>39</td>
</tr>
<tr>
<td>3.3 Implementation</td>
<td>41</td>
</tr>
<tr>
<td>3.3.1 Step I: Redistribution of MLWFs</td>
<td>41</td>
</tr>
<tr>
<td>3.3.2 Step II: Construction of Pair List and Proto-Subdomains</td>
<td>42</td>
</tr>
<tr>
<td>3.3.3 Step III: Communication of MLWFs</td>
<td>59</td>
</tr>
<tr>
<td>3.3.4 Step IV: Solution of Poisson’s Equation</td>
<td>62</td>
</tr>
<tr>
<td>3.3.5 Step V: Computation of Energy and Forces</td>
<td>69</td>
</tr>
<tr>
<td>3.3.6 Step VI: Redistribution of Wavefunction Forces</td>
<td>70</td>
</tr>
<tr>
<td>3.4 Conclusions</td>
<td>71</td>
</tr>
</tbody>
</table>

## 4 Accuracy and Performance

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Parameters for Performance Optimization</td>
<td>74</td>
</tr>
<tr>
<td>4.1.1 Convergence of Energy and Cell Forces</td>
<td>75</td>
</tr>
<tr>
<td>4.1.2 Convergence of Wavefunction Forces</td>
<td>77</td>
</tr>
<tr>
<td>4.2 Parallel Scaling and Performance</td>
<td>80</td>
</tr>
<tr>
<td>4.2.1 Internode Parallelization via MPI</td>
<td>80</td>
</tr>
<tr>
<td>4.2.2 Intranode Parallelization via OpenMP</td>
<td>88</td>
</tr>
<tr>
<td>4.3 Crystalline Symmetry and Performance</td>
<td>90</td>
</tr>
<tr>
<td>4.4 Conclusions and Future Outlook</td>
<td>92</td>
</tr>
</tbody>
</table>

## II Understanding Molecular Crystals and Liquids

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 Application to Ice Polymorphs at the Ice I/III–III Triple Point</td>
<td>96</td>
</tr>
<tr>
<td>5.1 Methodology</td>
<td>98</td>
</tr>
<tr>
<td>5.2 Density and Enthalpy at the Triple Point</td>
<td>101</td>
</tr>
<tr>
<td>5.3 Conclusion</td>
<td>105</td>
</tr>
</tbody>
</table>
# 6 Probing Nuclear Quantum Effects in Liquid Water and Ice Ih

- **6.1 Methodology** .................................................. 108
- **6.2 Structural Properties** ...................................... 110
  - **6.2.1 Radial Distribution Functions of Liquid Water** ........ 111
  - **6.2.2 Structure Factors of Ice Ih** ......................... 116
- **6.3 Electronic Properties** ...................................... 118
  - **6.3.1 Wannier Functions in Liquid Water** ................. 118
  - **6.3.2 Valence Band Maximum** ............................... 119
- **6.4 Autoprotolysis** ............................................ 123
- **6.5 Conclusion and Future Outlook** ......................... 126

# 7 Application to Organic Molecular Crystals

- **7.1 Molecular Crystals under Finite Temperatures** .......... 130
- **7.2 Simulation Methodology** .................................. 131
- **7.3 Classical Anharmonicity** ................................ 133
- **7.4 Nuclear Quantum Effects in Pyridine-I** .................. 137
  - **7.4.1 Experimental Behavior** ............................... 139
  - **7.4.2 Feynman Path-Integral Study** ....................... 140
- **7.5 Thermal Expansion of Pyridine-I** ....................... 143
  - **7.5.1 Role of Many-Body Dispersion** ..................... 144
  - **7.5.2 Role of Exact Exchange via the exx Module** ....... 146
- **7.6 Conclusions** .............................................. 146

# III Appendicies: Technical Details

- **A Tight Convergence to the Electronic Ground State** ...... 149
- **B Dispersion-Bound Molecular Crystals** ..................... 151
  - **B.1 Thermal Expansion of Pyridine-I in the Debye Model** 151
Chapter 1

Introduction

1.1 Motivation

Atoms and molecules are fundamental building units in Chemistry, Biology, and Materials Science. While all existing atoms can be listed on a brief periodic table of 118 elements, molecules—the various combinations of atoms—constitute an inexhaustible chemical space for exotic properties. As a chemist, one exploits the gigantic chemical space and synthesizes useful molecules based on empirical knowledge of interatomic interactions—chemical bonds. Interestingly, despite the fact that nearly all applications in Chemistry are performed in condensed-phase environments (e.g., solids, liquids, or solutions), their guiding chemistry involves merely isolated (gas-phase) molecular complexes (including starting reactants and final products). This approximated knowledge works remarkably well as the intramolecular interactions (such as chemical bonds) dominate the intermolecular ones (e.g., non-bonded dispersion interactions and H bonds). However, to understand relevant phenomena such as molecular crystal structure prediction as well as the structural and dynamical properties of molecular liquids—both of which are the main focus of this thesis—the intermolecular interactions often are non-negligible and are necessary; in these cases, the associated energy
scale is much smaller due to the absence of the formation/cleavage of intramolecular chemical bonds. As such, I will begin this chapter by briefly introducing molecular crystals and liquids (Secs. 1.1.1 and 1.1.2). Then, I will review the use of *ab initio* molecular dynamics (AIMD) simulations (Sec. 1.2) to probe realistic (experimental) thermodynamic conditions, under which molecular crystals and liquids are typically found. I will also discuss the electronic structure theory that underlies the AIMD simulations (Sec. 1.3). Finally, I will close this chapter by sketching the thesis outline (Sec. 1.4).

### 1.1.1 Molecular Crystal Structure Prediction and Polymorphism

Molecular crystals are ubiquitous and versatile materials with an impact ranging from medicine/biochemistry (pharmaceutical design, public health) to alternative energy (H$_2$ storage/delivery, organic photovoltaics) and industry/technology (molecular electronics, agrochemicals). It is of crucial importance to the above applications that molecular crystals often have many accessible polymorphs—alternative structures that are nearly identical in stability yet display drastically different physicochemical properties. For instance, polymorphism can affect device performance, *e.g.*, the conductivity of the organic semiconductor in Fig. 1.1 can be significantly increased when constructed from the plate (instead of needle) polymorph, as this form has superior hole ($h^+$) mobility. For another, coordination of polymorphs with distinct properties can even enable novel functionalities, as illustrated by the H$_2$ storage/delivery mechanism in Lithium Hydrazinidoborane (LiHB). In this case, the stable low-temperature $\beta$-LiHB polymorph is used to store H$_2$; upon heating, this polymorph is converted into the meta-stable $\alpha$-LiHB polymorph, which releases H$_2$ as a fuel source. A quite pertinent example of the devastating effects that can result from molecular crystal polymorphism resides in the extraordinary case of the human immunodeficiency virus
(HIV) inhibitor drug ritonavir, in which a previously unknown polymorph with a significantly lower bioavailability (Form II) began to appear throughout the bulk drug supply, compromising access to the life-saving polymorph (Form I). From these examples alone, it is clear that knowledge of the structures and stabilities of all thermodynamically (and even kinetically) competitive polymorphs is essential to accessing the vast and largely unexplored potential of molecular crystals. For example, I will discuss the relevant physical interactions that are necessary to describe the molecular crystal structures based on molecular crystals comprising small heterocyclic aromatic molecules and a detailed case study of pyridine-I polymorph.

1.1.2 Molecular Liquids

When a molecular crystal is placed under a temperature between its melting and boiling points, the thermal fluctuation is large enough to make the crystalline lattice instable yet insufficient to dissociate the lattice into a gas (i.e., nearly isolated environment for each molecule); as such, the system converts into the liquid phase. In a microscopic perspective, the comprising molecules in this liquid becomes diffusive as in a fluid, while their mutual distances remain similar to those found in the solid phase. When observed macroscopically, such molecular liquids can change shape in response to applied mechanical stress due to the diffusive nature but remains largely incompressible because of the Pauli repulsion between the contacting neighboring molecules.

In Science and Technology, molecular liquids are commonly used as solvents (e.g., liquids made up of water, alkane, ether, ketone, or other small organic molecules) to promote chemical reactions and biological processes. Among the common molecular liquids, liquid water is the ubiquitous “matrix of life”; liquid water also has many unique properties, including its potency as solvent, ability to form hydrogen bonds (H bonds), and amphoteric nature. Although liquid water have received a lot of atten-
Figure 1.1: Schematic illustration of the widespread use of molecular crystals and the critical roles played by energetically competitive polymorphs with different properties.

tion from the scientific community, an accurate first-principles based understanding remains highly challenging to date. This challenge mainly arises from the difficulty to simultaneously obtain accurate quantum mechanical intermolecular interactions and perform extensive statistical mechanical samplings under realistic thermodynamic conditions, which introduces a grand computational challenge. Among these relevant intermolecular interactions, an accurate description of the H-bond network is particularly delicate and requires a sophisticated treatment of the quantum mechanical exchange interactions. Such sophisticated treatment of the exchange interactions
in condensed-phase environment is both computationally demanding and inefficient in terms of scaling with system size using the conventional approach. To overcome this computational difficulty, an efficient theoretical and algorithmic framework is developed in this thesis, which is well suited for calculations on high-performance computing (HPC) architectures. Based on this framework, a first-principles based microscopic understanding of liquid water can be obtained. Besides the application to liquid water, this framework is suitable for other molecular liquids in general.

1.2 Isobaric-Isothermal Molecular Dynamics Simulation

To study the molecular solids and liquids under realistic experimental thermodynamic conditions—often under a particular temperature $T$ and pressure $p$, it is necessary to perform proper statistical mechanical samplings within the isobaric-isothermal $(NpT)$ ensemble. One way to perform such samplings is via molecular dynamics (MD) simulations. MD simulation under realistic thermodynamic conditions is commonly used to gain microscopic insights into material properties,\cite{10} chemical reactions,\cite{11} and biological processes.\cite{12} In MD simulations, the atoms are often treated as classical particles with the following Lagrangian (for a system of $N$ atoms interacting via a potential $U(\{\mathbf{R}_I\})$):

$$\mathcal{L}_{NVE} = \sum_{I=1}^{N} \frac{M_I \dot{\mathbf{R}}_I^2}{2} - U(\{\mathbf{R}_I\}),$$

leading to the Newton’s equations of motion as follows:

$$M_I \ddot{\mathbf{R}}_I = -\nabla_I U(\{\mathbf{R}_I\}),$$

in which $\mathbf{R}_I$ and $\{M_I\}$ are the position and mass of the $I$-th atom.
Time integration of Eq. (1.2) directly samples the microcanonical (NVE) ensemble, i.e., the ensemble with constant internal energy \(E\), volume \(V\), and number of atoms \(N\). Modifications by incorporating extended degrees of freedom on this standard MD simulation are needed to make the targeted \(NpT\) sampling possible. On the one hand, to make the system perform canonical (constant temperature) sampling that reproduces Boltzmann distribution, a thermostat can be attached to the atoms in an MD simulation as an extended degree of freedom \(s\) with an effective mass \(Q\). To do so, a conventional way is via Nosé approach\(^{13}\) which modifies the Lagrangian as:

\[
\mathcal{L}_{\text{Nosé}} = \sum_{I=1}^{N} \frac{M_I s^2 \dot{R}_I^2}{2} - \mathcal{U}(\{R_I\}) + \frac{Q}{2} \dot{s}^2 - \frac{g}{\beta} \ln s, \tag{1.3}
\]

where \(g = 3N\) is the number of degrees of freedom, \(\beta = \frac{1}{k_B T}\) is the inverse temperature, and \(k_B\) is Boltzmann’s constant. The equations of motion that follow from the Nosé Lagrangian involve scaling of time and are cumbersome. An equivalent approach was proposed by Hoover\(^{14}\) and is known as the Nosé-Hoover approach. Its equations of motion are:

\[
\dot{R}_I = \frac{P_I}{M_I} \tag{1.4}
\]

\[
\dot{P}_I = -\nabla I \mathcal{U}(\{R_I\}) - \xi P_I \tag{1.5}
\]

\[
\dot{\xi} = \left( \sum_I \frac{P_I^2}{M_I} - \frac{g}{\beta} \right) / Q \tag{1.6}
\]

\[
\dot{s}/s = \xi, \tag{1.7}
\]

in which \(P_I\) is the momentum of the \(I\)-th atom, \(\xi \equiv \dot{s}/s\) is an auxiliary variable for thermostat. By integrating Eqs. (1.4)–(1.7), the modified MD simulation samples the canonical (NVT) ensemble. On the other hand, constant-pressure sampling can be achieved using the Parrinello-Rahman barostat.\(^{15}\) For isobaric-isoenthalpic (\(NpH\))

---

*Here, we assume the so-called ergodic hypothesis—long time average reproduces the ensemble average.
ensemble, the NVE Lagrangian (Eq. (1.3)) can be modified to

\[
\mathcal{L}_{NpH} = \frac{1}{2} \sum_{I=1}^{N} M_I \dot{\mathbf{S}}_I^\top \mathbf{G} \dot{\mathbf{S}}_I - \mathcal{U}(\{\mathbf{R}_I\}) + \frac{1}{2} W \text{Tr} \dot{\mathbf{h}}^\top \dot{\mathbf{h}} - pV, \tag{1.8}
\]

where \(\mathbf{h}\) is the cell tensor, \(\mathbf{S}_I = \mathbf{h}^{-1} \mathbf{R}_I\) is the crystal coordinates of \(I\)-th atom, \(\mathbf{G} = \mathbf{h}^\top \mathbf{h}\) is the so-call metric tensor, \(W\) is the cell “mass”, \(p\) is the external pressure, and \(V = \det \mathbf{h}\) is the volume of the cell. This Lagrangian leads to the following equations of motion:

\[
M_I \ddot{\mathbf{S}}_I = -\mathbf{h}^{-1} (\nabla_{\mathbf{R}_I} E) - M_I \mathbf{G}^{-1} \dot{\mathbf{G}} \dot{\mathbf{S}}_I, \tag{1.9}
\]

\[
W \ddot{\mathbf{h}} = (\mathbf{\Pi} - p \mathbf{I}) (\mathbf{h}^\top)^{-1} V, \tag{1.10}
\]

in which \(\mathbf{\Pi}\) is the internal stress tensor and \(\mathbf{I}\) is a \(3 \times 3\) identity matrix. When both the thermostat and barostat are attached to an MD simulation (with appropriate setting of \(Q\) and \(W\)), \(NpT\) sampling can be achieved.

The accuracy of an MD simulation is dictated by the underlying interatomic potential or force field (FF) \(\mathcal{U}(\{\mathbf{R}_I\})\). FFs are typically constructed empirically with simple analytical functions with adjustable parameters tuned to reproduce experimental observations or high-level theoretical benchmark data. However, FFs often lacks transferability to thermodynamic conditions different from those used for the FF parametrization,\textsuperscript{16} e.g., a FF trained only based on ice Ih data may be inaccurate for a liquid water simulation. To overcome this difficulty, a physically sound approach for obtaining the interatomic potential is \textit{via} solving first-principles electronic structure theories, \textit{i.e.}, using \textit{ab initio} MD (AIMD) approaches.\textsuperscript{17}
1.3 Electronic Structures Based on Density Functional Theory

In view of its quite favorable balance of accuracy and computational cost, Kohn-Sham (KS) density functional theory (DFT) has become the most widely used electronic structure method for ab initio molecular dynamics (AIMD) simulations of large molecules and complex condensed-phase materials. Within the framework of DFT, the total ground-state energy \( E \) is given as the sum of the following contributions:

\[
E = E_{\text{kin},0} + E_{\text{ext}} + E_{\text{H}} + E_{\text{xc}},
\]

in which \( E_{\text{kin},0} \) is the KS kinetic energy (of the non-interacting electrons), \( E_{\text{ext}} \) is the external potential which accounts for the nuclear-electronic and nuclear-nuclear potential energies (as well as any external fields), \( E_{\text{H}} \) is the Hartree energy, i.e., the average (classical) Coulomb interaction energy of the electrons, and \( E_{\text{xc}} \) is the electronic exchange-correlation (xc) energy. Explicit forms for all of the energy contributions in Eq. (1.11) are known with the exception of \( E_{\text{xc}} \), the approximation of which is still the subject of active research to date.

Functional approximations to \( E_{\text{xc}} \) are often described as the rungs of “Jacob’s Ladder”, which connects the Hartree world to the exact solution of the time-independent Schrödinger equation. In this hierarchical classification of DFT, the first rung is given by the local (spin) density approximation (LDA). In LDA, the xc energy density \( \varepsilon_{\text{xc}}(\mathbf{r}) \) — a functional of electron density — in \( E_{\text{xc}} = \int d\mathbf{r} \rho(\mathbf{r}) \varepsilon_{\text{xc}}(\mathbf{r}) \) is approximated as a function of the local density \( \varepsilon_{\text{xc}}^{\text{LDA}}(\rho(\mathbf{r})) \). The form of the LDA xc functional is obtained from the solution for the homogeneous electron gas, and therefore works particularly well for systems with (nearly) uniform \( \rho(\mathbf{r}) \) such as the valence electrons in metallic solids. The next rung includes xc functionals based on the semi-local generalized gradient approximation (GGA), which utilize the gradient of the electron density.
density \( (\nabla \rho(r)) \) to correct the LDA description of systems with spatially varying \( \rho(r) \) (such as molecules and heterogeneous materials). At the current time, GGAs such as the non-empirical Perdew-Burke-Ernzerhof (PBE) xc functional\(^\text{28}\) are the computational workhorses for AIMD simulations of condensed-phase systems containing 100s–1000s of atoms. In this size regime, GGA-based approaches provide a favorable compromise between accuracy and computational cost, and have been quite successful in qualitatively (and sometimes even quantitatively) describing a number of systems and processes of interest throughout chemistry, physics, and materials science.

Despite such widespread success, GGA functionals are unable to account for non-local electron correlation effects, which are responsible for the ubiquitous class of dispersion (or van der Waals) interactions. As such, several approaches have been devised to incorporate these long-range forces into the framework of DFT\(^\text{29,32}\) which include the class of effective pairwise models\(^\text{33,37}\) methods that account for many-body dispersion interactions\(^\text{38,41}\) as well as non-local xc functionals\(^\text{43,45}\). We note in passing that third-rung meta-GGA functionals, which incorporate second-derivative information \( \text{via} \) the Laplacian \( (\nabla^2 \rho(r)) \) or the kinetic-energy density, are able to account for intermediate-range correlation effects and have experienced a resurgence with the recent introduction of the SCAN functional\(^\text{46}\), \textit{e.g.}, for bulk water systems\(^\text{9,47,48}\) and interfacial water\(^\text{49}\).

Another significant shortcoming associated with GGA (as well as meta-GGA) functionals is their proneness to significantly suffer from self-interaction error (SIE), an artifact in approximated DFT which manifests as a spurious interaction between an electron and itself.\(^\text{50,51}\) With such SIE, \( \rho(r) \) is too delocalized, which in turn often leads to deleterious effects such as inadequate descriptions of transition states and charge transfer complexes,\(^\text{52,55}\) underestimation of band gaps,\(^\text{54}\) overestimation of lattice parameters in a wide variety of solids,\(^\text{55}\) as well as excessive proton delocalization in liquid water,\(^\text{56,58}\) to name a few. The most commonly adopted approach
for ameliorating the SIE in semi-local KS-DFT is via the admixture of a fraction of exact exchange (EXX) in the underlying (GGA or meta-GGA) xc functional. These so-called hybrid xc functionals constitute the fourth rung in the KS-DFT hierarchy and can be written as (shown here as a correction to a GGA xc functional):

\[ E_{xc}^{\text{hybrid}} = a_x E_{xx} + (1 - a_x) E_{x}^{\text{GGA}} + E_{c}^{\text{GGA}}, \]  

(1.12)

in which \( a_x \) is a mixing parameter, \( E_{xx} \) is the EXX energy, and \( E_{x}^{\text{GGA}} \) and \( E_{c}^{\text{GGA}} \) are the GGA exchange and correlation energies, respectively. For a closed-shell system with \( N_o \) doubly occupied orbitals (bands), \( E_{xx} \) can be written as:

\[ E_{xx} = -\sum_{ij} \int dr \int dr' \frac{\phi_i^*(r) \phi_j^*(r') \phi_j(r) \phi_i(r')}{|r - r'|}, \]  

(1.13)

in which \( \phi_i \) and \( \phi_j \) represent the KS orbitals and the sum extends over all \( N_o \) states. The mixing parameter \( (a_x) \) in this expression for \( E_{xc}^{\text{hybrid}} \) depends on the hybrid xc functional approximation.\(^{59-61}\) By reducing the SIE, hybrid xc functionals are typically more accurate than GGA (or meta-GGA) approaches, in particular for the prediction of lattice parameters,\(^{55}\) reaction energy barriers, and band gaps.\(^{62}\) In this work, we limit our focus to the non-empirical PBE0\(^{60}\) hybrid xc functional, in which \( a_x = 1/4 \) and one adopts the PBE GGA functional\(^{28}\) for \( E_{x}^{\text{GGA}} \) and \( E_{c}^{\text{GGA}} \). Application of our approach (which is described below) to other popular hybrid xc functionals, such as B3LYP\(^{27,50}\) is straightforward.

### 1.4 Thesis Outline

This thesis has two main parts: (i) enabling accurate and efficient condensed-phase AIMD simulations at the hybrid DFT level of theory and (ii) understanding molecular crystals and liquids via such AIMD simulations. The first part comprises a
theoretical (Chapter 2) and algorithmic (Chapter 3) development of a linear-scaling module named exx to compute \( E_{xx} \). The exx module enables large-scale condensed-phase hybrid DFT based AIMD simulation using massively parallel supercomputer architectures, e.g., \( E_{xx} \) of a liquid water snapshot with a simulation cell containing \((\text{H}_2\text{O})_{256}\) can be computed (in a formally exact fashion) within 3 s walltime, which is much less than what would be required to compute the same quantity with standard methodologies\(^1\) (Chapter 4). In the second part, exx module is applied to study a set of molecular crystals and liquids at the hybrid DFT level of theory. In Chapter 5 the exx module allows us to perform AIMD simulation on ice Ih, II, and III using dispersion-inclusive hybrid DFT within the \( NpT \) ensemble at the the experimental triple point. In Chapter 6 the exx module allows us to use path-integral- (PI-) AIMD simulation to study the nuclear quantum effects (NQEs) in liquid water and ice Ih. In Chapter 7 the exx module is used to better understand the relative accuracy of GGA and hybrid functional schemes when studying molecular crystals. Such finding shed light on the understanding of the complex interplay between the physical effects that determine the structural and thermal properties in dispersion-bound molecular crystals like pyridine. Additional technical details are provided in the Appendices A–B.

\(^1\)In this case, the difference in computational cost is expected to be around two orders of magnitude.
Part I

Developing Efficient Hybrid Density Functional Theory
Chapter 2

Theoretical Framework for Large-Scale Condensed-Phase Systems

As discussed in Sec. 1.3, the main computational task in hybrid density functional theory (DFT) is to evaluate $E_{xx}$ in Eq. (1.13). This equation can be expressed in the following compact form:

$$E_{xx} = -\sum_{ij} \int dr \rho_{ij}(r) v_{ij}^{*}(r),$$

(2.1)

by defining the orbital-product density as

$$\rho_{ij}(r) \equiv \phi_i^*(r)\phi_j(r)$$

(2.2)

and the corresponding exact-exchange (EXX) orbital-product potential (i.e., the Coulomb potential felt by a test charge located at $r$ originating from the $\rho_{ij}(r')$ charge distribution) as

$$v_{ij}(r) \equiv \int dr' \frac{\rho_{ij}(r')}{|r - r'|}.$$  

(2.3)
2.1 Computational Challenge

Evaluation of \( v_{ij}(r) \) is therefore of central importance in EXX calculations, and for periodic systems, this quantity is usually computed through the convolution theorem (shown here at the \( \Gamma \)-point only), \( \text{i.e.,} \)

\[
v_{ij}(G) = 4\pi \frac{\rho_{ij}(G)}{|G|^2} \text{invFFT} \to v_{ij}(r),
\]

(2.4)

in which \( v_{ij}(G) \) and \( \rho_{ij}(G) \) are the Fourier coefficients of \( v_{ij}(r) \) and \( \rho_{ij}(r) \), respectively, and \( \text{invFFT} \) denotes the inverse (fast) Fourier transform (FFT). The computational scaling associated with both the \( \text{fwdFFT} \) (forward FFT) and the \( \text{invFFT} \) is \( \mathcal{O}(N_{\text{FFT}} \log N_{\text{FFT}}) \), where \( N_{\text{FFT}} \) is the size of the reciprocal space (planewave) grid which grows linearly with system size. Since the evaluation of \( E_{\text{xx}} \) in Eq. (2.1) requires a sum over the contributions from all \( N_o(N_o + 1)/2 \) unique pairs of occupied orbitals, the overall computational scaling becomes \( \mathcal{O}(N_o^2 N_{\text{FFT}} \log N_{\text{FFT}}) \). The resulting cubic scaling (neglecting logarithmic dependence) makes this reciprocal-space EXX algorithm quite computationally demanding and limits routine performance of hybrid DFT-based AIMD simulations on large-scale condensed-phase systems of interest. Hence, most applications of hybrid DFT still remain limited to small systems and condensed-phase properties at classical 0 K (neglecting the zero-point fluctuations).

Significant progress has been made to accelerate EXX calculations in the condensed-phase by employing various theoretical and numerical techniques, including: range-separation\(^6\) or truncation\(^6\) of the underlying Coulomb operator, implementation of massively parallel algorithms,\(^6^5\)\(^6^6\) employment of auxiliary atom-centered (localized) basis sets,\(^6^9\)\(^7^1\) adaptive compression (lower-rank decomposition) of the EXX operator (ACE),\(^7^2\) use of the projected commutator direct inversion of the iterative subspace (PC-DIIS) method to reduce the number of self-consistent field (SCF) iterations,\(^7^3\) utilization of sparsity through localization methods (\( \text{e.g.,} \),
maximally localized Wannier functions (MLWFs), \(^{8,74-76}\) recursive subspace bisection (RSB), \(^{77,78}\) selected columns of the density matrix (SCDM), \(^{79,81}\) and other localized representations\(^{82}\), and combinations thereof.\(^{64,69,83-85}\)

To enable large-scale hybrid DFT-based AIMD simulations in the condensed-phase environment, the most promising methods for reducing the intrinsic computational cost and scaling associated with EXX exploit sparsity via localized representations of the occupied space or density matrix. The RSB method of Gygi and coworkers uses a non-iterative algebraic decomposition of the wavefunction coefficients, which provides a transformation from the occupied KS eigenstates to a set of localized orbitals that are contained within a prescribed domain in real space. This method has already enabled a number of AIMD simulations using hybrid xc functionals \((e.g.,\) computational investigations into the finite-temperature density of ice, \(^{86}\) ion solvation, \(^{87,88}\) and the structural and vibrational properties of liquid water\(^{56-58}\)\) and is particularly convenient for simulating heterogeneous systems such as solid-liquid interfaces due to the ease of selecting the prescribed localization domains. The SCDM method by Damle, Lin, and Ying exploits the sparsity of the off-diagonal elements of the density matrix, \(^{79-81}\) and does not rely on an initial guess to iteratively localize the occupied space. As such, this approach sidesteps issues related to gauge invariance and can furnish more robust localized orbitals than other iterative localization schemes.\(^{90}\) The MLWF formalism introduced by Marzari and Vanderbilt uses an iterative scheme to obtain a localized representation of the occupied KS orbitals by minimizing a total spread functional \((e.g.,\) the sum of the spreads of the individual localized orbitals) and therefore extends the well-known Boys orbital localization scheme\(^{91}\) used in quantum chemistry into the condensed phase.\(^{74}\) MLWFs have shown great promise as both qualitative and quantitative analysis tools due to their similarity to the orbitals encountered in molecular orbital (MO) theory \((i.e.,\) bonding pairs, lone pairs, etc.) and the fact that they allow one to obtain molecular multipole moments\(^{92-94}\) partition the charge
density and/or electrostatic potential and even compute non-bonded dispersion interactions in complex condensed-phase environments. Numerous algorithms (such as wannier90) for obtaining MLWFs have been incorporated into a number of existing community codes such as Quantum ESPRESSO (QE), SIESTA, ABINIT, NWChem, GPAW, CP2K, and VASP, which makes this localization scheme readily available and quite practical for a posteriori analyses of DFT-based calculations and AIMD simulations. Furthermore, the MLWF localization scheme is particularly suitable for large-scale hybrid DFT based AIMD simulations since a Car-Parrinello-like propagation of the MLWFs has already been demonstrated, which makes the computational cost associated with orbital localization negligible between AIMD steps. In light of this computationally efficient orbital localization scheme, the wide availability of MLWFs, and the promise of a robust tool for on-the-fly analytics, we will now focus our discussion on the development and implementation of a formally exact and linear-scaling MLWF-based EXX algorithm which can be used to perform large-scale condensed-phase AIMD simulations at the hybrid DFT level of theory.

In this chapter, we will focus on Car-Parrinello molecular dynamics (CPMD) simulations of sufficiently large finite-gap condensed-phase systems such that the first Brillouin zone can be accurately sampled at the Γ-point. Extensions to Born-Oppenheimer molecular dynamics (BOMD) and metallic systems are possible and will be discussed in future work. Working at the Γ-point allows us to consider real-valued orbitals only, i.e., \( \phi_i(r) = \phi_i^*(r) \), from which it follows that \( \rho_{ij}(r) = \rho_{ji}(r) \) and \( v_{ij}(r) = v_{ji}(r) \) in Eqs. (2.2)–(2.1). Without loss of generality, we will also assume that the total wavefunction is closed shell (spin-unpolarized). Under these conditions, one can show that the set of MLWFs, \( \{ \tilde{\phi} \} \), which are obtained via an orthogonal (unitary) transformation of the occupied KS eigenstates, i.e.,

\[
\tilde{\phi}_i(r) = \sum_j U_{ij} \phi_j(r),
\]  

(2.5)
have a significantly smaller support than the entire simulation cell and are, in fact, exponentially localized in real space. These features of the MLWF representation of the occupied space provide a theoretical and computational framework for exploiting the natural sparsity in the real-space evaluation of the EXX energy (and forces, vide infra) that we will explore in the first part of this thesis. To demonstrate that the use of MLWFs leads to a linear-scaling EXX approach, consider the $E_{xx}$ expression in Eq. (2.1). Since this expression is invariant to any orthogonal transformation of the occupied orbitals, evaluation of this quantity within the MLWF representation is formally exact (see Sec. 2.2.3). The first level of computational savings originates from the fact that a given MLWF only appreciably overlaps with a subset of neighboring MLWFs, which makes the number of non-vanishing EXX pair interactions per orbital independent of the system size and therefore reduces the total number of orbital pairs required in the summation over $i$ and $j$ in Eq. (2.1). In addition, one can further exploit the fact that an exact evaluation of $E_{xx}$ only requires that the spatial integral in Eq. (2.1) be performed on the support of the orbital-product density (which is now sparse in the MLWF representation) and can therefore be restricted to a real-space domain that is independent of the system size. Taken together, these observations can be leveraged to construct a computationally efficient linear-scaling MLWF-based algorithm for computing $E_{xx}$.

The initial concept and several pilot algorithms for this formally exact and linear-scaling MLWF-based EXX approach have already been successfully used to enable a number of large-scale hybrid-DFT based applications, e.g., computational investigations into the electronic structure of semi-conducting solids, the structural properties of ambient liquid water, the structure and dynamics of aqueous ionic solutions as well as the thermal properties of the pyridine-I polymorph and stability in other molecular crystals. In this chapter, we present a general theoretical and algorithmic framework for computing the EXX energy and orbital-
dependent potential (wavefunction forces) in general Bravais lattice based simulation cells, thereby enabling linear-scaling hybrid DFT-based AIMD simulations of large-scale condensed-phase systems of interest in the microcanonical (\textit{NVE}), canonical (\textit{NVT}), and isobaric-isothermal (\textit{NpT}) ensembles. As such, the current development will enable us to utilize the fourth-rung of DFT (in the Jacob’s Ladder introduced in Sec. 1.3) in the study of the structure, properties, and dynamics of a number of important condensed-phase systems, as well as perform AIMD simulations across extended length- and time-scales which have been prohibitively difficult to access to date. We also describe the implementation and performance of our massively parallel MLWF-based EXX algorithm in the pseudopotential- and planewave-based open-source QE package.\textsuperscript{99,100} The method described herein can be incorporated into any planewave-based DFT code or combined with linear-scaling GGA codes, such as PARSEC,\textsuperscript{126,127} BigDFT,\textsuperscript{128} ONETEP,\textsuperscript{129} or CONQUEST,\textsuperscript{130} to achieve a fully (overall) linear-scaling hybrid DFT approach. It is also noteworthy that our linear-scaling EXX approach also sets the stage for performing quantum chemical electronic structure calculations in the condensed phase, since a majority of the theoretical and algorithmic developments presented herein are directly applicable to the iterative solution of the Hartree-Fock (HF) equations and can be adapted to enable a hierarchy of post-HF local electron correlation methods as well as fifth-rung xc functionals (\textit{e.g.}, efficient GW method using MLWFs\textsuperscript{131}).

In the remainder of the chapter (Sec. 2.2), we describe the theoretical foundation for performing CPMD simulations at the hybrid DFT level of theory within the MLWF representation. Our massively parallel algorithmic implementation in the open-source QE package will be described in Chapter 3. This is followed by a detailed systematic analysis of the computational performance and scaling of the current implementation in Chapter 4 which conclude the first part of this thesis.
2.2 Theoretical Foundation

In this section, we describe the theory behind our real-space MLWF-based framework for performing large-scale AIMD simulations of finite-gap condensed-phase systems at the hybrid DFT level of theory. We will focus the discussion below on the equations of motion underlying constant-pressure CPMD simulations in the $NpH$ ensemble. Adaptation of this approach to sample other ensembles is straightforward and proceeds by simply thermostatting the ionic degrees of freedoms ($NpT$), enforcing a fixed cell ($NVE$), or both ($NVT$). Furthermore, variable-cell (VC) and/or ionic structural optimizations are equally accessible with the use of second-order damped dynamics (or other optimization techniques such as conjugate gradient). Although we limit our scope herein to CPMD, which provides a computationally efficient orbital localization (propagation) scheme, a cost-effective and competitive extension to BOMD has been achieved by our group and will be addressed in a forthcoming paper.

2.2.1 Index Conventions

We will utilize the following conventions for the various indices encountered in this chapter:

- $i, j, k$: indices for the $N_o$ occupied orbitals (or MLWFs)
- $a, b, c$: indices corresponding to the Cartesian directions $x, y,$ and $z$
- $\alpha, \beta, \gamma$: indices corresponding to the cell (lattice) vectors $L_1, L_2,$ and $L_3$
- $I, J, K$: indices for the $N_A$ ions
- $q$: index for the points on the real-space grid
- $l, m$: indices for the spherical harmonics
2.2.2 Constant-Pressure Car-Parrinello Molecular Dynamics

Equations of Motion

In constant-pressure CPMD, fictitious dynamics are introduced on the $N_o$ occupied KS orbitals $\{\phi_i(r)\}$ and simulation cell tensor $h$ via artificial (fictitious) masses $\mu$ and $W$, respectively. In this chapter, $h$ is a $3 \times 3$ matrix defined as $h \equiv (L_1 \ L_2 \ L_3)$ or $h_{aa} \equiv (L_a)_a$, where $L_1$, $L_2$, and $L_3$ are the corresponding cell (lattice) vectors. The simulation cell volume will be denoted by $V = \det (h)$. Constant-pressure ($NpH$) CPMD simulations with the Parrinello-Rahman barostat\textsuperscript{15} (cf. Eqs. (1.9)–(1.10)) are governed by the following equations of motion for the electronic, ionic, and cell degrees of freedom:\textsuperscript{17}

\begin{align}
\mu \ddot{\phi}_i(r) &= - \left( \frac{\delta E}{\delta \phi^*_i(r)} \right) + \sum_j \Lambda_{ij} \phi_j(r) \tag{2.6} \\
M_I \ddot{S}_I &= - h^{-1} (\nabla_{R_I} E) - M_I \mathcal{G}^{-1} \mathcal{G} \dot{S}_I \tag{2.7} \\
W \ddot{h} &= (\Pi - p I) (h^T)^{-1} V, \tag{2.8}
\end{align}

in which Newton’s dot notation was used to indicate time derivatives, $E$ is the total DFT (electronic) energy in Eq. (1.11), $-(\delta E/\delta \phi^*_i(r))$ is the force acting on the $i$-th occupied KS wavefunction, $\Lambda_{ij}$ is a Lagrange multiplier enforcing orthonormality in $\{\phi_i(r)\}$, $-\nabla_{R_I} E$ is the force acting on the $I$-th ion (which is located at $R_I$ with mass $M_I$), $\mathcal{G} = h^T h$ is the so-called metric tensor, $\Pi$ is the total internal stress tensor, and $p$ is the applied (external) pressure. For the fluctuating simulation cells encountered in constant-pressure CPMD, it is more convenient to work in crystal (fractional) coordinates for the ions, $S_I$, which are independent of the dynamical variables associated with the cell degrees of freedom and are related to the Cartesian coordinates, $R_I$, via $R_I = h S_I$ or $S_I = h^{-1} R_I$. 

20
The components of the total internal stress tensor, \( \Pi \), can be further decomposed into nuclear and electronic contributions as follows:

\[
\Pi_{ab} = \Pi_{ab}^{(n)} + \Pi_{ab}^{(e)}.
\]  

(2.9)

The nuclear contribution, \( \Pi_{ab}^{(n)} \), originates from the ionic kinetic energy via

\[
\Pi_{ab}^{(n)} = \frac{1}{V} \sum_I M_I \sum_{\alpha\beta} h_{a\alpha} \dot{S}_{I,\alpha} \dot{S}_{I,\beta} h_{b\beta},
\]

(2.10)

while the electronic part, \( \Pi_{ab}^{(e)} \), arises from cell derivatives, \( \sigma^{a\alpha} \), of the total DFT (electronic) energy:

\[
\Pi_{ab}^{(e)} = -\frac{1}{V} \sum_{\alpha} \left( \frac{\partial E}{\partial h_{a\alpha}} \right) h_{b\alpha} \equiv -\frac{1}{V} \sum_{\alpha} \sigma^{a\alpha} h_{b\alpha}.
\]

(2.11)

At the hybrid DFT level, the equations of motion governing constant-pressure CPMD in Eqs. (2.6)–(2.8) will depend on \( E_{xc} \) via the derivatives of the DFT energy found in the wavefunction forces and cell derivatives in the electronic part of the stress tensor. As such, we will discuss the EXX contribution to both of these quantities below.

**EXX Contribution to the Wavefunction Forces**

In semi-local (GGA-based) DFT, \( E_{xc} = E_{xc}[\rho(r)] \) is a functional of the total (electron) charge density, given by \( \rho(r) = 2 \sum_i \phi_i^*(r) \phi_i(r) \). As such, one can write the \( E_{xc} \) contribution to the (negative of the) wavefunction force (for the \( i \)-th KS orbital) as the action of the so-called xc potential, \( v_{xc}(r) \), on the KS orbital:

\[
\left( \frac{\delta E_{xc}}{\delta \phi_i^*(r)} \right) = \left( \frac{\delta E_{xc}}{\delta \rho(r)} \right) \left( \frac{\delta \rho(r)}{\delta \phi_i^*(r)} \right) \equiv 2v_{xc}(r)\phi_i(r).
\]

(2.12)
Since the explicit functional dependence of $E_{xx}$ on $\rho(r)$ is not known, one needs special procedures such as the optimized effective potential (OEP) method\textsuperscript{133} to derive the EXX contribution to the wavefunction forces within a strict KS-DFT scheme. In this chapter, we adopt a generalized KS-DFT scheme (\textit{i.e.}, by allowing for an orbital-dependent $v_{xc}(r)$), which requires significantly less computational effort and yields essentially the same ground-state energies as the OEP formalism. In this approach (which is currently the standard practice in the field), we compute the corresponding orbital-dependent EXX wavefunction forces, $D_{xx}^i(r) = -(\partial E_{xx}/\partial \phi_i^*(r))$, by taking the functional derivative of $E_{xx}$ in Eq. (2.1) with respect to $\phi_i^*(r)$:

$$D_{xx}^i(r) = \sum_j v_{ij}(r)\phi_j(r) \equiv \sum_j D_{ij}^{xx}(r), \quad (2.13)$$

where we have used the expressions in Eqs. (2.2) and (2.3) for the orbital-product density and potential, $\rho_{ij}(r)$ and $v_{ij}(r)$, and defined $D_{ij}^{xx}(r)$ as the action of $v_{ij}(r)$ on $\phi_j(r)$. From Eq. (2.13), it is again clear that the evaluation of the orbital-product potential, $v_{ij}(r)$, is of central importance to the calculation of $D_{xx}^i(r)$.

**EXX Contribution to the Stress Tensor**

As seen in Eq. (2.11), the EXX contribution to the electronic part of the total internal stress tensor is through the corresponding cell derivatives of the DFT energy. Using Eqs. (2.1)–(2.3), the EXX contribution to $\Pi^{(e)}$ can therefore be computed \textit{via} $\sigma^{aa}_{xx} = (\partial E_{xx}/\partial h_{aa})$, given as:

$$\sigma^{aa}_{xx} = -\sum_{ij} \frac{\partial}{\partial h_{aa}} \int d\mathbf{r} \int d\mathbf{r}' \frac{\rho_{ij}(r)\rho_{ij}(r')}{|\mathbf{r} - \mathbf{r}'|}. \quad (2.14)$$

To compute these cell derivatives, it is again more convenient to work in crystal coordinates, as was done above in the $NpH$ equations of motion for the ionic degrees of freedom. For the electrons, the transformation from crystal coordinates, $s$, to
Cartesian coordinates, \( r \), is completely analogous and is given by \( r = hs \) or \( s = h^{-1}r \). Since the Jacobian for this transformation is given by \( \det (dr/ds) = \det (h) = V \), the relationship between a KS orbital in Cartesian and crystal coordinates is given by \( \phi_i(r) = \phi_i(s)/\sqrt{V} \), from which it follows that

\[
\rho_{ij}(r) = \frac{1}{V} \rho_{ij}(s). \tag{2.15}
\]

Using this expression and the fact that \( dr = V ds \), we can transform Eq. (2.14) into crystal coordinates as follows:

\[
\sigma_{aa}^{xx} = -\sum_{ij} \frac{\partial}{\partial h_{aa}} \int ds \int ds' \rho_{ij}(s)\rho_{ij}(s') \frac{|h(s - s')|}{|h|}, \tag{2.16}
\]

in which all factors of \( V \) arising from the transformation of the orbital-dependent potentials and the differentials have canceled. Since the crystal coordinates are independent of the dynamical variables associated with the cell, the only remaining dependence on the cell tensor is in the denominator of the integrand in Eq. (2.16). Letting \( \Delta s = s - s' \), we can perform the derivative in question:

\[
\frac{\partial |h\Delta s|^{-1}}{\partial h_{aa}} = \frac{\partial (\Delta s^T h^T h \Delta s)^{-1/2}}{\partial h_{aa}} = -\sum_{\beta} \frac{(\Delta s^\alpha h_{a\beta} \Delta s_{\beta})}{|h\Delta s|^3}. \tag{2.17}
\]

Plugging this expression back into Eq. (2.16) and transforming back to Cartesian coordinates yields

\[
\sigma_{aa}^{xx} = \sum_{ij} \sum_b (h^{-1})_{ab} \int dr \int dr' \rho_{ij}(r)\rho_{ij}(r') \frac{\Delta r_b \Delta r_a}{|\Delta r|^3}, \tag{2.18}
\]

in which \( \Delta r = r - r' \). Further reduction of this expression requires splitting \( \Delta r_b = r_b - r'_b \) across two integrals, using the fact that \( \Delta r' = r' - r = -\Delta r \), and then noticing that both of these integrals are equivalent upon swapping the \( r \) and \( r' \) dummy
variables, which allows us to write:

\[ \sigma^{aa}_{xx} = 2 \sum_{ij} \sum_{b} (h^{-1})_{ab} \int d\mathbf{r} \int d\mathbf{r}' \rho_{ij}(\mathbf{r}) \rho_{ij}(\mathbf{r}') \frac{r_{b} \Delta r_{a}}{|\Delta \mathbf{r}|^3}. \]  

(2.19)

This expression can be further simplified by writing:

\[ \sigma^{aa}_{xx} = 2 \sum_{ij} \sum_{b} (h^{-1})_{ab} \int d\mathbf{r} r_{b} \rho_{ij}(\mathbf{r}) \left[ \int d\mathbf{r}' \rho_{ij}(\mathbf{r}') \frac{\Delta r_{a}}{|\Delta \mathbf{r}'|^3} \right], \]  

(2.20)

and then realizing that the term inside of the square brackets is precisely (to within a sign) the derivative of \( v_{ij}(\mathbf{r}) \) with respect to the \( a \)-th Cartesian component, \( i.e., \)

\[ \frac{\partial v_{ij}(\mathbf{r})}{\partial r_{a}} = \frac{\partial}{\partial r_{a}} \int d\mathbf{r}' \rho_{ij}(\mathbf{r}') \frac{r_{a} - r'_{a}}{|\mathbf{r} - \mathbf{r}'|^3}. \]  

(2.21)

This provides us with the following working expression for the cell derivatives required in EXX-based calculations:

\[ \sigma^{aa}_{xx} = -2 \sum_{ij} \sum_{b} (h^{-1})_{ab} \int d\mathbf{r} r_{b} \rho_{ij}(\mathbf{r}) \left( \frac{\partial v_{ij}(\mathbf{r})}{\partial r_{a}} \right). \]  

(2.22)

From Eq. (2.22), it is clear that once \( v_{ij}(\mathbf{r}) \) is evaluated (which is required for \( E_{xx} \) and \( D^{i}_{xx}(\mathbf{r}) \)), its first derivative provides the EXX contribution to the stress tensor, again highlighting the central role played by \( v_{ij}(\mathbf{r}) \) in EXX calculations.

While the full stress tensor is required during constant-pressure CPMD simulations with anisotropic cells, not all entries of the stress tensor are required for modeling isotropic systems, such as liquids and solids with cubic simulation cells, or systems that might be unstable with respect to shear. In a simple cubic cell, for example, \( h_{aa} = L \delta_{aa} \) (in which \( L \) is the side length) and \( \sigma^{aa}_{xx} \) takes on the following simplified
form (cf. Eq. (2.16)):

\[
\sigma_{xx}^{\alpha\alpha} = -\delta_{\alpha\alpha} \sum_{ij} \frac{\partial}{\partial L} \int ds \int ds' \rho_{ij}(s) \rho_{ij}(s') \frac{L}{|s - s'|} \\
= \frac{\delta_{\alpha\alpha}}{L} \sum_{ij} \int dr \int dr' \rho_{ij}(r) \rho_{ij}(r') \frac{1}{|r - r'|} = -\frac{E_{xx}}{L} \delta_{\alpha\alpha},
\]

(2.23)

since \( r = Ls \). Hence, the EXX contribution to \( \Pi^{(e)} \) in Eq. (2.11) is given by

\[
(\Pi_{xx}^{(e)})_{ab} = -\frac{1}{V} \sum_{\alpha} \sigma_{xx}^{\alpha\alpha} h_{\alpha} = \frac{E_{xx}}{V} \delta_{ab},
\]

(2.24)

from which it follows that the EXX contribution to the internal pressure also takes on the following simple form:

\[
\Pi_{xx} = \frac{1}{3} \text{Tr} \Pi^{(e)}_{xx} = \frac{E_{xx}}{V}.
\]

(2.25)

As such, the EXX contributions to \( \Pi^{(e)} \) and \( p \) are trivial and only require evaluation of \( E_{xx} \).

### 2.2.3 Linear Scaling via Orbital Localization

The efficient evaluation of \( v_{ij}(r) \)—which is a required ingredient for computing \( E_{xx} \), \( D_{xx}^{i}(r) \), and \( \sigma_{xx} \)—is key to enabling large-scale condensed-phase AIMD simulations at the hybrid DFT level of theory. In this section, we will describe a formally exact and linear-scaling EXX method that exploits the natural sparsity of the quantum mechanical exchange interaction in real space via the use of a localized (MLWF) representation of the occupied orbitals. Within this framework, \( \tilde{v}_{ij}(r) \) (which is the MLWF analog of \( v_{ij}(r) \) introduced in Eq. (2.30) below) only needs to be computed for overlapping pairs of MLWFs on a real-space domain that is independent of the system size, thereby paving the way to a linear-scaling EXX method in the condensed phase
(see Chapter 3 for algorithmic details). As such, the cornerstone of our method is the efficient real-space evaluation of $\tilde{v}_{ij}(\mathbf{r})$, which is accomplished herein via the solution of Poisson’s equation on a system-size independent domain for each overlapping MLWF pair.

For a finite-gap condensed-phase system, the occupied KS orbitals (or bands) can be mapped via an orthogonal transformation onto a unique set of MLWFs, $\tilde{\phi}_i(\mathbf{r}) = \sum_j U_{ij} \phi_j(\mathbf{r})$, that are exponentially localized in real space and have a significantly smaller support than the entire simulation cell. As such, the MLWF representation of the occupied space allows us to exploit the underlying sparsity in the quantum mechanical exchange interaction and provides a theoretical and computational framework for substantially reducing the computational scaling and cost associated with EXX-based approaches.

To see how MLWFs can be leveraged to attain a linear-scaling EXX algorithm, we first transform the canonical KS $E_{xx}$ expression in Eq. (1.13) into the MLWF representation. Since $\phi_j(\mathbf{r}) = \sum_i (U^{-1})_{ji} \tilde{\phi}_i(\mathbf{r})$ (cf. Eq. (2.5)), $E_{xx}$ can be written as follows:

$$E_{xx} = - \sum_{ij} \sum_{kk'k''} \int \! d\mathbf{r} \int \! d\mathbf{r}' \frac{\tilde{\phi}_k(\mathbf{r})\tilde{\phi}_{k'}(\mathbf{r}')\tilde{\phi}_{k''}(\mathbf{r})\tilde{\phi}_{k'''}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \times (U^{-1})_{ik}(U^{-1})_{jk'}(U^{-1})_{jk''}(U^{-1})_{ik''}.$$  

(2.26)

Utilizing the fact that $UU^T = UU^{-1} = I$ for an orthogonal matrix, summation over $i$ and $j$ in this expression leads to $\sum_{ij}(U^{-1})_{ik}(U^{-1})_{jk'}(U^{-1})_{jk''}(U^{-1})_{ik''} = \sum_i U_{ki}(U^{-1})_{ik''} \sum_j U_{k'j}(U^{-1})_{jk''} = \delta_{kk''} \delta_{kk''}$, from which we see that

$$E_{xx} = - \sum_{ij} \int \! d\mathbf{r} \int \! d\mathbf{r}' \frac{\tilde{\phi}_i(\mathbf{r})\tilde{\phi}_j(\mathbf{r}')\tilde{\phi}_j(\mathbf{r})\tilde{\phi}_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}$$

(2.27)
upon dummy variable substitutions of \( k \rightarrow i \) and \( k' \rightarrow j \). This proof demonstrates that the expression for evaluating \( E_{xx} \) is invariant to the orthogonal transformation between the KS and MLWF representations. In fact, this invariance property of \( E_{xx} \) also holds for any arbitrary orbital representation \( \{ \psi_i(r) \} \) that is derived from an orthogonal rotation \( U' \) within the occupied KS subspace (i.e., \( \psi_i(r) = \sum_j U'_{ij} \phi_j(r) \)). In analogy to Eq. (2.1), the MLWF expression for \( E_{xx} \) in Eq. (2.27) can also be written in the following compact form:

\[
E_{xx} = -\sum_{ij} \int dr \, \tilde{\rho}_{ij}(r) \tilde{v}_{ij}(r),
\]

(2.28)

in terms of the MLWF-product density,

\[
\tilde{\rho}_{ij}(r) \equiv \tilde{\phi}_i(r) \tilde{\phi}_j(r) = \tilde{\rho}_{ji}(r),
\]

(2.29)

and the corresponding MLWF-product potential,

\[
\tilde{v}_{ij}(r) \equiv \int dr' \, \frac{\tilde{\rho}_{ij}(r')}{|r - r'|} = \tilde{v}_{ji}(r).
\]

(2.30)

We note in passing that while \( E_{xx} \) is invariant to any orthogonal transformation, the values of \( \tilde{\rho}_{ij}(r) \) and \( \tilde{v}_{ij}(r) \)—despite the fact that they have the same expression as that given in Eqs. (2.2) and (2.3)—depend on the employed representation. It is this freedom in the choice of the orthogonal transformation that allows one to select an appropriate localized orbital representation (e.g., MLWF) for exploiting the underlying sparsity in the quantum mechanical exchange interaction. Throughout the first part of this thesis, we will dress each MLWF-specific quantity with a tilde to distinguish them from their analogous expressions in the canonical KS representation.

Given the expression for \( E_{xx} \) in the MLWF representation (cf. Eqs. (2.27)–(2.28)), the corresponding EXX contributions to the wavefunction forces and cell derivatives
that are required to propagate the constant-pressure CPMD equations (Eqs. (2.6)–(2.8)) can be derived following the same procedures given above in Sec. 2.2.3. In this regard, the wavefunction force on the $i$-th MLWF, $\tilde{D}_{xx}^i(r) = -(\delta E_{xx}/\delta \tilde{\phi}_i(r))$, can be obtained from Eqs. (2.27)–(2.30) as:

$$\tilde{D}_{xx}^i(r) = \sum_j \tilde{v}_{ij}(r) \tilde{\phi}_j(r) \equiv \sum_j \tilde{D}_{xx}^{ij}(r),$$

(2.31)

where $\tilde{D}_{xx}^{ij}(r)$ has been defined as the action of $\tilde{v}_{ij}(r)$ on $\tilde{\phi}_j(r)$. Here, $\tilde{D}_{xx}^i(r)$ and $\tilde{D}_{xx}^{ij}(r)$ also depend on the MLWF representation and therefore take on different values when compared to their KS analogs (i.e., $D_{xx}^i(r)$ and $D_{xx}^{ij}(r)$ in Eq. (2.13)). On the other hand, the EXX contribution to the cell derivatives, $\sigma_{xx}^{aa} = (\partial E_{xx}/\partial h_{aa})$, are independent of the orbital representation (due to the fact that both $E_{xx}$ and $h$ are invariant to orthogonal transformations of the occupied space), and are given by:

$$\sigma_{xx}^{aa} = -2 \sum_{ij} \sum_b (h^{-1})_{ab} \int d\mathbf{r} r_b \tilde{\rho}_{ij}(\mathbf{r}) \left( \frac{\partial \tilde{v}_{ij}(\mathbf{r})}{\partial r_a} \right).$$

(2.32)

From Eqs. (2.28), (2.31), and (2.32), it is again clear that evaluation of the MLWF-product potential, $\tilde{v}_{ij}(\mathbf{r})$, is the cornerstone of our MLWF-based EXX approach.

With all of the expressions required for the evaluation of $E_{xx}$, $\tilde{D}_{xx}^i(r)$, and $\sigma_{xx}$ in hand, we will now discuss in detail how MLWFs lead to a linear-scaling EXX algorithm by exploiting the underlying sparsity in the exchange interaction. Since the set of MLWFs are exponentially localized in real space and therefore have a significantly smaller support than the entire simulation cell, this allows us to exploit two levels of sparsity during the computational evaluation of all required EXX-related quantities. The first level of computational savings originates from the fact that a given MLWF, $\tilde{\phi}_i(\mathbf{r})$, will only appreciably overlap with a number, $\tilde{n}_i$, of neighboring MLWFs. For all other MLWFs, the product density, $\tilde{\rho}_{ij}(\mathbf{r}) = \tilde{\phi}_i(\mathbf{r})\tilde{\phi}_j(\mathbf{r}) = 0$, and hence the
corresponding product potential, \( \tilde{v}_{ij}(r) = 0 \). In these cases, the contributions to \( E_{xx} \), \( \tilde{D}_{xx}^i(r) \), and \( \sigma_{xx} \) are formally zero, thereby directly reducing the number of terms that are required in the summation over \( j \) in the evaluation of Eqs. (2.28), (2.31), and (2.32). As such, the number of EXX pair interactions per orbital becomes independent of system size (assuming a fixed number density), which reduces the total number of orbital pairs, \( N_{\text{pair}} \), from \( O(N_o^2) \) to \( O(N_o) \), i.e., \( N_{\text{pair}} = N_o(N_o + 1)/2 \implies \tilde{n}N_o \). In this last expression, \( \tilde{n} = \max_i \{\tilde{n}_i\} < N_o \) is independent of the system size, hence \( \tilde{n}N_o \) represents an upper bound to the number of EXX pair interactions in our approach. Since the contributions from the omitted MLWF pairs to \( E_{xx} \), \( \tilde{D}_{xx}^i(r) \), and \( \sigma_{xx} \) are zero, this reduction in \( N_{\text{pair}} \) still allows for a formally exact evaluation of all EXX-related quantities. Although this leads to significant computational savings, the overall scaling associated with evaluating these quantities is still formally quadratic with respect to system size as the real-space domain associated with the simulation cell, \( \Omega \), grows linearly with the size of the system.

To achieve linear scaling with system size, one can further exploit the fact that the set of exponentially localized MLWFs have a substantially smaller support than \( \Omega \), which allows us to employ real-space domains that are independent of system size and still maintain a formally exact evaluation of \( E_{xx} \), \( \tilde{D}_{xx}^i(r) \), and \( \sigma_{xx} \). To harness this second level of computational savings, we define an MLWF-orbital domain as \( \Omega_i = \{r \in \Omega \mid |\tilde{\phi}_i(r)| > 0\} \), which encompasses the support of \( \tilde{\phi}_i(r) \) in real space. As such, this domain is focused around the so-called MLWF center, \( \tilde{C}_i \), which is given by the expectation value (or first moment) of \( r \), i.e., \( \tilde{C}_i = \langle \tilde{\phi}_i | r | \tilde{\phi}_i \rangle = \int dr \ r \tilde{\rho}_{ii}(r) \). The MLWF-product density domain is therefore defined as \( \Omega_{ij} = \Omega_i \cap \Omega_j = \{r \in \Omega \mid |\tilde{\rho}_{ij}(r)| > 0\} \), which encompasses the support of \( \tilde{\rho}_{ij}(r) \), i.e., the points in real space where \( \tilde{\phi}_i(r) \) and \( \tilde{\phi}_j(r) \) are both non-negligible. When \( i = j \), the corresponding center of charge for \( \tilde{\rho}_{ii}(r) \) is \( \tilde{C}_{ii} = \int dr \ r \tilde{\rho}_{ii}(r) / \int dr \tilde{\rho}_{ii}(r) = \tilde{C}_i \), which is simply the center of the \( i \)-th MLWF since \( \tilde{\rho}_{ii}(r) \) integrates to unity. For \( i \neq j \), \( \tilde{\rho}_{ij}(r) \) corresponds
to a localized charge distribution with a vanishing monopole (i.e., \( Q_{00} = 0 \)) due to the orthogonality of the MLWFs, and the center of this charge distribution cannot be analogously defined as \( \int d\mathbf{r} \mathbf{r} \tilde{\rho}_{ij}(\mathbf{r}) / \int d\mathbf{r} \tilde{\rho}_{ij}(\mathbf{r}) \). As such, we utilize the standard gauge in molecular quantum mechanics for an electrically neutral system, wherein the “center of charge” is taken as the position at which the nuclear (ionic) dipole moment vanishes, to define \( \tilde{C}_{ij} = \int d\mathbf{r} |\tilde{\rho}_{ij}(\mathbf{r})| / \int d\mathbf{r} |\tilde{\rho}_{ij}(\mathbf{r})| \) as the corresponding center for \( \tilde{\rho}_{ij}(\mathbf{r}) \). By making all sectors of this charge distribution positive, \( |\tilde{\rho}_{ij}(\mathbf{r})| \) now has a sizable monopole and a well-defined center of charge given by \( \tilde{C}_{ij} \). By construction, this choice of gauge recovers the correct center of charge when \( i = j \), i.e., \( \tilde{C}_{ii} = \tilde{C}_i \), and is therefore consistent with the expression used above for \( \tilde{\rho}_{ii}(\mathbf{r}) \).

Within this framework, both \( \Omega_i \) and \( \Omega_{ij} \) are system-size independent and substantially smaller than \( \Omega \). Furthermore, since \( \Omega_{ij} \) is defined as the overlapping region between two exponentially decaying MLWFs, \( \tilde{\phi}_i(\mathbf{r}) \) and \( \tilde{\phi}_j(\mathbf{r}) \), the extent of this domain is smaller than both \( \Omega_i \) and \( \Omega_j \), which holds true even when \( i = j \). From Eqs. \( (2.28), (2.31), \) and \( (2.32) \), one sees that a formally exact evaluation of \( E_{xx} \) and \( \sigma_{xx} \) only requires that the spatial integral be performed over \( \Omega_{ij} \), while \( \tilde{D}_{ij}^{xx}(\mathbf{r}) \) requires the action of \( \tilde{v}_{ij}(\mathbf{r}) \) over \( \Omega_j \). This implies that one only needs the corresponding MLWF-product potential, \( \tilde{v}_{ij}(\mathbf{r}) \) in Eq. \( (2.30) \), on \( \Omega_{ij} \) for a numerically exact evaluation of \( E_{xx} \) and \( \sigma_{xx} \), and on \( \Omega_j \) for \( \tilde{D}_{ij}^{xx}(\mathbf{r}) \). As such, the evaluation of \( \tilde{v}_{ij}(\mathbf{r}) \) can also be restricted to system-size-independent real-space domains, despite the fact that this quantity is formally non-zero across \( \Omega \) and asymptotically goes as \( 1/r \) (for \( i = j \) in the case of the bare Coulomb operator; for \( i \neq j \) the potential will decay faster). This leads to even further computational savings as \( \tilde{v}_{ij}(\mathbf{r}) \) can be obtained exactly by solving Poisson’s equation (PE) over \( \Omega_{ij} \) in the near field,

\[
\nabla^2 \tilde{v}_{ij}(\mathbf{r}) = -4\pi \tilde{\rho}_{ij}(\mathbf{r}) \quad \mathbf{r} \in \Omega_{ij},
\]
subject to Dirichlet boundary conditions defined by a multipole expansion (ME) of \( \tilde{\rho}_{ij}(r) \) in the far field, \textit{i.e.},

\[
\tilde{v}_{ij}(r) = 4\pi \sum_{lm} \frac{Q_{lm}}{2l+1} \frac{Y_{lm}(\theta, \varphi)}{r^{l+1}} \quad r \notin \Omega_{ij},
\]

(2.34)

where \( \tilde{C}_{ij} \) is taken as the origin, \( r = (r, \theta, \varphi) \) is given in spherical polar coordinates, \( Y_{lm}(\theta, \varphi) \) are the spherical harmonics, and

\[
Q_{lm} = \int_{\Omega_{ij}} d\mathbf{r} Y_{lm}^*(\theta, \varphi) r^l \tilde{\rho}_{ij}(r),
\]

(2.35)

are the multipole moments corresponding to \( \tilde{\rho}_{ij}(r) \). We note in passing that the ME in Eq. (2.34) serves a dual purpose and will also be employed during the evaluation of \( \tilde{D}_{xx}^{ij}(r) \), which requires \( \tilde{v}_{ij}(r) \) on \( \Omega_j \). In other words, \( \tilde{D}_{xx}^{ij}(r) \) is computed with \( \tilde{v}_{ij}(r) \) on \( \Omega_{ij} \) \textit{via} the solution to the PE in Eq. (2.33), and \( \tilde{v}_{ij}(r) \) on the \( \Omega_j \setminus \Omega_{ij} \) domain, \textit{i.e.}, for all points in \( \Omega_j \) that are not contained in \( \Omega_{ij} \), \textit{via} the ME in Eq. (2.34). As stated in Ref. [75] it is typically sufficiently to include terms in ME (Eqs. 2.34–2.35) up to \( l = 6 \) in most applications such as liquid water systems.

This discussion again clearly highlights that an efficient real-space evaluation of \( \tilde{v}_{ij}(r) \)—on compact and system-size-independent domains—is the cornerstone of our linear-scaling and formally exact MLWF-based EXX approach. In the next chapter, we will focus our discussion on the algorithmic implementation of this approach, which can be used to perform large-scale condensed-phase AIMD simulations at the hybrid DFT level of theory.

2.3 Conclusions

In this chapter, I have presented the computational challenge associated with using conventional cubic-scaling reciprocal-space hybrid DFT algorithm to study large-scale
condensed-phase systems and the theoretical framework to overcome this challenge. This theoretical framework is discussed in the context of CPMD simulation within the \textit{NpT} ensemble which allows realistic sampling under common experimental thermodynamic conditions; this framework can also be easily adapted for other relevant statical mechanical samplings such as \textit{NpH}, \textit{NVT}, and \textit{NVE} ensembles, as well as doing variable-cell structure optimizations. The computational efficiency of this theoretical framework relies on the ability to choose a localized representation for the occupied orbitals. Such representation is guaranteed by the existence of a finite band gap which covers the majority of the molecular crystals and liquids in general. In this case, we choose the MLWFs among the available localized representations, since MLWFs can be generated on-the-fly during a CPMD simulation using a second-order damped dynamics scheme. Within the MLWF representation, this theoretical framework exploits orbital localization in real space for two levels of savings—coming from the natural sparsity of the quantum mechanical exchange interactions—during the computational evaluation of all required EXX-related quantities. The first saving is the ability to consider only overlapping MLWF pairs using a system-size-independent radial distance between centers of each MLWF pairs. The second saving is the existence of, again, a system-size-independent cover (within a rigid translation) for the region of MLWF pair overlap. Taken together both savings, this theoretical framework provides a linear-scaling algorithm that shows the promise of enabling hybrid DFT based AIMD simulations for large-scale condensed-phase systems. In addition, we derived the expression of all required EXX-related quantities—wavefunction forces and stress tensor—to propagate constant-pressure CPMD simulation. From this derivation, we identified the central importance of the MLWF-product potential \( \tilde{v}_{ij}(\mathbf{r}) \)—obtained from the solution of the PE, \( \nabla^2 \tilde{v}_{ij}(\mathbf{r}) = -4\pi \tilde{\rho}_{ij}(\mathbf{r}) \)—in evaluating the EXX contributions relevant to hybrid DFT based CPMD equations of motion.
Based on this theoretical foundation, we will discuss practical implementation of this theoretical framework in Chapter 3.
Chapter 3

Algorithmic Framework: the exx Module

In this chapter, we describe the implementation of our linear-scaling MLWF-based EXX algorithm in the CP module of QE. This algorithm has been implemented as a standalone module named exx which is integrated with an MLWF-enabled semi-local DFT module via a portable input/output interface (see flowchart in Fig. 3.1). At each CPMD step, the main input required for exx includes the current set of MLWFs, \( \{ \tilde{\phi}_i(\vec{r}) \} \), while the output produced by this module includes \( E_{\text{xx}} \), \( \tilde{D}_{\text{xx}}(\vec{r}) \), and \( \sigma_{\text{xx}} \). As such, adaptation of this module to other periodic DFT codes is straightforward, once the capability to produce MLWFs “on-the-fly” during CPMD simulations is available (\textit{vide infra}). In fact, the current exx module only requires that the input orbitals are sufficiently local and form an orthonormal set, and can therefore accommodate (with appropriate modifications) other orbital localization schemes such as RSB and SCDM. To enable large-scale EXX-based AIMD using this approach, we employ a hybrid message-passing interface (MPI) and open multi-processing (OpenMP) parallelization scheme that allows us to differentially exploit both internode and intranode computational resources provided by massively parallel (super)computer ar-
chitectures. In addition to the shared-memory paradigm in OpenMP, this massively parallel MPI-based approach requires specific data distribution schemes (described below) to minimize communication overhead and maximize computational efficiency, thereby allowing us to treat systems with 1000s of atoms with hybrid DFT. To provide the big picture on top of the many technical implementation details, the main results of this chapter are summarized below:

1. Since the \texttt{exx} module depends on a localized orbital representation, we describe the MLWF module in \texttt{CP} as a prerequisite.

2. In view of efficiently using massively parallel supercomputer architectures, we introduce a custom data distribution scheme such that each MPI process has complete real-space information of its assigned orbitals.

3. We then describe how workload is distributed to 1000s of processes with a static load-balancing algorithm.

4. We also introduce reusable auxiliary real-space subdomains to effectively exploit the localization of MLWFs for both communication and computation.

5. We also discuss how the \texttt{exx} module handles general (non-orthorhombic) and fluctuating (during constant-pressure simulations) unit cells via using an effective approach and lattice symmetry constraints, respectively.

6. At the end of the \texttt{exx} module, all EXX contributions required to propagate the constant-pressure CPMD simulation at hybrid DFT level are obtained with compatible data distribution scheme to the \texttt{QE} convention.

### 3.1 Prerequisites

To start a CPMD (or BOMD) simulation, one needs to reach the electronic ground state for a given initial configuration of the system via a self-consistent field (SCF) cal-
culation. In the CP module of QE, the iterative solution of the non-linear KS equations is accomplished using either conjugate gradient (CG) or second-order damped dynamics (SODD) to minimize the fictitious kinetic energy associated with the electronic degrees of freedom while keeping the cell and ions fixed. In doing so, large-scale and cubic-scaling matrix operations such as diagonalization of the Fock matrix are completely sidestepped during the SCF procedure; as such, this approach does not require unoccupied (virtual) states and provides a solid foundation upon which one can build a fully linear-scaling DFT (or HF) code base. In fact, this CP-like approach to the SCF solution of the KS equations can also be carried out within the MLWF framework by performing a nested SODD minimization of the Marzari-Vanderbilt functional to incrementally localize the occupied orbitals between each SCF step. This approach leads to the desired set of MLWFs upon reaching SCF convergence, and therefore provides a cost-effective alternative to the standard a posteriori procedure of localizing the canonical (Bloch) occupied orbitals from a fully converged SCF calculation.

To reach a SCF with our MLWF-based EXX approach that avoids performing a preliminary (and computationally expensive) EXX calculation in reciprocal space, we adopt this CP-like approach to incrementally localize the occupied orbitals during the SCF procedure. Since these incrementally localized orbitals are not equivalent to the final set of MLWFs at a given SCF step, the orbital-dependent EXX contributions to \( v_{xc} \) are approximately evaluated via the aforementioned sparse computation of \( D_{xx}^i(r) \) in Eq. (2.31). Such errors are once again inconsequential as the incremental refinement of the localized orbitals (and therefore \( D_{xx}^i(r) \)) with each step leads to the desired set of MLWFs upon SCF convergence. We note in passing that our MLWF-based EXX approach is based on incremental “on-the-fly” refinements of the localized orbitals during the SCF procedure, and is therefore not suitable for standard diagonalization-based SCF methods. In the latter case, global rotations of the occupied and virtual
subspaces lead to marked delocalization of the occupied orbitals, and therefore require significant orbital relocalization (essentially from scratch) at each SCF step.

In practice, EXX-based SCF calculations in CP take advantage of the incremental nature of the aforementioned MLWF refinement process by starting the KS iterations with a relatively inexpensive semi-local xc functional (e.g., PBE\cite{PBE} for PBE0\cite{PBE0}), which stabilizes the total charge density, $\rho(\mathbf{r})$, and incrementally reduces the sum of the spreads associated with each occupied orbital. Once the semi-local DFT iterations reach $10 \times$ the target SCF convergence threshold, the orbitals are typically quite localized and closely resemble the final set of MLWFs corresponding to the semi-local xc functional. At this point, the exx module is activated to perform the remaining steps required to reach SCF convergence, upon which one obtains the corresponding final set of MLWFs at the hybrid DFT level. For all systems tested, this approach has a significantly reduced computational cost when compared to localizing the occupied orbitals (from a fully converged canonical PBE calculation) from scratch and then using these orbitals to perform a PBE0 SCF calculation to convergence. We note in passing that the computational cost associated with the initial SCF procedure is completely negligible when compared to the overall CPMD simulation, which is the focus of this work. Further optimization of this incremental SCF procedure to enable high-throughput hybrid DFT-based single-point energy evaluations on large-scale condensed-phase systems as well as linear-scaling EXX-based BOMD will be addressed in future work.\cite{futureWork}

In this chapter, we have extended this incremental localization procedure to full-fledged CPMD simulations in which the ionic and cell degrees of freedom are propagated according to the appropriate equations of motion. In this case, the MLWFs are refined via a small number of SODD iterations (3–4 in the case of the liquid water examples that will be shown in Sec. 4.2.1) between each CPMD step, which results in minimal computational overhead when compared to the cost of the EXX calculation.
Figure 3.1: Flowchart of the exx module (dashed green box) in CP. As described in the main text, the input required by this module includes the current set of MLWFs, $\{\tilde{\phi}_i(r)\}$, at each CPMD step. The output produced by exx includes the EXX energy ($E_{xx}$), the EXX contribution to the wavefunction forces ($\{\tilde{D}_{xx}^i(r)\}$), and the EXX contribution to the cell forces ($\sigma_{xx}$). Purple (brown) circles indicate that a given quantity is represented according to the GRID (ORBITAL) data distribution scheme (see Fig. 3.2), while the pale yellow circles represent data that are globally broadcasted. The steps outlined in red represent an loop over non-redundant (overlapping) MLWF pairs, and contain MLWF communication, PE solution, as well as computation of all EXX-related quantities. The $\overrightarrow{r}$ notation indicates local (relative) Cartesian coordinates in a given subdomain. For a detailed description of each step in the exx modules, see Secs. 3.3.1–3.3.6.
Moving forward, this incremental localization can actually be completely avoided by using the field-theoretic approach proposed by Tuckerman and coworkers, which introduces additional fictitious dynamics for a set of gauge fields to enable “on-the-fly” propagation of the orthogonal MLWF transformation ($U$) matrix. Implementation of this approach would reduce the (relatively small) number of SODD iterations between each CPMD step.

### 3.2 Data Distribution Schemes

As mentioned above, we employ a hybrid MPI/OpenMP parallelization scheme to enable large-scale EXX-based AIMD on massively parallel (super)computer architectures containing 1000s of nodes. Our algorithm, which is described in Sec. 3.3 below, is primarily based upon the MPI distributed-memory paradigm, which requires specific data distribution schemes to minimize communication overhead and maximize computational efficiency. During a GGA-based CPMD simulation in QE, the orbitals, charge density, and potential are constantly transformed between real- and reciprocal-space via the `fwdFFT` and `invFFT` operations. With all real-space quantities numerically represented on a grid (mesh) that is discretized along the corresponding lattice vectors, QE employs the GRID data distribution scheme to scatter these quantities across $N_{\text{proc}}$ MPI processes (ranks). In the GRID data distribution scheme (see Fig. 3.2), the real-space grid is partitioned into $N_{\text{slab}}$ slabs along the $L_3$ axis. Assuming that $N_{\text{proc}} = N_{\text{slab}}$, each MPI process will hold the data corresponding to all distributed real-space quantities on one slab of the real-space grid. In doing so, this data distribution scheme facilitates efficient parallel FFT by dividing the 3D FFT into a set of 2D FFTs (each of which can be executed by a given MPI process with a given slab) and then performing a 1D FFT along the direction of the slab partition. As depicted in Fig. 3.1, the input to the `exx` module in QE includes the current set of MLWFs,
Figure 3.2: Schematic illustration of the GRID and ORBITAL data distribution schemes in QE. For simplicity, consider a system consisting of a single water molecule with $N_o = 4$ MLWFs ($\tilde{\phi}_i(r)$), a simulation cell consisting of a real-space simple-cubic grid that has been partitioned into $N_{\text{slab}} = 4$ slabs along the $z$-direction ($Z_i$), and a pool of $N_{\text{proc}} = 4$ MPI processes ($P_i$). As depicted at the top of this figure, each of these MPI processes (and the corresponding data it holds in local memory) is assigned a color: $P_1$ (red), $P_2$ (green), $P_3$ (blue), and $P_4$ (yellow). As input into the exx module, the $\tilde{\phi}_i(r)$ are provided in the GRID data distribution scheme, in which a given MPI process, $P_i$, holds the data corresponding to all $N_o$ MLWFs on one slab, $Z_i$, of the real-space grid. During Step I of the exx module (Sec. 3.3.1), the $\tilde{\phi}_i(r)$ are redistributed according to the ORBITAL scheme, in which a given MPI process, $P_i$, holds the data corresponding to only one MLWF, $\tilde{\phi}_i(r)$, across all $N_{\text{slab}}$ slabs of the real-space grid. As described in (Secs. 3.3.2–3.3.5), Steps II–V involve selective communication of the $\tilde{\phi}_i(r)$ between MPI processes and computation of all EXX-related quantities ($E_{\text{xx}}$, $\sigma_{xx}$, and $\{\tilde{D}_{\text{xx}}^i(r)\}$). At the end of Step V, the $\{\tilde{D}_{\text{xx}}^i(r)\}$ are stored according to the ORBITAL scheme and are then redistributed into the GRID scheme during Step VI (Sec. 3.3.6), the final step of the exx module.

$\{\tilde{\phi}_i(r)\}$, at each CPMD step. These MLWFs are distributed across MPI processes according to the GRID data distribution scheme, in which a given process holds the data corresponding to all MLWFs on a given slab. Although the GRID scheme is convenient for efficient parallel FFT, this data distribution model is far from ideal for an efficient massively parallel implementation of our MLWF-based EXX approach. As such, we have introduced an alternative ORBITAL data distribution scheme (see Fig. 3.2) in QE, in which a given MPI process now holds quantities like $\tilde{\phi}_i(r)$ and $\tilde{D}_{\text{xx}}^i(r)$ for a single MLWF across the entire real-space grid. The details behind the transforma-
tion between the GRID and ORBITAL data distribution schemes are provided below in Secs. 3.3.1 and 3.3.6.

In doing so, the ORBITAL data distribution scheme is particularly suited for our real-space MLWF-based EXX algorithm, since this approach is centered around the concept of orbital sparsity and the efficient evaluation (via the PE) of $\tilde{v}_{ij}(r)$. For one, the ORBITAL scheme allows us to utilize a significantly larger number of MPI processes, $N_{\text{proc}} \gg N_{\text{slab}}$, as the number of overlapping MLWF pairs (which grows linearly with system size) quickly exceeds $N_{\text{slab}}$ (which grows with the cubic root of the system size). The ORBITAL scheme also allows us to exploit intranode parallelization with $N_{\text{thread}}$ OpenMP threads during the most computationally intensive steps in our algorithm, e.g., solving the PE to obtain $\tilde{v}_{ij}(r)$ (see Sec. 3.3.4). As a result, this hybrid MPI/OpenMP parallelization scheme not only provides us with access to even more computational resources during EXX-based simulations, but also allows us to sidestep the prohibitively large data communication overhead associated with an MPI-based solution to the PE.

3.3 Implementation

In this section, we provide a detailed description for each of the steps inside the exx module in QE. Our discussion will follow the flowchart depicted in Fig. 3.1 which takes the current set of MLWFs in real space, $\{\tilde{\phi}_i(r)\}$ (distributed according to the GRID scheme), as input, and outputs $E_{\text{xx}}, \{\tilde{D}_{\text{xx}}^i(r)\}$ (also in the GRID scheme), and $\sigma_{\text{xx}}$.

3.3.1 Step I: Redistribution of MLWFs

Given the current set of MLWFs in real space, $\{\tilde{\phi}_i(r)\}$, which are distributed among MPI processes according to the GRID scheme, the first step in the exx module is the
forward redistribution of these quantities into the ORBITAL data distribution scheme. For this purpose, each MPI process collects an assigned \( \tilde{\phi}_i(r) \) across the entire real-space grid via an all-to-all internode communication step, as shown in Fig. 3.2. This all-to-all communication is performed twice per CPMD step, once here in the forward redistribution of \( \{ \tilde{\phi}_i(r) \} \) from the GRID to the ORBITAL scheme, and once in Step VI in the inverse redistribution of \( \{ \tilde{D}_{xx}^i(r) \} \) from the ORBITAL to the GRID scheme (see Sec. 3.3.6). As will be discussed in Chapter 4, this communication overhead contributes a substantial fraction of the overall cost associated with the exx module, and can be significantly reduced by a more sophisticated communication scheme (in future work) over select subsets of the MPI process pool, i.e., those containing the regions of real-space that overlaps with the MLWF-orbital domain, \( \Omega_i \), introduced above in Sec. 2.2.3.

In the exx algorithm, the assignment of MLWFs to a given MPI process is based on the relative number of MPI processes to MLWFs, which we denote by \( \zeta \equiv N_{\text{proc}}/N_o \). When \( \zeta = 1 \), there is one MPI process per MLWF, and each process, \( P_i \), will be assigned a unique MLWF, \( \tilde{\phi}_i \). With limited computational resources (\( N_{\text{proc}} < N_o \)), \( \zeta < 1 \) and multiple MLWFs are assigned to each process; as such, a balanced distribution of MLWFs across MPI processes is only possible when \( N_{\text{proc}} \) is a divisor of \( N_o \). In the strong scaling limit, our exx algorithm allows for \( \zeta \) to take on integer values greater than one, in which a given MLWF is assigned to multiple processes.

### 3.3.2 Step II: Construction of Pair List and Proto-Subdomains

With the MLWFs distributed among MPI processes according to the ORBITAL scheme, we now explain how the exx module exploits the sparsity of the MLWFs and utilizes system-size independent subdomains of \( \Omega \) during the computation of all EXX-related quantities. To accomplish this goal, we will first describe the construction of the
so-called unique MLWF-pair list, $\mathcal{L}$, which not only contains the relevant set of overlapping MLWF pairs, but also determines how the computational workload associated with these pairs is distributed among the pool of available MPI processes. This is followed by a detailed description of the set of “proto-subdomains” employed in the exx module, which represent computationally efficient alternatives to the formal $\Omega_i$ and $\Omega_{ij}$ subdomains introduced in Sec. 2.2.3. We conclude this section with a discussion of our current strategy regarding the use of such system-size independent subdomains in the presence of cell fluctuations during constant-pressure EXX-based CPMD simulations in the $NpH$ and/or $NpT$ ensembles.

Construction of the MLWF-Pair List

To exploit the first level of computational savings, which originates from the fact that MLWFs are exponentially localized and only overlap with a limited number of neighbors, two MLWFs, $\tilde{\phi}_i(r)$ and $\tilde{\phi}_j(r)$, are considered an overlapping pair if $|\tilde{C}_i - \tilde{C}_j| < R_{\text{pair}}$. A judicious choice for $R_{\text{pair}}$ is required for accurately calculating all EXX-related quantities, and an analysis of the convergence of $E_{\text{xx}}$ with respect to $R_{\text{pair}}$ will be provided in Sec. 4.1.1.

At the current point in the algorithm, each MLWF $\tilde{\phi}_i(r)$ is stored according to the ORBITAL data distribution scheme on one (or more) MPI processes (depending on the value of $\zeta = N_{\text{proc}}/N_o$ employed during runtime, see Sec. 3.3.1). For simplicity, we will discuss the $\zeta = 1$ case first, in which there is one MPI process per MLWF, and each process, $P_i$, is assigned a unique MLWF, $\tilde{\phi}_i(r)$. As such, $P_i$ lacks direct access to $\tilde{\phi}_j(r)$ for $j \neq i$, which is required for the construction of $\tilde{\rho}_{ij}(r)$ and the subsequent computation of $\tilde{v}_{ij}(r)$ for evaluating the $ij$ pair contribution to $E_{\text{xx}}, \sigma_{xx}, \tilde{D}_{xx}^{ij}(r)$, and $\tilde{D}_{xx}^{ji}(r)$. Although $\tilde{\rho}_{ii}(r)$ can be constructed locally on $P_i$ to evaluate the $ii$ (self pair) contribution to each of these quantities, all of the other $ij$ pair contributions will require communication between MPI processes.
To design an MLWF-based EXX algorithm that achieves a minimal time to solution while efficiently utilizing all parallel computational resources, one needs to (i) minimize the total computational workload, (ii) minimize the number of inter-process communication events, and (iii) maintain a balanced workload among the pool of available MPI processes. To accomplish this goal, we now describe the procedure employed in the exx module to construct the so-called unique MLWF-pair list, \( \mathcal{L} \), which defines the computation and communication protocol in our algorithm. To construct \( \mathcal{L} \), the indices corresponding to all overlapping MLWF pairs (as determined by the aforementioned criteria based on \( \mathcal{R}_{\text{pair}} \)) are first assembled into the non-unique MLWF-pair list, \( \mathcal{L}_0 \), which contains all possible permutations \( ij \) and \( ji \) of these overlapping MLWF pairs. Since the \( ij \) and \( ji \) pair contributions to \( E_{xx} \) are equivalent (cf. Eq. (2.28)), it is clear that \( \mathcal{L}_0 \) is redundant and contains twice as many pairs as needed.

Before discussing the procedure used to determine \( \mathcal{L} \), we first demonstrate that exploiting such redundancy within the more parallelizable ORBITAL data distribution scheme leads to the requirement for two interprocess communication events per unique MLWF pair. To see this more clearly, one only needs to consider a minimalistic system which contains a single \( ij \) pair of overlapping MLWFs. Throughout this example, the spatial arguments of all EXX-related quantities will be suppressed (e.g., \( \tilde{\phi}_i \) will be used instead of \( \tilde{\phi}_i(r) \)), since all computation and communication events will be performed using system-size independent subdomains (vide infra). With \( \tilde{\phi}_i \) located on \( P_i \) and \( \tilde{\phi}_j \) located on \( P_j \), first consider the case where the inherent pair redundancy is not exploited. In this case, \( \tilde{\phi}_i \) (\( \tilde{\phi}_j \)) is first communicated to \( P_j \) (\( P_i \)) for a total of two interprocess communication events. At this point, each MPI process constructs the corresponding MLWF-product density, \( \tilde{\rho}_{ij} = \tilde{\rho}_{ji} \), and proceeds to compute \( \tilde{v}_{ij} = \tilde{v}_{ji} \) by solving two equivalent PEs. Since the solution of the PE is the dominant computational step in our EXX algorithm, this will count for a to-
tal of two computation events. With $\tilde{v}_{ij} = \tilde{v}_{ji}$ available on both $P_i$ and $P_j$, each process is now in position to compute the $ij$ and $ji$ contributions to $E_{xx}$, $\sigma_{xx}$, and both $\tilde{D}_{xx}^{ij} = \tilde{v}_{ij} \tilde{\phi}_j$ and $\tilde{D}_{xx}^{ji} = \tilde{v}_{ji} \tilde{\phi}_i$ via Eqs. (2.28), (2.32), and (2.31). As depicted in the flowchart for the exx module in Fig. 3.1, the $\{\tilde{D}_{xx}^{i}\}$ are needed in the ORBITAL data distribution scheme before these quantities are finally redistributed to the GRID scheme to ensure compatibility with the other modules in QE (see Fig. 3.2). As such, the local evaluation of $\tilde{D}_{xx}^{ij}$ on $P_i$ and $\tilde{D}_{xx}^{ji}$ on $P_j$ directly provides these quantities in the requisite ORBITAL data distribution scheme without the need for any additional communication. Hence, the total cost per unique MLWF pair amounts to two units of communication followed by two units of computation.

Since the removal of all MLWF-pair redundancy is crucial for minimizing the total number of computational events (and hence the overall time to solution), we now consider the case where this inherent pair redundancy is exploited. In this case, only $\tilde{\phi}_j$ would be sent to $P_i$ with an associated cost of one unit of communication, and $P_i$ will therefore be solely responsible for computing all EXX-related quantities. Since the $ij$ and $ji$ pair contributions to both $E_{xx}$ and $\sigma_{xx}$ are equivalent, these quantities can be computed on $P_i$ via a single computation event (i.e., the solution to the corresponding PE), and then sent to any other process with minimal communication (i.e., one double-precision number for $E_{xx}$ and $3 \times 3$ double-precision numbers for $\sigma_{xx}$). Although $\tilde{D}_{xx}^{ij} \neq \tilde{D}_{xx}^{ji}$, this also poses no problem as $P_i$ has direct access to $\tilde{\phi}_i$ and $\tilde{\phi}_j$, and hence both $\tilde{D}_{xx}^{ij}$ and $\tilde{D}_{xx}^{ji}$ can be computed locally. With the requirement that the $\{\tilde{D}_{xx}^{i}\}$ are stored in the ORBITAL data distribution scheme, this will incur an additional communication event as $\tilde{D}_{xx}^{ji}$ is shipped back to $P_j$. Hence, exploiting the inherent MLWF-pair redundancy reduces the computational workload by half (as expected), but it does not change the requirement for two communication events per unique MLWF pair.
During the non-redundant evaluation of the \(ij\) and \(ji\) pair contributions in this example, the fact that \(P_j\) was idle while \(P_i\) performed all of the required computations creates an imbalance in the computational workload assigned to each \texttt{MPI} process. With the freedom to assign the computational workload associated with the \(ij\) pair to either \(P_i\) or \(P_j\), the \texttt{exx} module is now tasked with determining how the total computational workload will be distributed among the pool of available \texttt{MPI} processes. Armed with knowledge of the total number of non-unique MLWF pairs in the system (\textit{via} \(L_0\)) as well as the use of system-size independent subdomains to regularize the computational cost associated with the solution to each PE (\textit{vide infra}), the process for doing so involves a static load-balancing algorithm which seeks to minimize the overall time to solution by reducing the imbalances present in the computational workload, and hence the number of idle processes. Although it is certainly possible in the current version of the algorithm, we chose not to involve a third process, \(P_k\), in the evaluation of the \(ij\) pair contribution, as this would introduce two additional communication events associated with sending \(\tilde{\phi}_i\) to \(P_k\) and \(\tilde{D}^{ij}_{xx}\) back to \(P_i\). In this regard, the local computation of \(\tilde{D}^{ij}_{xx}\) on \(P_i\) not only avoids additional unnecessary communication events, but also allows for reduced storage requirements as this quantity can be cumulatively incremented (over multiple \(j\)) within a single array corresponding to \(\Omega_i\).
Algorithm 1 Refinement of $L_0$ to $L$

\begin{verbatim}
any_removal ← TRUE

while any_removal do
    any_removal ← FALSE
    for $i = 1, N_o$ do
        for $j \neq i \in L_0[i]$ do
            if $i \in L_0[j]$ then
                $L_0[j] \leftarrow L_0[j] \setminus \{i\}$
                any_removal ← TRUE
                break
            end if
        end for
    end for
end while

$\mathcal{L} \leftarrow L_0$
\end{verbatim}

This static load-balancing algorithm can be represented by the so-called unique MLWF-pair list, $\mathcal{L}$, the construction of which is described in Algorithm 1 (for the general case) as well as the left panel of Fig. 3.3 (for the illustrative case of a single water molecule \textit{vide infra}). Here, we start with the $L_0$ array, which contains all possible permutations of overlapping MLWF pairs, \textit{i.e.}, $L_0[i]$ (the $i$-th row of $L_0$) is populated with a list of indices, $\{j\}$, corresponding to all $\tilde{\phi}_j$ that overlap with $\tilde{\phi}_i$. For each $i$, the indices $j \in L_0[i]$ are sorted into ascending order based on their vicinity to $\tilde{\phi}_i$ \textit{via} $|\tilde{\mathcal{C}}_i - \tilde{\mathcal{C}}_j|$. By construction, each $L_0[i]$ also contains $i$ (self pair) and will retain this non-redundant index throughout the refinement of $L_0$ to $\mathcal{L}$ in Algorithm 1. While there are still redundant pairs in $L_0$, this algorithm will consecutively sweep over MLWFs to locate redundant pairs such as $ij$ and $ji$; in our approach, this is tantamount to finding both $j \in L_0[i]$ and $i \in L_0[j]$. Once located, the algorithm eliminates
Figure 3.3: Graphical depiction of the unique MLWF-pair list construction process and corresponding MLWF communication scheme in the \texttt{exx} module of a single water molecule.

This redundancy from $L_0$ by removing the index $i$ from $L_0[j]$. At the end of these sweeps, all of the redundancies in $L_0$ are removed and we are left with $L$, the unique MLWF-pair list. This list contains the minimum number of computational tasks required to evaluate all EXX-related quantities and dictates how this computational workload will be distributed among the pool of available MPI processes. By virtue of the ORBITAL data distribution scheme, $L$ also encodes the communication protocol that will be followed throughout the remainder of the \texttt{exx} module (see Sec. 3.3.3). With $\zeta = 1$, this amounts to sending $\tilde{\phi}_j \rightarrow P_i$ and $\tilde{D}_{ij}^{\text{xx}} \rightarrow P_j$ for each unique $ij$ pair, as depicted in the right panel of Fig. 3.3.

As a simple illustrative example for this algorithm shown in Fig. 3.3, we will again consider a single water molecule with $N_o = 4$ MLWFs ($\tilde{\phi}_i$) and a pool of $N_{\text{proc}} = 4$ MPI processes ($P_i$), i.e., $\zeta = 1$. Starting with the non-unique MLWF-pair list, $L_0$, which contains all possible permutations of overlapping MLWF pairs, the step-by-step procedure employed to transform $L_0$ into the unique MLWF-pair list, $L$, is depicted in the left panel. Since all of the MLWFs are mutually overlapping in a single water molecule, $L_0[i]$ contains the indices, $\{j\}$, corresponding to all MLWFs (including $j = i$) sorted according to $|\tilde{C}_i - \tilde{C}_j|$. During the process of reducing $L_0$ to $L$, $i$ is always selected (bold black font) to remain in $L_0[i]$, while the remaining non-unique
indices are shown in gray. In the first sweep (SWEEP01) of Algorithm 1, the next element $j \in L_0[i]$ is selected to remain in $L_0[i]$, while the now redundant index $i$ is removed (red slash) from $L_0[j]$. During the first sweep in this example, the indices 2, 4, 2, and 3 were selected to remain in $L_0[1]$, $L_0[2]$, $L_0[3]$, and $L_0[4]$, respectively, while the corresponding redundant indices, 1, 2, 3, and 4, were removed from $L_0[2]$, $L_0[4]$, $L_0[2]$, and $L_0[3]$. This process is repeated until all of the MLWF-pair redundancy is removed from $L_0$, upon which one is left with the final $L$. For a single water molecule, only two sweeps are required to reach this stage; at this point, all of the unique MLWF pairs have been assigned to a given $L[i]$, and no redundant indices remain. With each MLWF stored according to the ORBITAL data distribution scheme (in which $\tilde{\phi}_i$ is assigned to $P_i$), the final $L$ determines how the computational workload will be distributed among the pool of available MPI processes. Even in this simple example, there exists a mismatch in the number of pairs assigned to each process, with three MLWF pairs assigned to $P_1$ and $P_3$, and only two MLWF pairs assigned to $P_2$ and $P_4$; such discrepancies are expected (even for homogeneous condensed-phase systems) and represent a manageable imbalance in the computational workload. By virtue of the ORBITAL data distribution scheme, $L$ also determines the corresponding MLWF communication protocol, which is depicted in the right panel for the single water molecule. With each $\tilde{\phi}_j$ located on a given $P_j$, the MLWFs are directly communicated to the $P_i$, which has been assigned to evaluate the $ij$ contribution to all EXX-related quantities. With $\tilde{\phi}_j$ in hand, $P_i$ proceeds to form $\tilde{\rho}_{ij}$, solve the corresponding PE for $\tilde{v}_{ij}$, and then compute the $ij$ contribution to $E_{xx}$, $\sigma_{xx}$, $\tilde{D}_{xx}^{ij}$, and $\tilde{D}_{xx}^{ji}$. Once $\tilde{D}_{xx}^{ji}$ is formed, this quantity is sent back to $P_j$. As an example, consider $L[2]$, which contains two indices (2 and 4). As such, $P_2$ is responsible for computations involving the $ij = 22$ and $ij = 24$ pairs. Since $P_2$ already holds $\tilde{\phi}_2$, $\tilde{\rho}_{22}$ (required for computing $\tilde{v}_{22}$) can be constructed locally without the need for any interprocess communication.
To construct $\tilde{\rho}_{24}$ (and hence $\tilde{v}_{24}$), $\tilde{\phi}_4$ is sent to $P_2$ from $P_4$ (solid arrow). Once the corresponding $D^{12}_{xx}$ is formed, it is shipped back to $P_4$ (dashed arrow).

By construction, static load-balancing algorithms (such as Algorithm 1) yield fairly well-balanced workload distributions by mitigating potential imbalances during the refinement of $L_0$ to $L$. Here, we note that the distance-based sorting of the indices in each row of $L_0$ is crucial for avoiding severe workload imbalances; such issues originate from sequential index ordering and can also be avoided by performing random sweeps over row indices. Since the number of overlapping pairs per MLWF changes throughout an MD simulation, such static approaches are not flexible enough to determine an optimal workload balance in all cases. For a detailed discussion regarding the performance of this static load-balancing algorithm during CPMD simulations of liquid water, as well as future possible improvements on this approach, see Sec. 4.2.1.

When computational resources are limited, the exx module can utilize less MPI processes during runtime (i.e., $\zeta < 1$). In this case, multiple MLWFs are contiguously assigned to each process, and a balanced distribution of the workload (within the framework defined by Algorithm 1) is only possible when $N_{\text{proc}}$ is a divisor of $N_o$; as such, this is the current recommended setting whenever applicable (see Sec. 3.3.1). In the isolated water molecule example in Fig. 3.3, the use of $\zeta = 1/2$ would start with $P_1$ holding $\tilde{\phi}_1$ and $\tilde{\phi}_2$, and $P_2$ holding $\tilde{\phi}_3$ and $\tilde{\phi}_4$. After running Algorithm 1 to generate $L$, the workload associated with a given MLWF is mapped onto the process holding this orbital. This results in five units of computation assigned to each MPI process: two self pairs, one local pair (in which both MLWFS are held on the same process), and two non-local pairs (in which one of the MLWFs is held on a different process), e.g., $P_1$ would be responsible for the $ij = 11, 22, 12, 14, 24$ MLWF pairs. In this case, the workload is optimally balanced and the maximum number of computation events per process is only $5/3 \times$ (instead of $2 \times$) larger than $\zeta = 1$. This allows for a
more computationally efficient means to performing an EXX calculation, albeit with a longer time to solution.

With access to massively parallel resources (ζ > 1 ∈ N), each \( \tilde{\phi}_i \) is now duplicated and stored in \( P_i, P_{i+N_o}, \ldots, P_{i+(\zeta-1)N_o} \). After running Algorithm [1] to generate \( \mathcal{L} \), the workload associated with a given MLWF is split into \( \zeta \) parts, each of which is assigned to one of the processes holding this orbital. For the isolated water molecule, the use of \( \zeta = 2 \) (\( \zeta = 1 \)) results in processes assigned with 1−2 (2−3) computational tasks. This reduces the maximum number of computation events per process from 3 to 2, and hence lowers the overall time to solution. However, this gain comes at the expense of increasing the workload imbalance from 1/3 (i.e., processes with the lightest workload idling for \( \approx 1/3 \) of the time) to 1/2, and is therefore a less efficient use of the available computational resources.

**Construction of Proto-Subdomains**

As discussed throughout this work, the efficient evaluation of \( \tilde{v}_{ij}(r) \) is the cornerstone of our MLWF-based EXX approach. To exploit the sparsity of the MLWFs and still retain a formally exact evaluation of \( \tilde{v}_{ij}(r) \), this quantity is computed via the solution to the corresponding PE for all points in \( \Omega \) that are contained in \( \Omega_{ij} \) (see Eq. (2.33)). Since the PE is a boundary-value problem, the required boundary conditions are provided by the ME of \( \tilde{\rho}_{ij}(r) \) about \( \tilde{C}_{ij} \) on the thin shell of the real-space grid surrounding \( \Omega_{ij} \) (see Eqs. (2.34)–(2.35)). By computing \( \tilde{v}_{ij}(r) \) for all \( r \in \Omega_{ij} \), both \( E_{xx} \) and \( \sigma_{xx} \) can be computed exactly (cf. Eqs. (2.28) and (2.32)). However, the exact evaluation of the EXX contribution to the wavefunction forces, \( \tilde{D}_{ij}^{ij}(r) \) and \( \tilde{D}_{ji}^{ji}(r) \), requires \( \tilde{v}_{ij}(r) \) for all \( r \in \Omega_j \) and \( r \in \Omega_i \), respectively (see Eq. (2.31)). Since \( \Omega_{ij} \subset \Omega_j \) and \( \Omega_{ij} \subset \Omega_i \), \( \tilde{D}_{ij}^{ij}(r) \) and \( \tilde{D}_{ji}^{ji}(r) \) are evaluated with \( \tilde{v}_{ij}(r) \) from the solution to the PE for all \( r \in \Omega_{ij} \). For \( r \in \Omega_j \setminus \Omega_{ij} \) and \( r \in \Omega_i \setminus \Omega_{ij} \), \( \tilde{v}_{ij}(r) \) can be conveniently and accurately supplied by a sufficiently converged ME of \( \tilde{\rho}_{ij}(r) \). As such, the ME serves
the dual purpose of providing the necessary boundary conditions for the PE as well as the far-field $\tilde{v}_{ij}(r)$ required for a formally exact computation of both $\tilde{D}_{xx}^{ij}(r)$ and $\tilde{D}_{xx}^{ji}(r)$.

To exploit this second level of computational savings, which originates from the fact that a formally exact evaluation of all EXX-related quantities can be restricted to real-space domains that are system-size independent and significantly smaller than $\Omega$, the exx module introduces an alternative formulation of the $\Omega_i$ and $\Omega_{ij}$ subdomains described directly above. To begin, we first note that subdomains like $\Omega_i$ and $\Omega_{ij} = \Omega_i \cap \Omega_j$ are formally defined as the points in $\Omega$ for which $|\tilde{\phi}_i(r)| > 0$ and $|\tilde{\rho}_{ij}(r)| = |\tilde{\phi}_i(r)\tilde{\phi}_j(r)| > 0$, respectively. As such, both of these subdomains can have irregular and even disjoint shapes. In practice, a subdomain like $\Omega_{ij}$ could be represented by the points in $\Omega$ for which $|\tilde{\rho}_{ij}(r)|$ is larger than some predetermined numerical cutoff, but this is a cumbersome and computationally demanding definition that would require screening substantial sectors of $\Omega$ for each pair of MLWFs during every CPMD step. To combat this issue and still maintain a formally exact evaluation of all required quantities, one could simply utilize two concentric spherical subdomains per $ij$ pair, i.e., $\Theta(\tilde{C}_{ij}, R_{PE}^{ij})$ and $\Theta(\tilde{C}_{ij}, R_{ME}^{ij})$, which are spheres centered at $\tilde{C}_{ij}$ with radii $R_{PE}^{ij}$ and $R_{ME}^{ij}$ chosen to be large enough to encompass $\Omega_{ij}$ and $\Omega_i \cup \Omega_j$, respectively. In doing so, the corresponding PE, $\nabla^2 \tilde{v}_{ij}(r) = -4\pi \tilde{\rho}_{ij}(r)$, could then be solved without any domain truncation error on $\Theta(\tilde{C}_{ij}, R_{PE}^{ij})$, which is significantly smaller than $\Omega$. Computing the ME of $\tilde{\rho}_{ij}(r)$ on the $\Theta(\tilde{C}_{ij}, R_{ME}^{ij}) \setminus \Theta(\tilde{C}_{ij}, R_{PE}^{ij})$ shell would again provide the necessary boundary conditions for the PE as well as the far-field $\tilde{v}_{ij}(r)$ needed for evaluating both $\tilde{D}_{xx}^{ij}(r)$ on $\Omega_j$ and $\tilde{D}_{xx}^{ji}(r)$ on $\Omega_i$. Since both of these subdomains are contained in $\Theta(\tilde{C}_{ij}, R_{ME}^{ij})$, $\tilde{D}_{xx}^{ij}(r)$ and $\tilde{D}_{xx}^{ji}(r)$ can also be computed in a numerically exact fashion on a subset of points contained in $\Omega$.

To efficiently utilize this concept of concentric spherical subdomains in the exx module, we assemble two fixed-size proto-subdomains, $\Theta(C_0, R_{PE})$ and $\Theta(C_0, R_{ME})$,
Figure 3.4: Graphical depiction of the proto-subdomains used in the exx module. Dots are used to denote the MLWF centers, \( \tilde{C}_i \), which are approximated by the closest points, \( C_i \), on the real-space grid, \( \Omega \). The dashed blue and red circles bound the two concentric spherical proto-subdomains, \( \Theta(C_0, R_{PE}) \) and \( \Theta(C_0, R_{ME}) \), which are assembled around \( C_0 \) (purple star) with radii \( R_{PE} \) and \( R_{ME} \), respectively. Pair-exchange interactions involving \( \tilde{\phi}_i(r) \) include all overlapping MLWFs (yellow dots), whose centers are located within a distance, \( R_{pair} \), of \( \tilde{C}_i \approx C_i \) (black dot) that is large enough to account for all \( \tilde{\phi}_k(r) \) with \( \Omega_{ik} \neq \emptyset \). For the \( ij \) overlapping pair, the \( \Theta(C_0, R_{PE}) \) and \( \Theta(C_0, R_{ME}) \) proto-subdomains are translated across \( \Omega \) via a rigid grid offset, \( \tau_{ij} \), to form \( \Theta(C_{ij}, R_{PE}) \) and \( \Theta(C_{ij}, R_{ME}) \), which are centered at \( C_{ij} \approx \tilde{C}_{ij} \) (purple square). To evaluate the \( ij \) contribution to all EXX-related quantities, the corresponding PE, \( \nabla^2 \tilde{v}_{ij}(r) = -4\pi \tilde{\rho}_{ij}(r) \), is solved for \( \tilde{v}_{ij}(r) \) on the \( \Theta(C_{ij}, R_{PE}) \) subdomain, which encompasses \( \Omega_{ij} \), the support of \( \tilde{\rho}_{ij}(r) \) (shaded in dark green). Boundary conditions for the PE (as well as the far-field \( \tilde{v}_{ij}(r) \)) are computed via a ME of \( \tilde{\rho}_{ij}(r) \) on the \( \Theta(C_{ij}, R_{ME}) \) \( \setminus \Theta(C_{ij}, R_{PE}) \) shell surrounding \( \Theta(C_{ij}, R_{PE}) \).

centered around a predetermined origin, \( C_0 \), which is chosen to be one of the grid points in \( \Omega \). When dealing with all computations involving a given \( ij \) pair, these
proto-subdomains are simply translated to $\tilde{C}_{ij}$, which will be approximated (with no discernible error) by $C_{ij}$, the closest grid point in $\Omega$ (see Fig. 3.4 and Sec. 3.3.3). Since these fixed-size proto-subdomains will be used for all $ij$ pairs, their radii should be chosen such that $R_{PE} = \max_{ij} \{ R_{ij}^{PE} \}$ and $R_{ME} = \max_{ij} \{ R_{ij}^{ME} \}$. With judicious choices for $R_{PE}$ and $R_{ME}$ (see Secs. 4.1.1–4.1.2), these proto-subdomains allow for an exact evaluation of all EXX-related quantities, and have several algorithmic advantages that will be described below.

With $R_{PE}$ and $R_{ME}$ in hand, we now describe the construction of these proto-subdomains around $C_0$, an arbitrary center that is coincident with a grid point in $\Omega$. In this work, the grid point closest to the center of $\Omega$ was chosen as the reference $C_0$, since this allows us to avoid the use of both the minimum image convention and wrap-around (periodic) boundary conditions during grid point screening. Assembly of the proto-subdomains begins by looping over grid points, $r \in \Omega$, and determining whether or not a given grid point is contained within $\Theta(C_0, R_{PE})$ or $\Theta(C_0, R_{ME}) \setminus \Theta(C_0, R_{PE})$. For each grid point contained in either proto-subdomain, we increment the corresponding counter ($q'$ or $q''$) and store its relative (local) Cartesian coordinates (in $\mathbf{r}_{PE}$ or $\mathbf{r}_{ME}$) and (global) grid point indices (in $g_{PE}^0$ or $g_{ME}^0$), as depicted in Algorithm 2.
Algorithm 2 Proto-Subdomain Construction

$q' \leftarrow 0; q'' \leftarrow 0$

foreach $r \in \Omega$ do

    if $|r - C_0| \leq R_{PE}$ then
        $q' \leftarrow q' + 1$
        $\vec{r}_{PE}[q'] \leftarrow r - C_0$
        $g^0_{PE}[q'] \leftarrow \text{NINT} \left[ N_{\text{grid},\alpha} (h^{-1} r) \right], \quad \alpha = 1, 2, 3$
    
    else if $R_{PE} < |r - C_0| \leq R_{ME}$ then
        $q'' \leftarrow q'' + 1$
        $\vec{r}_{ME}[q''] \leftarrow r - C_0$
        $g^0_{ME}[q''] \leftarrow \text{NINT} \left[ N_{\text{grid},\alpha} (h^{-1} r) \right], \quad \alpha = 1, 2, 3$

    end if

end for

$N_{PE} \leftarrow q'$

$N_{ME} \leftarrow q' + q''$

Incrementing these counters throughout the loop over $r \in \Omega$ yields $N_{PE}$ and $N_{ME}$, the (fixed) number of points in $\Theta(C_0, R_{PE})$ and $\Theta(C_0, R_{ME})$. By storing the relative Cartesian coordinates, $r - C_0$, we now have a set of local coordinates that are invariant to rigid translations of $\Theta(C_0, R_{PE})$ and $\Theta(C_0, R_{ME})$, thereby avoiding the need to recompute these coordinates for every $ij$ pair. This also provides a convenient platform for precomputing a number of quantities (e.g., $r$ in spherical polar coordinates, the set of spherical harmonics, etc.) that are required during the ME of $\tilde{p}_{ij}(r)$ (cf. Eqs. (2.34)–(2.35)). For each point in the proto-subdomains, we also store its global grid point indices, which are given by three integer values, $(g^0_1, g^0_2, g^0_3)$, representing the position of a given grid point along the cell (lattice) vectors, $L_1$, $L_2$, and $L_3$. For a general (Bravais lattice based) grid, which has $N_{\text{grid},\alpha}$ equispaced grid points along each of the $L_\alpha$ lattice vectors (with grid spacing $\delta \xi_\alpha = |L_\alpha|/N_{\text{grid},\alpha}$), the
global grid index along $L_\alpha$ is given by $g^0_\alpha = N_{\text{grid}, \alpha}(h^{-1}r)_\alpha = N_{\text{grid}, \alpha}s_\alpha$. Since $r$ is always coincident with a grid point in $\Omega$, $\{g^0_\alpha\}$ is formally an array of integers; this is enforced in a floating-point environment using the nearest integer function, $\text{NINT}$.

This accumulated data is then concatenated to form two $3 \times N_{\text{ME}}$ arrays as follows: the local coordinates are stored in a double-precision array,

$$\bar{r}[q] = \begin{cases} \bar{r}_{\text{PE}}[q] & q = 1, \ldots, N_{\text{PE}} \\ \bar{r}_{\text{ME}}[q - N_{\text{PE}}] & q = N_{\text{PE}} + 1, \ldots, N_{\text{ME}} \end{cases}, \quad (3.1)$$

while the global grid indices are stored in an integer array,

$$g^0[q] = \begin{cases} g^0_{\text{PE}}[q] & q = 1, \ldots, N_{\text{PE}} \\ g^0_{\text{ME}}[q - N_{\text{PE}}] & q = N_{\text{PE}} + 1, \ldots, N_{\text{ME}} \end{cases}. \quad (3.2)$$

By storing all of the subdomain data in this scheme, only a single local index, $q$, is required for labeling the elements in these arrays, while still maintaining access to the $\Theta(C_0, R_{\text{PE}})$ and $\Theta(C_0, R_{\text{ME}})$ proto-subdomains (as well as the $\Theta(C_0, R_{\text{ME}}) \setminus \Theta(C_0, R_{\text{PE}})$ shell) through $N_{\text{PE}}$ and $N_{\text{ME}}$, the number of elements in each proto-subdomain. As such, this scheme provides us with a compact representation for the sparse quantities required in our EXX algorithm as well as a convenient mapping between data stored in the proto-subdomain representation and the real-space grid ($\Omega$) representation. This is crucial for loading and off-loading data to and from $\Omega$, communicating only the relevant sectors of $\Omega$ for sparse quantities such as $\tilde{\rho}_{ij}(\bar{r})$ and $\tilde{v}_{ij}(\bar{r})$, as well as determining grid-point neighbors during the finite-difference solution of the PE and the evaluation of the first derivatives of $\tilde{v}_{ij}(\bar{r})$ for $\sigma_{xx}$. 

56
Subdomains in Constant-Pressure CPMD

During an MLWF-based CPMD simulation of an insulating system, the band gap is not expected to have substantial variations. As such, the individual MLWF spreads will fluctuate, but the size/extent of the support associated with these exponentially decaying functions will remain essentially constant throughout the trajectory. For simulations with a fixed cell (e.g., NVE and NVT), this implies that the sizes of the $\Theta(C_{ij}, R_{PE})$ and $\Theta(C_{ij}, R_{ME})$ subdomains (i.e., the translated $\Theta(C_0, R_{PE})$ and $\Theta(C_0, R_{ME})$ proto-subdomains) will remain essentially fixed throughout the simulation and still encompass $\Omega_{ij}$ and $\Omega_i \cup \Omega_j$. As such, all proto-subdomain related quantities, which include the radii ($R_{PE}$ and $R_{ME}$), the number of local grid points ($N_{PE}$ and $N_{ME}$), the local (relative) Cartesian coordinates ($\{\vec{r}\}$), and the global grid indices ($\{g_0\}$), are kept fixed throughout the simulation and precomputed prior to the first MD step.

When a fluctuating cell is employed (e.g., during NpH and NpT simulations), the size and shape of $\Omega$ can vary significantly, while $\Omega_i$ is again expected to retain a similar size/extent (but potentially a different shape) throughout the trajectory. To understand how fluctuations in the cell size affect the choice of subdomains, consider the calculation of $E_{xx}$ for a simple cubic cell undergoing uniform contraction and expansion. In this case, one potential choice is to maintain a fixed $R_{PE}$ (and therefore a spherical subdomain shape) throughout the MD simulation, which has several mathematical and algorithmic consequences worth mention. For one, the use of a fixed $R_{PE}$ leads to different values for $N_{PE}$ as the cell fluctuates, with $N_{PE}$ being larger for cell contractions than expansions; this results in different grid point spacings ($\delta\xi$) and hence varying levels of accuracy in the discretization of the Laplacian operator. Such Pulay-like errors during the solution of the PE for cells with different sizes are largely mitigated in our MLWF-based EXX approach by the use of a sufficiently accurate finite-difference representation of the Laplacian (with an associated error of
\(O([\delta \xi]^6])\) and can be further reduced at linear cost with the use of a higher-order stencil (see Sec. 3.3.4). Algorithmically speaking, the use of a fixed \(R_{\text{PE}}\) also has the disadvantages of (i) requiring the computation of \(\{\vec{r}\}\) and \(\{g^0\}\) for each CPMD step, (ii) introducing an imbalance in the computational workload and associated memory requirements between CPMD steps, and (iii) complicating the use of extrapolation schemes for the \(\tilde{v}_{ij}(r)\) initial guess during the iterative solution of the PE.

To combat these issues, we have opted to employ subdomains with a fixed number of grid points throughout the simulation, as determined by the values of \(N_{\text{PE}}\) and \(N_{\text{ME}}\) corresponding to the initial proto-subdomains. More specifically, we retain the following proto-subdomain related quantities throughout the MD simulation: the number of grid points (\(N_{\text{PE}}\) and \(N_{\text{ME}}\)), the global grid indices (\(\{g^0\}\)), and the relative scaled coordinates (\(\{s\} = \{h_0^{-1}\vec{r}\}\), where \(h_0\) is the initial cell tensor). In other words, the subdomains employed in our approach do not have fixed radii (\(R_{\text{PE}}\) and \(R_{\text{ME}}\)) and are therefore no longer (necessarily) spherical; instead, these subdomains take on the shape of the fluctuating cell (e.g., sphere \(\rightarrow\) ellipsoid). This scheme directly addresses all of the algorithmic disadvantages that accompany the use of subdomains with fixed radii. For one, there is no need for the additional computational overhead associated with computing \(\{\vec{r}\}\) and \(\{g^0\}\) by screening \(\Omega\) at each CPMD step, as \(\{\vec{r}\}\) is straightforwardly obtained \textit{via} \(\{\vec{r}\} = \{h\vec{s}\}\) (where \(h\) is the current cell tensor) and \(\{g^0\}\) is simply stored in persistent memory. In addition, the complications associated with workload/memory imbalances as well as the use of extrapolation schemes between CPMD steps are completely eliminated with the use of a fixed number of grid points in each subdomain. Although this approach incurs the Pulay-like errors discussed above (since the grid point spacing is still a function of the cell size and shape), this error is effectively eliminated with a sufficiently accurate discretization of the Laplacian.

58
Another salient feature of this approach is the fact that our subdomains are no longer restricted to be spherical and can now take on the shape of the fluctuating cell; this is essential for encompassing the support of $\tilde{\phi}_i(\mathbf{r})$ when the shape of $\Omega_i$ also deforms with the cell (e.g., in atomic solids). To maintain a formally exact evaluation of all EXX-related quantities in the presence of severely anisotropic cell fluctuations, this approach should be further modified to ensure that the substantially deformed subdomains still encompass $\Omega_{ij}$ and $\Omega_i \cup \Omega_j$. This can be accomplished with the re-assembly (from scratch) of appropriately sized spherical proto-subdomains based on a predefined strain criteria or with a given stride (e.g., every 1000 CPMD steps) throughout the simulation, which still retains all of the algorithmic advantages mentioned above.

3.3.3 Step III: Communication of MLWFs

By virtue of the ORBITAL data distribution scheme, the unique MLWF-pair list, $\mathcal{L}$, not only determines the computational workload associated with each MPI process, but also encodes the communication protocol that will be followed throughout the remainder of the exx module (see Fig. 3.3). With a support that is significantly smaller than $\Omega$, the communication of any given MLWF on the entire real-space grid is clearly neither efficient nor necessary in our EXX algorithm. To reduce the communication overhead associated with each overlapping $ij$ pair, the exx module employs the proto-subdomains ($\Theta(C_0, R_{PE})$ and $\Theta(C_0, R_{ME})$) introduced in Sec. 3.3.2. As discussed above, these system-size-independent proto-subdomains provide a compact data representation for the storage and communication of sparse quantities such as $\tilde{\phi}_i$, $\tilde{\rho}_{ij}$, $\tilde{v}_{ij}$, and $\tilde{D}_{xx}^{ij}$ (or $\tilde{D}_{xx}^{ji}$). To utilize $\Theta(C_0, R_{PE})$ and $\Theta(C_0, R_{ME})$ in practice, these proto-subdomains must be translated across $\Omega$ to form the subdomains, $\Theta(C_{ij}, R_{PE})$ and $\Theta(C_{ij}, R_{ME})$, required for evaluating all quantities associated with a given $ij$ pair, as shown in Fig. 3.4.
Before describing the translation of these proto-subdomains, we now discuss the employed convention used herein for \( \tilde{C}_{ij} \), and remind the reader of the flexibility one has in defining this quantity for neutral charge distributions like \( \tilde{\rho}_{ij}(r) \) (see Sec. 2.2.3). In the \texttt{exx} module, \( \tilde{C}_{ij} \) is defined as the midpoint between the \( i \)-th and \( j \)-th MLWF centers, \textit{i.e.}, \( \tilde{C}_{ij} = (\tilde{C}_i + \tilde{C}_j)/2 \), which represents an excellent approximation to the aforementioned gauge used in molecular quantum mechanics and an algorithmically convenient choice. By utilizing the MLWF centers, this definition for \( \tilde{C}_{ij} \) accounts for the spatial distribution of each MLWF through its first moment, and becomes equivalent to the conventional definition, \( \tilde{C}_{ij} = \int \text{d}r \, r |\tilde{\rho}_{ij}(r)| / \int \text{d}r \, |\tilde{\rho}_{ij}(r)| \), for a number of different symmetric cases (\textit{e.g.}, when both \( \tilde{\phi}_i(r) \) and \( \tilde{\phi}_j(r) \) have the same spread and are spherically symmetric with respect to \( \tilde{C}_i \) and \( \tilde{C}_j \), when \( \tilde{\rho}_{ij}(r) \) is centrosymmetric with respect to the midpoint, etc.). Furthermore, this choice for \( \tilde{C}_{ij} \) recovers the correct center of charge for \( \tilde{\rho}_{ii}(r) \), \textit{i.e.}, \( \tilde{C}_{ij} \to \tilde{C}_{ii} = \tilde{C}_i \). Algorithmically speaking, this convention for \( \tilde{C}_{ij} \) is also quite useful, as it only requires knowledge of the MLWF centers, which is available throughout a CPMD simulation.

As mentioned above, the \texttt{exx} module employs one additional simplification when dealing with MLWF and MLWF-pair centers: these quantities are approximated by the closest grid points in \( \Omega \) and denoted throughout by either \( C_i \) or \( C_{ij} \). This algorithmic simplification leads to no appreciable error during evaluation of \( E_{\text{xx}}, \sigma_{\text{xx}}, \) and \( \{ \tilde{D}_{\text{xx}} \} \), and allows us to rigidly translate the proto-subdomains to the appropriate center, \( C_{ij} \), as needed. For a general (Bravais lattice based) grid, the component of the required grid translation vector, \( \tau_{ij} \), along a given lattice vector, \( L_\alpha \), is given by:

\[
\tau_{ij}^\alpha = \text{NINT} \left[ N_{\text{grid},\alpha} \left( h^{-1} (C_{ij} - C_0) \right)_\alpha \right],
\]

\[
= \text{NINT} \left[ |L_\alpha| \left( h^{-1} (C_{ij} - C_0) \right)_\alpha / \delta \xi_\alpha \right].
\]
Application of $\tau_{ij}$ to a given proto-subdomain leaves the radius and local Cartesian coordinates untouched, and simply offsets the global grid indices as follows:

$$g_{ij}^\alpha[q] = \text{MOD} \left[ g_0^\alpha[q] + \tau_{ij}^\alpha, N_{\text{grid},\alpha} \right] ,$$

thereby resulting in a subdomain centered at $C_{ij}$. The use of $\text{MOD}$ in Eq. (3.5) enforces the appropriate wrap-around boundary conditions; as such, this equation is specific to the grid convention used in QE, in which the grid points (along $L_\alpha$) are numbered from $0, 1, \ldots, N_{\text{grid},\alpha} - 1$. For codes that number these grid points from $1, 2, \ldots, N_{\text{grid},\alpha}$, Eq. (3.5) should be modified as follows: $g_{ij}^\alpha[q] = \text{MOD} \left[ g_0^\alpha[q] + \tau_{ij}^\alpha - 1, N_{\text{grid},\alpha} \right] + 1$.

With each MLWF stored according to the ORBITAL data distribution scheme, the MPI process ($\zeta = 1$) or processes ($\zeta > 1$) that are currently storing $\tilde{\phi}_i(r)$ on $\Omega$ are responsible for sending this MLWF to another MPI process (or processes) according to the computation and communication protocol outlined by $\mathcal{L}$. In order to do so, $\tilde{\phi}_i(r)$ is off-loaded onto appropriately translated subdomains, $\Theta(C_{ij}, R_{\text{ME}})$, corresponding to the overlapping $ij$ pairs that will be handled remotely (i.e., on other processes); all of the information required to do so is provided by local access to $\mathcal{L}$ and $\{C_{ij}\}$, as both of these arrays have been broadcast to all processes. For each of these overlapping $ij$ pairs, a sparse quantity like $\tilde{\phi}_i(r)$ is now stored in the more compact $\Theta(C_{ij}, R_{\text{ME}})$ representation via the use of three relatively small arrays: $\bar{r}$, $g^{ij}$, and $\tilde{\phi}_i(\bar{r}) \equiv \tilde{\phi}_i(\bar{r} + C_{ij})$, with associated sizes (types) of $3 \times N_{\text{ME}}$ (double-precision), $3 \times N_{\text{ME}}$ (integer), and $1 \times N_{\text{ME}}$ (double-precision), respectively. Here, we remind the reader that all of the data on the $\Theta(C_{ij}, R_{\text{PE}})$ subdomain and $\Theta(C_{ij}, R_{\text{ME}}) \setminus \Theta(C_{ij}, R_{\text{PE}})$ shell are contained within $\Theta(C_{ij}, R_{\text{ME}})$, and can easily be accessed using the local grid indexing scheme outlined in Eqs. (3.1)-(3.2). As mentioned above, the local Cartesian coordinates stored in the $\bar{r}$ array are invariant to translations of the proto-subdomains; as such, this information does not need to be recomputed for each translated sub-
domain and can be broadcast across all processes. Communication of the MLWFs on these compact subdomains then proceeds according to $\mathcal{L}$ among the pool of available MPI processes. Once $\tilde{\phi}_i(\vec{r})$ is received by a given process, $\tilde{\rho}_{ij}(\vec{r})$ is assembled on the $\Theta(C_{ij}, R_{PE})$ subdomain and the exx module begins the process of solving the corresponding PE.

### 3.3.4 Step IV: Solution of Poisson’s Equation

Based on $\mathcal{L}$, each MPI process, $P_i$, now holds an assigned MLWF-product density $\tilde{\rho}_{ij}(\vec{r})$ as well as the relevant quantities that map the $\Theta(C_{ij}, R_{PE})$ and $\Theta(C_{ij}, R_{ME})$ subdomains onto $\Omega$ (i.e., $N_{PE}$, $N_{ME}$, $\{\vec{r}\}$, and $\{g^{ij}\}$). As such, $P_i$ has all of the required information to compute $\tilde{v}_{ij}(\vec{r})$ on the $\Theta(C_{ij}, R_{PE})$ and $\Theta(C_{ij}, R_{ME}) \backslash \Theta(C_{ij}, R_{PE})$ subdomains. On $\Theta(C_{ij}, R_{PE})$, $\tilde{v}_{ij}(\vec{r})$ is obtained via the solution of the PE in Eq. (2.33). On $\Theta(C_{ij}, R_{ME}) \backslash \Theta(C_{ij}, R_{PE})$, $\tilde{v}_{ij}(\vec{r})$ is obtained via the ME in Eqs. (2.34)–(2.35), which provides the appropriate boundary conditions for the PE as well as the far-field potential.

While the ME of $\tilde{\rho}_{ij}(\vec{r})$ (about $C_{ij}$) can be straightforwardly computed using the local coordinates, $\{\vec{r}\}$, the PE requires a discrete representation of the Laplacian operator for computing numerical second derivatives on these subdomains. Since the subdomains employed in the exx module are coincident with the underlying real-space grid, it is most computationally efficient to employ a discrete representation for the Laplacian that is aligned with the corresponding lattice vectors, $L_1$, $L_2$, and $L_3$. Throughout this discussion we will consider the general case in which these lattice vectors are non-orthogonal; hence, the subdomains are not aligned with the unit Cartesian directions, $\hat{x}$, $\hat{y}$, and $\hat{z}$, as one would encounter with orthorhombic (e.g., simple cubic) cells.

To proceed, we adapt the unit lattice vectors as a basis for this tilted (non-Cartesian) space, i.e., $\tilde{L}_\alpha = L_\alpha / |L_\alpha|$, such that a given position vector in $\mathbb{R}^3$ can
be written using either Cartesian ($\mathbf{r} = \{x, y, z\}$) or tilted ($\mathbf{\xi} = \{\xi_1, \xi_2, \xi_3\}$) coordinates. To facilitate the direct solution of the PE on these subdomains (i.e., without the need for interpolation of $\tilde{\rho}_{ij}(\mathbf{r})$ and $\tilde{v}_{ij}(\mathbf{r})$ to and from an auxiliary Cartesian grid), we need to define a transformation which connects the Laplacian operator in these two representations, i.e., $\nabla^2_{\mathbf{r}} \leftrightarrow \nabla^2_{\mathbf{\xi}}$. This transformation can be derived from the corresponding Jacobian matrix, $\mathbf{J} = (\partial \mathbf{\xi} / \partial \mathbf{r})$, which connects the Cartesian and tilted bases. Since $\xi_\alpha = |\mathbf{L}_\alpha| s_\alpha = |\mathbf{L}_\alpha| \sum_a (h^{-1})_{aa} r_a$, an explicit form for the Jacobian is obtained as follows:

$$
J_{a\alpha} = \left( \frac{\partial \xi_\alpha}{\partial r_a} \right) = \frac{\partial}{\partial r_a} \left[ |\mathbf{L}_\alpha| \sum_b (h^{-1})_{ab} r_b \right] = |\mathbf{L}_\alpha| \sum_b (h^{-1})_{ab} \delta_{ba} = |\mathbf{L}_\alpha| (h^{-1})_{aa} .
$$

(3.6)

This expression can then be used to write the Laplacian in the tilted representation, i.e.,

$$
\nabla^2_{\mathbf{r}} = \nabla_{\mathbf{r}} \cdot \nabla_{\mathbf{r}} = \sum_a (\nabla_{\mathbf{r}})_a (\nabla_{\mathbf{r}})_a = \sum_a \left[ \sum_\alpha J_{a\alpha} (\nabla_{\mathbf{\xi}})_\alpha \right] \left[ \sum_\beta J_{a\beta} (\nabla_{\mathbf{\xi}})_\beta \right] = \sum_{a\beta} F_{a\beta} (\nabla_{\mathbf{\xi}})_\alpha (\nabla_{\mathbf{\xi}})_\beta .
$$

(3.7)

in which $F_{a\beta} \equiv \sum_a J_{a\alpha} J_{a\beta}$. Using the Clairaut-Schwarz theorem, the Laplacian in Eq. (3.7) can be further split into a sum over pure and mixed second partial derivatives as follows:

$$
\nabla^2_{\mathbf{r}} = \sum_a \left[ F_{aa} \frac{\partial^2}{\partial s_a^2} + 2 \sum_{\beta > a} F_{a\beta} \frac{\partial^2}{\partial s_a \partial \xi_{\alpha}} \right] .
$$

(3.8)
Here, we note that the number of non-vanishing $F_{\alpha\beta}$ determines the number of mixed partial derivatives that need to be considered when evaluating Eq. (3.8). As such, the computational cost associated with evaluating the Laplacian operator is correlated with the number of non-vanishing $F_{\alpha\beta}$. For example, during simulations involving orthorhombic cells (assuming that the lattice vectors are aligned with the Cartesian axes), $F_{\alpha\beta} = 0$ for $\alpha \neq \beta$ and the corresponding Laplacian comprises only pure derivatives. For more detailed performance considerations regarding general Bravais lattice based cells, see Sec. 4.3.

While the pure derivatives in Eq. (3.8) can be efficiently represented by standard central-difference formulae along the lattice vectors (vide infra), there exists considerable flexibility when evaluating the mixed derivatives in this expression. Here, we adopt the treatment outlined by Kronik and coworkers, and define the following unit auxiliary vector:

$$\hat{L}_{\alpha\beta} = \frac{\hat{L}_{\alpha} + \kappa_{\alpha\beta} \hat{L}_{\beta}}{\sqrt{\hat{L}_{\alpha} + \kappa_{\alpha\beta} \hat{L}_{\beta}}} \equiv \frac{\hat{L}_{\alpha} + \kappa_{\alpha\beta} \hat{L}_{\beta}}{d_{\alpha\beta}},$$

(3.9)

in which $\kappa_{\alpha\beta}$ is a non-zero mixing coefficient. With this expression in hand, the first and second partial derivatives with respect to $\xi_{\alpha\beta}$ (a coordinate along $\hat{L}_{\alpha\beta}$) take on the following form:

$$\frac{\partial}{\partial \xi_{\alpha\beta}} = \frac{1}{d_{\alpha\beta}} \left( \frac{\partial}{\partial \xi_{\alpha}} + \kappa_{\alpha\beta} \frac{\partial}{\partial \xi_{\beta}} \right),$$

(3.10)

and

$$\frac{\partial^2}{\partial^2 \xi_{\alpha\beta}} = \frac{1}{d_{\alpha\beta}^2} \left( \frac{\partial^2}{\partial^2 \xi_{\alpha}} + \kappa_{\alpha\beta}^2 \frac{\partial^2}{\partial^2 \xi_{\beta}} + 2\kappa_{\alpha\beta} \frac{\partial^2}{\partial \xi_{\alpha} \partial \xi_{\beta}} \right),$$

(3.11)
the latter of which can be rearranged to express the mixed partial derivative,

\[
\frac{\partial^2}{\partial \xi_\alpha \partial \xi_\beta} = \frac{1}{2 \kappa_{\alpha\beta}} \left( -\frac{\partial^2}{\partial \xi_\alpha^2} - \kappa_{\alpha\beta}^2 \frac{\partial^2}{\partial \xi_\beta^2} + \delta_{\alpha\beta}^2 \frac{\partial^2}{\partial \xi_{\alpha\beta}^2} \right),
\]

(3.12)

in terms of pure derivatives along the \(\hat{\mathbf{L}}_\alpha\) and \(\hat{\mathbf{L}}_\beta\) unit lattice vectors, and the \(\hat{\mathbf{L}}_{\alpha\beta}\) unit auxiliary vector. This reformulation of the mixed derivative in Eq. (3.12) is crucial for an efficient evaluation of the Laplacian in Eq. (3.8) since the pure derivatives along \(\hat{\mathbf{L}}_\alpha\) and \(\hat{\mathbf{L}}_\beta\) can now be merged in these two expressions.

By construction, the finite-difference stencils required to evaluate \(\frac{\partial^2}{\partial \xi_\alpha^2}\) and \(\frac{\partial^2}{\partial \xi_\beta^2}\) are coincident with the \(\mathbf{L}_\alpha\) and \(\mathbf{L}_\beta\) lattice vectors, with corresponding grid spacings (\(\delta \xi_\alpha\) and \(\delta \xi_\beta\)) that are fixed by the planewave (kinetic energy) cutoff and the FFT algorithm. For the pure derivative along \(\hat{\mathbf{L}}_{\alpha\beta}\), a judicious choice for \(\kappa_{\alpha\beta}\) ensures that: (i) \(\frac{\partial^2}{\partial \xi_{\alpha\beta}^2}\) is coincident with a real-space grid direction (distinct from \(\mathbf{L}_1, \mathbf{L}_2,\) and \(\mathbf{L}_3\)) and (ii) \(\hat{\mathbf{L}}_{\alpha\beta}\) has the minimum possible grid spacing, \(\delta \xi_{\alpha\beta}\) (which improves the accuracy of the finite-differencing scheme). For cells with \(\delta \xi_\alpha \approx \delta \xi_\beta\) (which is the case for standard planewave based approaches), an explicit choice for \(\kappa_{\alpha\beta}\) which accomplishes both of these goals is given by:

\[
\kappa_{\alpha\beta} = -\frac{\delta \xi_\beta}{\delta \xi_\alpha} \text{sgn} \left[ \hat{\mathbf{L}}_\alpha \cdot \hat{\mathbf{L}}_\beta \right] = -\frac{\delta \xi_\beta}{\delta \xi_\alpha} \text{sgn} \left[ \cos \Phi_{\alpha\beta} \right],
\]

(3.13)

wherein \(\Phi_{\alpha\beta}\) is the angle between \(\hat{\mathbf{L}}_\alpha\) and \(\hat{\mathbf{L}}_\beta\). Since \(\delta \xi_\alpha, \delta \xi_\beta,\) and \(\Phi_{\alpha\beta}\) are all quantities determined by the lattice vectors, this expression for \(\kappa_{\alpha\beta}\) uses the sign of \(\cos \Phi_{\alpha\beta}\) to determine the direction of \(\hat{\mathbf{L}}_{\alpha\beta}\) with the minimum grid spacing, \(\delta \xi_{\alpha\beta}\). When \(\Phi_{\alpha\beta} > 90^\circ\) \((i.e., \Phi_{\alpha\beta}\) is an obtuse angle), then \(\text{sgn}[\cos \Phi_{\alpha\beta}] = -1\) and \(\kappa_{\alpha\beta} = \delta \xi_\beta/\delta \xi_\alpha > 0\); as such, the mixing coefficient in front of \(\hat{\mathbf{L}}_\beta\) is positive in Eq. (3.9) and \(\hat{\mathbf{L}}_{\alpha\beta}\) becomes the grid-resolved bisector of the obtuse \(\Phi_{\alpha\beta}\) angle (see Fig. 3.5). When \(\Phi_{\alpha\beta} < 90^\circ\)

\[
^*\text{This choice also extends to cases where } \delta \xi_\alpha/\delta \xi_\beta < 1.5/|\cos \Phi_{\alpha\beta}| \text{ and } \delta \xi_\beta/\delta \xi_\alpha < 1.5/|\cos \Phi_{\alpha\beta}|. \text{ As such, this expression for } \kappa_{\alpha\beta} \text{ should be valid in any standard planewave based DFT implementation.}
\]

65
(i.e., $\Phi_{\alpha\beta}$ is an acute angle), $\kappa_{\alpha\beta} = -\delta\xi_{\beta}/\delta\xi_{\alpha} < 0$ and the mixing coefficient in front of $\hat{L}_\beta$ is negative; in this case, $\hat{L}_{\alpha\beta}$ becomes the grid-resolved bisector of the obtuse supplementary angle of $\Phi_{\alpha\beta}$ (see Fig. [3.5]). With $\hat{L}_{\alpha\beta}$ as the grid direction bisecting the appropriate obtuse angle, the auxiliary grid spacing, $\delta\xi_{\alpha\beta} = |\hat{L}_\alpha + \kappa_{\alpha\beta}\hat{L}_\beta| = \delta\xi_\alpha d_{\alpha\beta}$, is minimized.

After some algebra, one can plug Eq. (3.12) into Eq. (3.7) (or Eq. (3.8)) and exploit the fact that $F_{\beta\alpha} = F_{\alpha\beta}$, $\kappa_{\beta\alpha} = 1/\kappa_{\alpha\beta}$, and $d_{\beta\alpha} = d_{\alpha\beta}/|\kappa_{\alpha\beta}|$ to arrive at the following working expression for the Laplacian operating on a general Bravais lattice based grid:

$$\nabla^2 r = \sum_\alpha \left[ \left(-\sum_\beta \frac{F_{\alpha\beta}}{\kappa_{\alpha\beta}} \frac{\partial^2}{\partial \xi^2_\alpha} + \sum_{\beta > \alpha} F_{\alpha\beta} d_{\alpha\beta}^2 \frac{\partial^2}{\partial \xi^2_{\alpha\beta}} \right) \right].$$

(3.14)

In this form, the Laplacian operator depends exclusively on pure second partial derivatives, which can be straightforwardly evaluated with the finite-difference techniques described below.

To discretize the Laplacian operator in Eq. (3.14), we employ the standard central-difference approach, which computes the pure second derivative of a function, $f(\xi)$ (with respect to an arbitrary direction, $\xi_\gamma$), at a given grid point, $\xi_0$, as follows:

$$\left. \frac{\partial^2 f(\xi)}{\partial \xi^2_\gamma} \right|_{\xi = \xi_0} = \sum_{q=-n}^{n} w_q f(\xi_0 + q \delta \xi_\gamma \hat{L}_\gamma)/\delta \xi^2_\gamma. \quad (3.15)$$

In this expression, the associated discretization error depends on the number, $n$, of (equispaced) neighboring grid points on each side of $\xi_0$, and $w_q$ is the central-difference coefficient for the $q$-th neighboring grid point; only $w_{|q|}$ is required due to the central symmetry ($w_q = w_{-q}$) of the finite-difference stencil. Discretization of this derivative results in a $2n + 1$-point stencil along each grid direction, $L_\gamma$. For an orthogonal cell, the Laplacian in Eq. (3.14) requires three pure second derivatives along the lattice vectors, $L_1$, $L_2$, and $L_3$, which corresponds to a finite-difference stencil covering
Figure 3.5: Discretization of the Laplacian (shown here in 2D for simplicity) on an orthogonal cell (left) and two tilted (non-orthogonal) cells with obtuse (center) and acute (right) values of $\Phi_{12}$, the angle between the $L_1$ and $L_2$ lattice vectors. Each discretized Laplacian is centered at a given grid point, $\xi_0$ (highlighted in yellow), and is represented by a stencil which covers the neighboring $\pm n$ (shown here for $n = 3$) grid points in each required derivative direction. For non-orthogonal cells, pure second derivatives with respect to $\xi_{12}$ are computed along the unit auxiliary vector ($\hat{L}_{12}$, see Eqs. (3.9) and (3.13)) corresponding to the grid-resolved bisector of the obtuse $\Phi_{12}$ angle (or its supplementary angle when $\Phi_{12}$ is acute).

$3 \times 2n + 1 = 6n + 1$ grid points (see Fig. 3.5). For a non-orthogonal cell with $N_{\text{off}}$ non-zero off-diagonal elements in the upper (or lower) triangle of $F$, the Laplacian will require an additional second derivative along each of the $N_{\text{off}}$ auxiliary vectors, $L_{\alpha\beta}$, and the corresponding finite-difference stencil will cover $(3 + N_{\text{off}}) \times 2n + 1 = 6n + 2nN_{\text{off}} + 1$ points (see Fig. 3.5). As such, the treatment of non-orthogonal cells will incur a larger computational cost that scales with the number of points in the finite-difference stencil. We note in passing that the choice of $n = 3$ (with corresponding central-difference coefficients given by $w_0 = -49/18$, $w_1 = +3/2$, $w_2 = -3/20$, and $w_3 = +1/90$) yields a second derivative with an associated discretization error of
\( O (\delta \zeta^6) \); this choice is the default option in the \texttt{exx} module as it yields a well-converged value for \( E_{\text{xx}} \).

With this discrete representation of the Laplacian, we can express the PE in Eq. (2.33), \( \nabla^2 \tilde{v}_{ij}(\vec{r}) = -4\pi \tilde{\rho}_{ij}(\vec{r}) \), as the following set of sparse linear equations on the \( \Theta(C_{ij}, R_{PE}) \) subdomain:

\[
\nabla_{PE}^2 \tilde{v}_{ij} = -4\pi \left( \tilde{\rho}_{ij} - \tilde{\rho}^b_{ij} \right). \tag{3.16}
\]

In this expression, \( \nabla_{PE}^2 \) is a sparse \( N_{PE} \times N_{PE} \) matrix containing the discretized Laplacian (whose stencil coverage has been restricted to \( \Theta(C_{ij}, R_{PE}) \)), \( \tilde{v}_{ij} \) is a \( N_{PE} \times 1 \) vector representing the (currently unknown) MLWF-product potential, and \( \tilde{\rho}_{ij} \) is a \( N_{PE} \times 1 \) vector containing the MLWF-product density. The final term on the right-hand side, \( \tilde{\rho}^b_{ij} \equiv -1/(4\pi)(\nabla^2 - \nabla_{PE}^2)\tilde{v}_{ij} \), is the so-called boundary charge, which accounts for the parts of the Laplacian stencil that extend outside of \( \Theta(C_{ij}, R_{PE}) \) (and into \( \Theta(C_{ij}, R_{ME}) \setminus \Theta(C_{ij}, R_{PE}) \) shell) and have been truncated in the \( \nabla_{PE}^2 \) representation of this operator. In doing so, \( \tilde{\rho}^b_{ij} \) accounts for the Dirichlet boundary conditions provided by the ME (in Eq. (2.34)) form for the potential on the \( \Theta(C_{ij}, R_{ME}) \setminus \Theta(C_{ij}, R_{PE}) \) shell surrounding \( \Theta(C_{ij}, R_{PE}) \), and therefore allows for an exact solution of the PE using a Laplacian whose stencil coverage has been restricted to the \( \Theta(C_{ij}, R_{PE}) \) subdomain. This restricts the PE to the support of \( \tilde{\rho}_{ij} \) and substantially reduces the dimensionality and associated computational cost of solving the PE for each overlapping MLWF pair.

The system of sparse linear equations in Eq. (3.16) is then solved (for \( \tilde{v}_{ij} \)) using an iterative conjugate-gradient (CG) approach. Since the solution of the PE is notoriously difficult to parallelize efficiently over \texttt{MPI} tasks, the CG-based PE solver in the \texttt{exx} module is largely parallelized over \texttt{OpenMP} threads to allow for an efficient real-space evaluation of \( \tilde{v}_{ij} \). The efficient solution of the PE for each overlapping MLWF
pair is the cornerstone of our MLWF-based EXX algorithm, and the performance of the CG-based PE solver will be discussed in Sec. 4.2.2. During CPMD simulations, the number of CG iterations required to solve a given PE can be substantially reduced by using a polynomial extrapolation of the potential from the previous CPMD steps as the initial guess. More detailed considerations of this extrapolation scheme as well as extensions to BOMD will be discussed in future work.

3.3.5 Step V: Computation of Energy and Forces

After a process, \( P_i \), arrives at the solution to the PE for one of its assigned pairs \( i.e. \), for a given \( j \in L[i] \), this process now holds the corresponding MLWF-product potential \( \tilde{v}_{ij}(r) \) on the entire \( \Theta(C_{ij}, R_{ME}) \) subdomain. With \( \tilde{v}_{ij}(r) \) on the \( \Theta(C_{ij}, R_{PE}) \) subdomain \( (via \) the solution of the PE) and \( \tilde{v}_{ij}(\bar{r}) \) on the \( \Theta(C_{ij}, R_{ME}) \setminus \Theta(C_{ij}, R_{PE}) \) shell \( (via \) the ME of \( \tilde{\rho}_{ij}(r) \)), \( P_i \) is now ready to evaluate the \( ij \) contribution to the EXX energy \( (E_{xx}) \), cell forces \( (\sigma_{xx}) \), and wavefunction forces \( (\tilde{D}^{ij}_{xx}(\bar{r}) \) and \( \tilde{D}^{ji}_{xx}(\bar{r}) \)).

The evaluation of \( E_{xx} \) and \( \sigma_{xx} \) is quite straightforward \( via \) Eqs. (2.28) and (2.32), respectively. Here, we remind the reader that a formally exact evaluation of these quantities only requires integration on the \( \Theta(C_{ij}, R_{PE}) \) subdomain; this integration over the support of \( \tilde{\rho}_{ij}(r) \) is also parallelized over OpenMP threads and is therefore quite computationally efficient. Partial summations over the assigned \( ij \) pairs on each process are accumulated to form \( E_{xx} \) and \( \sigma_{xx} \) with minimal associated communication \( (i.e., \) one double-precision number for \( E_{xx} \) and \( 3 \times 3 \) double-precision numbers for \( \sigma_{xx} \) per MPI process). To evaluate the Cartesian gradient, \( \partial \tilde{v}_{ij}(\bar{r})/\partial r_a \), needed during the evaluation of \( \sigma_{xx} \), first derivatives of \( \tilde{v}_{ij}(\bar{r}) \) are computed along each of the lattice vectors and transformed using the Jacobian in Eq. (3.6) \( via \) \( \nabla_r = J \nabla_\xi \). Each of these first derivatives is obtained using an unrestricted evaluation of the corresponding central-difference stencils; the default option in the exx module is \( n = 3 \) \( (with \) an associated discretization error of \( O(\delta \xi^6) \)), as this choice also yields well-converged
values for $\sigma_{xx}$. For these first derivatives, the stencil is anti-symmetric ($w_q = w_{-q}$) and the corresponding central-difference coefficients\cite{136} are given by $w_0 = 0$, $w_1 = +3/4$, $w_2 = -3/20$, and $w_3 = +1/60$.

With $\tilde{v}_{ij}(\mathbf{r})$ in hand, $P_i$ is also in position to compute both $\tilde{D}^{ij}_{xx}(\mathbf{r}) = \tilde{v}_{ij}(\mathbf{r})\tilde{\phi}_j(\mathbf{r})$ and $\tilde{D}^{ji}_{xx}(\mathbf{r}) = \tilde{v}_{ij}(\mathbf{r})\tilde{\phi}_i(\mathbf{r})$ on the entire $\Theta(C_{ij}, R_{ME})$ subdomain. For each $\tilde{D}^{ji}_{xx}(\mathbf{r})$ computed on $P_i$, this quantity is shipped back to $P_j$ (assuming $\zeta = 1$ here for simplicity), which requires communication of $N_{ME}$ double-precision numbers for each $\tilde{D}^{ji}_{xx}(\mathbf{r})$; this array is equivalent in size to $\tilde{\phi}_j(\mathbf{r})$ and represents the necessary second communication event described in Sec. 3.3.2. Since $P_i$ has access to $\tilde{v}_{ij}(\mathbf{r}) \forall j \in \mathcal{L}[i]$, this process accumulates $\tilde{D}^i_{xx}(\mathbf{r}) = \sum_j \tilde{D}^{ij}_{xx}(\mathbf{r})$ into a local temporary array that is the size of the global real-space grid; as $P_i$ receives a given $\tilde{D}^{ij}_{xx}(\mathbf{r})$ array from $P_j$, this quantity is also accumulated into this local temporary array. When all $\tilde{D}^{ij}_{xx}(\mathbf{r})$ contributions are accounted for, this temporary array on $P_i$ holds the final $\tilde{D}^i_{xx}(\mathbf{r})$ according to the ORBITAL data distribution scheme.

### 3.3.6 Step VI: Redistribution of Wavefunction Forces

At this stage, all of the EXX-related quantities have been evaluated; $E_{xx}$ and $\sigma_{xx}$ have been accumulated and broadcast to all processes, while $\{\tilde{D}^i_{xx}(\mathbf{r})\}$ is stored in the ORBITAL data distribution scheme. As such, the remaining task for the exx module is the transformation of $\{\tilde{D}^i_{xx}(\mathbf{r})\}$ to the GRID data distribution scheme for compliance with the core functions in QE (see Sec. 3.2). This redistribution is essentially the reverse operation of the GRID $\rightarrow$ ORBITAL redistribution of the MLWFs described in Fig. 3.2 and Sec. 3.3.1.

At this stage, the QE executable exits the exx module with $E_{xx}$, $\sigma_{xx}$, and $\{\tilde{D}^i_{xx}(\mathbf{r})\}$ (in the GRID data distribution scheme) as output. These EXX-related quantities are then added to their semi-local exchange analogs with the appropriate EXX fraction, $a_x$, given in Eq. (1.12). With the EXX contribution to the cell and wavefunction
forces, the constant-pressure CPMD equations of motion are now propagated forward via Eqs. (2.6)–(2.8).

3.4 Conclusions

In this chapter, I have demonstrated a detailed discussion on a practical implementation of a formally exact linear-scaling EXX algorithm, named \texttt{exx} within the \texttt{CP} module of the \texttt{QE} package. Our discussion started with the necessary prerequisites of a CPMD simulation that is implemented within the \texttt{CP} module of \texttt{QE} and the minor computational overhead to generate the MLWFs using second-order damped dynamics for adaptive refinement while propagating the CPMD trajectory. Then, we introduced a custom \texttt{ORBITAL} data distribution scheme and its transformation from the default \texttt{GRID} scheme used in \texttt{CP} (and \texttt{QE} in general). The need for this \texttt{ORBITAL} scheme is to provide a convenient starting point for massively parallelize the computational workload of overlapping MLWF pairs to thousands of \texttt{MPI} processes (each with nested \texttt{OpenMP} threads). Based on the \texttt{ORBITAL} data distribution scheme, a unique pair list \( \mathcal{L} \) is constructed to statically distribute the set of PEs from overlapping MLWF pairs across the pool of available \texttt{MPI} processes. For simplicity, we employ two types of spherical subdomains to cover MLWF-product densities and potentials, respectively. When distributing orbitals (and later on the wavefunction forces) following \( \mathcal{L} \) the MLWFs are compressed by off-loading from the global real-space grid to the MLWF-product potential subdomain to reduce the communication overhead. The subdomains can be conveniently and efficiently mapped using the Wannier centers of the targeted pair from two spherical proto-subdomains precomputed before the AIMD simulation. After each \texttt{MPI} process solves its assigned PEs (on the associated MLWF-product density subdomains), it computes the corresponding EXX contributions to energy \( (E_{\text{xx}}) \), cell derivatives \( (\sigma_{\text{xx}}) \), and wavefunction forces \( \{ \tilde{D}_{\text{xx}}(r) \} \) (whose
far-field contributions comes from ME of the MLWF-product density). While $E_{xx}$ and $\sigma_{xx}$ are collected, summed, and broadcasted to all processes, $\{\tilde{D}_i^{xx}(r)\}$—after appropriate interprocess collection—is redistributed to compatible GRID data distribution to merge with the GGA wavefunction forces in QE. As such, the massive parallel exx module provides a general purpose tool for large-scale condensed-phase hybrid DFT based AIMD simulations with arbitrary Bravais lattice based unit cells. In Chapter 4 we will discuss the performance optimization of the exx module while simultaneously keeping a formally exact evaluation of the EXX related quantities.
Chapter 4

Accuracy and Performance

During the implementation (Chapter 3) of the exx module, we have introduced three parameters (Sec. 3.3.2): $R_{\text{pair}}$, $R_{\text{PE}}$, and $R_{\text{ME}}$. $R_{\text{pair}}$ defines the upper bound of the MLWF center distances for an orbital pair to be considered overlapping. For each of these overlapping $ij$ pairs, $R_{\text{PE}}$ and $R_{\text{ME}}$ are radii of spherical subdomains both centered at $C_{ij}$ to cover the product density $\tilde{\rho}_{ij}(r)$ and individual orbitals ($\tilde{\phi}_i(r)$ and $\tilde{\phi}_j(r)$), respectively, for each $ij$ pair. An accurate EXX calculation requires converging these parameters. However, the cost of solving the PE—the computational bottleneck—grows with both $N_{\text{pair}}$ and $N_{\text{PE}}$, which scale asymptotically cubically with $R_{\text{pair}}$ (for homogeneous systems) and $R_{\text{PE}}$, respectively. Moreover, the communication overhead for sending the MLWFs on $\Theta(C_{ij}, R_{\text{ME}})$ scales cubically with $R_{\text{ME}}$ due to the associated asymptotically cubic growth in $N_{\text{ME}}$. In order to perform efficiently large-scale hybrid-DFT based applications without appreciable loss of accuracy, judicious choices of these three parameters are required to balance the accuracy and performance, which is the focus of this section. In Sec. 4.1, we introduce a systematic selection of these parameters based on user-defined error thresholds on $E_{\text{xx}}$ and $\{\tilde{D}_{xx}^i(r)\}$. Then, in Sec. 4.2, we discuss the parallel scaling performance for both the intranode OpenMP and the internode MPI parallelization paradigms. Finally, in
Sec. 4.3, we discuss additional performance considerations arisen from the underlying Bravais lattice.

4.1 Parameters for Performance Optimization

To demonstrate the systematic determination of the targeted parameters, we choose a snapshot from an (H\textsubscript{2}O)\textsubscript{64} liquid in a cubic cell with \( L = 23.52 \) bohr as an example system. We begin with a reference PBE0 calculation using largest possible values of these parameters such that the proto-subdomains remain encompassed within the cell (\textit{i.e.}, \( R_{\text{pair}} = R_{\text{ME}} = L/2 = 11.76 \) bohr and \( R_{\text{PE}} = L/2 - n \max_{\alpha} \{\delta \xi_\alpha\} = 11.11 \) bohr), and obtain the resulting energy \( E_{xx}^{\text{ref}} \). The subtraction of \( n \max_{\alpha} \{\delta \xi_\alpha\} \) from \( L/2 \) in \( R_{\text{PE}} \) is to provide sufficient spatial region for the boundary condition needed by the solution of the PE.

In this reference calculation, we observed that the majority of the EXX contributions (\textit{e.g.}, \( \sim 85\% \) of \( E_{xx}^{\text{ref}} \)) come from the self-exchange terms, which only requires a minor portion (\( \sim 13\% \)) of the gross computational cost. The minor contribution to \( E_{xx} \) from the non-self pairs is due to the smaller orbital overlaps compared to the self ones. This smaller overlap also results in smaller support of the corresponding MLWF product density. As such, sufficiently converged \( R_{\text{PE}} \) for the self-exchange calculation is, in general, overestimated and hence inefficient for the evaluation of the pair-exchange interactions.

In addition to the choice of \( R_{\text{PE}} \), smaller \( R_{\text{ME}} \) can also be used for pair-exchange MLWF-product potential (\( \tilde{v}_{ij}, i \neq j \)) than the self-exchange ones (\( \tilde{v}_{ii} \)) due to the absence of monopole term in \( \tilde{\rho}_{ij}(\vec{r}) \) for the pair-exchange case (see Sec. 2.2.3), which leads to quicker decay in \( \tilde{v}_{ij}(\vec{r}) \) and hence smaller support for \( \tilde{D}_{xx}^{ij}(\vec{r}) = \tilde{v}_{ij}(\vec{r}) \tilde{\phi}_j(\vec{r}) \).

Although reducing \( R_{\text{ME}} \) does not speed up the computational bottleneck (\textit{i.e.}, the solution of the PEs), it lowers the MLWF communication overhead and the storage
of $\phi_j(\vec{r})$ and $\tilde{D}_{ji}(\vec{r})$ on $\Theta(C_{ij}, R_{ME})$; both savings are key to enabling large-scale EXX calculations. To exploit these savings for pair-exchange calculations, we made separate choices for $R_{PE}$ and $R_{ME}$ for self-exchange ($R_{PE}^s$ and $R_{ME}^s$) and pair-exchange ($R_{PE}^p$ and $R_{ME}^p$) cases, respectively.

Here, we choose these parameters based on the convergence of the EXX related quantities without analyzing individual support ($\Omega_{ij}$, $\Omega_i$, and $\Omega_j$) of each $ij$ pair. We determine these required parameters ($R_{pair}$, $R_{PE}^s$, $R_{PE}^p$, $R_{ME}^s$, and $R_{ME}^p$) by testing the convergence of $E_{xx}$ (see Sec. 4.1.1) and $\{\tilde{D}_{ij}(\vec{r})\}$ (see Sec. 4.1.2).

### 4.1.1 Convergence of Energy and Cell Forces

We assess the convergence of $E_{xx}$ as a function of $R_{pair}$, $R_{PE}^s$, and $R_{PE}^p$ by scanning one radius at a time, keeping the other controlled parameters at their largest possible values using in the reference calculation for $E_{xx}^{ref}$. In this test, $E_{xx}$ are computed non-self-consistently using the MLWFs corresponding to the reference calculation since the effect of the self-consistency is negligible vide infra. We discuss the convergence of $E_{xx}$ based on its relative error to $E_{xx}^{ref}$ (as plotted in Fig. 4.1). In this case, we choose $R_{pair} = 8.0$ bohr, $R_{PE}^s = 6.0$ bohr, and $R_{PE}^p = 5.0$ bohr as the parameter set for water systems; as such, the overall relative error is within $\sim 0.02\%$, which is typically sufficient for obtaining accurate ground state energies and ionic forces in condensed-phase systems. Since the chosen parameter set is in close proximity to the reference values, the overall relative error is a sum of error contributions from each parameters. To achieve higher level of convergence in $E_{xx}$, a rule of thumb is to increase $R_{PE}^p$ since the self-exchange energy is typically more dominant and comprises less terms (hence cheaper) than the pair-exchange one. In this example, one may reduce the overall error by half (to $\sim 0.01\%$) by going from $R_{PE}^s = 6.0$ bohr to $R_{PE}^s = 7.0$ bohr at a negligible increase ($< 1\%$) in computational cost of the exx module.
Figure 4.1: Convergence test on $E_{xx}$ based on a liquid $(\text{H}_2\text{O})_{64}$ snapshot. The relative errors of $E_{xx}$ compared to $E_{xx}^{\text{ref}}$ are evaluated individually with respect to $R_{\text{pair}}$ (black), $R_{\text{PE}}^s$ (blue), and $R_{\text{PE}}^p$ (blue dashed). The overall relative error ($\sim 0.02\%$) of our chosen parameters is depicted with the green line.

The above test determined the sizes of the $\Theta(C_{ij}, R_{\text{PE}}^s)$ and $\Theta(C_{ij}, R_{\text{PE}}^p)$ subdomains that are able to cover the MLWF product-densities and to restrict the integration of $E_{xx}$ and $\sigma_{xx}$ in Eqs. (2.28) and (2.32), respectively. As such, no further test is done on the convergence of $\sigma_{xx}$; however, when extreme accuracy on $\sigma_{xx}$ becomes relevant, similar convergence test should also be performed based on $\sigma_{xx}$ in analogy to $E_{xx}$. We note in passing that using the set of converged radii, the cost of computing $\sigma_{xx}$ from $\tilde{v}_{ij}$ via Eq. (2.32) took $< 0.5\%$ of the gross computational cost of the $\text{exx}$ module; hence, the stress contribution can always be evaluated without introducing noticeable overhead.
In this test for $E_{xx}$, we have neglected the effect of orbital self-consistency. Now, we address this effect by allowing fully self-consistent calculation of $E_{xx}$ using the chosen parameters for $R_{\text{pair}}$, $R_{\text{PE}}^s$, and $R_{\text{PE}}^p$, while keeping $R_{\text{ME}}^s$ and $R_{\text{ME}}^p$ at the reference values. In doing so, we found that the self-consistency varies $E_{xx}$ by a negligible amount ($< 0.01\%$). We note in passing that to achieve very tight self-consistency also depends on the localization of the MLWFs (see Appendix A for a detailed discussion). This negligible change also manifests that the far-field (beyond $R_{\text{PE}}$) wavefunction forces evaluated by ME are sufficiently accurate. This self-consistent calculation will later be referred to as the “intermediate-reference” calculation, which provides the reference EXX contributions to the wavefunction forces $\{D_{xx}^{i,\text{ref}}(r)\}$ and the MLWFs for non-self-consistent evaluations of $\{D_{xx}^i(r)\}$ for given $R_{\text{ME}}^s$ and $R_{\text{ME}}^p$ in Sec. 4.1.2. In this case, the original reference calculation, in which all parameters are chosen to their largest possible values, cannot be used in place of the intermediate reference due to its lack of flexibility to vary $R_{\text{ME}}$ (i.e., $R_{\text{ME}}$ is bounded below by the reference setting of $R_{\text{PE}} = 11.11 \text{ Å}$ to provide boundary conditions for the PE).

4.1.2 Convergence of Wavefunction Forces

Based on the MLWFs and $\{D_{xx}^{i,\text{ref}}(r)\}$ provided in the intermediate-reference calculation (see Sec. 4.1.1), we are in position to check the convergence of the EXX contributions to the wavefunctions forces $\{D_{xx}^i(r)\}$ (obtained non-self-consistently from the intermediate-reference calculation) associated with the given $R_{\text{ME}}^s$ and $R_{\text{ME}}^p$. For simplicity, we measure the convergence of $\{D_{xx}^i(r)\}$ based on the average of the relative
mean absolute deviations ($\{\gamma_{xx}^i\}$) over the MLWFs defined as:

$$\gamma_{xx} = \frac{1}{N_o} \sum_i \gamma_{xx}^i$$

$$\gamma_{xx}^i = \frac{\int dr \left| \tilde{D}_{xx}^{i,\text{ref}}(r) - \tilde{D}_{xx}^i(r) \right|}{\int dr \left| \tilde{D}_{xx}^{i,\text{ref}}(r) \right|}. \quad (4.1)$$

When the convergence is achieved, both $R_{\text{ME}}^s$ and $R_{\text{ME}}^p$ are large enough, $\{\tilde{D}_{xx}^i(r)\} = \{\tilde{D}_{xx}^{i,\text{ref}}(r)\}$, $\gamma_{xx}^i = 0 \ \forall \ i$, and hence $\gamma_{xx} = 0$. For smaller $R_{\text{ME}}^s$ and/or $R_{\text{ME}}^p$, the larger deviation between $\{\tilde{D}_{xx}^i(r)\}$ and $\{\tilde{D}_{xx}^{i,\text{ref}}(r)\}$ will lead to more positive $\gamma_{xx}$. Based on this measure, we explore the individual convergence of $R_{\text{ME}}^s$ and $R_{\text{ME}}^p$ (i.e., vary one parameter, while keeping the other at its reference value) as shown in Fig. 4.2.

Since it is less obvious to determine a reasonable error tolerance for wavefunction forces, we performed an auxiliary test on the constant of motion (CoM) of NVT CPMD simulation at PBE0 level for a $(\text{H}_2\text{O})_{32}$ liquid at 400 K using a planewave cutoff at 40 Ry with several combinations of $R_{\text{ME}}^s$ and $R_{\text{ME}}^p$; we note that in this case $L = 18.67$ bohr and the largest possible $R_{\text{ME}} = L/2 = 9.3$ bohr. If these radii are not chosen large enough tiny errors in the wavefunction forces near the tail of the MLWFs can cause minor drifts in the CoM in CPMD simulations. We found that $R_{\text{ME}}^s = 9.3$ bohr and $R_{\text{ME}}^p = 7.0$ bohr provided a reasonably conserved CoM (with a drift $\sim 0.2$ meV/ps-atom during this test simulation, which corresponds to a tolerance of $\sim 0.2\%$ as shown in Fig. 4.2; this test leads to our choice of $R_{\text{ME}}^s = 10.0$ Å (if the cell can accommodate) and $R_{\text{ME}}^p = 7.0$ Å. We note in passing that for larger systems such as $(\text{H}_2\text{O})_{64}$ liquid, $R_{\text{ME}}^s$ can be chosen to be 11.0 bohr such that the overall error tolerance on the wavefunction forces can be lowered by half $\sim 0.1\%$, since it does not add to the communication overhead and creates only minor increase in the memory usage.
Figure 4.2: Convergence tests on $\{\widetilde{D}_{\text{xx}}(r)\}$ based on a liquid (H$_2$O)$_{64}$ snapshot. The convergence is measured via $\gamma_{\text{xx}}$—the average of the relative mean absolute deviations $\{\gamma_{\text{xx}}^i\}$ defined in Eq. (4.1), and is evaluated individually with respect to $R_{\text{ME}}^s$ (red) and $R_{\text{ME}}^p$ (red dashed). Based on our chosen parameters, the error threshold ($\gamma_{\text{xx}} = 0.2\%$) is depicted with the green line along with a standard deviation of $\sim 0.03\%$ shown in green dashed lines.

We expect these five radial cutoffs (i.e., $R_{\text{pair}} = 8.0$ bohr, $R_{\text{PE}}^s = 6.0$ bohr, $R_{\text{PE}}^p = 5.0$ bohr, $R_{\text{ME}}^s = 10.0$ bohr, and $R_{\text{ME}}^p = 7.0$ bohr) determined here for a system of bulk liquid water to be semi-quantitatively transferable to other large band gap homogeneous materials. For systems with relatively small gap, a test calculation should be performed to determine optimal values of these radial cutoffs prior to performing large scale production runs such as AIMD simulations.
4.2 Parallel Scaling and Performance

The judicious choice of the five radii allows one to obtain formally exact evaluation of the EXX related quantities while simultaneously maintaining minimal computational cost. Despite that this computational cost is minimized and scales linearly with system size, it is still highly demanding in absolute scale. Hence, the feasibility of performing large-scale hybrid DFT-based AIMD simulations using the \texttt{exx} module relies critically on the efficiency of the massively parallel algorithm. Our implementation (described in Chapter 3) seamlessly distributes the major computational workload coming from the evaluation of tens of thousands of MLWF-product potentials among several thousands of MPI processes. Within each MPI process, the CG-based PE solver is further parallelized over many intranode OpenMP threads. The efficiency in the scaling of the internode MPI (first-layer) and intranode OpenMP (second-layer) parallelization schemes were tested using the IBM Blue Gene/Q platform Mira and will be discussed below in Sec. 4.2.1 and 4.2.2, respectively.

4.2.1 Internode Parallelization \textit{via} MPI

For the first-layer parallelization, the \texttt{exx} module uses internode MPI communication to spread the tens of thousands of PEs to thousands of computing nodes, which is hence the heart of this massively parallel algorithm. To provide a detailed characterization for this critical parallelization level, we chose to perform both a strong-scaling test (\textit{i.e.}, fixed problem size with varying number of processing elements) and a weak-scaling analysis (\textit{i.e.}, vary problem size with fixed ratio of the problem size to the number of processing elements). In case of water, we use the number of water molecules ($N_{\text{water}}$)—which determines the number of valence orbitals ($N_o = 4N_{\text{water}}$)—as a measure of system (problem) size; as such, the strong-scaling and weak-scaling tests were performed with fixed $N_{\text{water}}$ and $\zeta = N_{\text{proc}}/N_o$, respectively. To investigate the strong
and weak scaling, we performed a total of 12 CPMD simulations on liquid water with
density of 0.993 g/cm$^3$, in which we vary system sizes $N_{\text{water}} = 64, 128, 256$ and
parallelization level $\zeta = 1/2, 1, 2, 4$. In these calculations, we used one MPI process
per node and used all 64 hyperthreads available on each node for intranode OpenMP
parallelization to run 50 CPMD steps at 330 K.
<table>
<thead>
<tr>
<th>Parameters</th>
<th>QE Module Timings (in seconds)</th>
<th>Breakdown of exx Timings (in seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t\text{GGA} &amp; t\text{MLWF} &amp; t\text{exx} &amp; t\text{exx}/t\text{GGA} &amp; t\text{comp} (f\text{comp}) &amp; t\text{comm} (f\text{comm}) &amp; t\text{idle} (f\text{idle})</td>
<td></td>
</tr>
<tr>
<td>N_{\text{water}} &amp; N_z &amp; \zeta &amp; N_{\text{proc}} &amp; N_{\text{tg}} &amp; 2.81 &amp; 0.16 &amp; 7.34 &amp; 2.61 &amp; 4.07 (55.4) &amp; 0.96 (13.1) &amp; 2.31 (31.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64 &amp; 140 &amp; 1/2 &amp; 128 &amp; 1 &amp; 1.97 &amp; 0.17 &amp; 3.83 &amp; 1.94 &amp; 2.05 (53.4) &amp; 0.52 (13.5) &amp; 1.27 (33.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64 &amp; 140 &amp; 1 &amp; 256 &amp; 1 &amp; 1.02 &amp; 0.16 &amp; 2.74 &amp; 2.68 &amp; 1.06 (38.9) &amp; 0.39 (14.3) &amp; 1.28 (46.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64 &amp; 140 &amp; 4 &amp; 1024 &amp; 4 &amp; 0.63 &amp; 0.16 &amp; 1.70 &amp; 2.70 &amp; 0.54 (32.0) &amp; 0.37 (21.6) &amp; 0.79 (46.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>128 &amp; 176 &amp; 1/2 &amp; 256 &amp; 1 &amp; 4.60 (55.6) &amp; 1.21 (14.7) &amp; 2.46 (29.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>128 &amp; 176 &amp; 1 &amp; 512 &amp; 2 &amp; 2.64 &amp; 0.43 &amp; 4.35 &amp; 1.65 &amp; 2.35 (54.1) &amp; 0.64 (14.8) &amp; 1.36 (31.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>128 &amp; 176 &amp; 2 &amp; 1024 &amp; 4 &amp; 3.04 &amp; 1.94 &amp; 1.57 &amp; 0.41 &amp; 1.25 (41.0) &amp; 0.51 (16.9) &amp; 1.28 (42.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>128 &amp; 176 &amp; 4 &amp; 2048 &amp; 8 &amp; 1.96 &amp; 2.04 &amp; 0.96 &amp; 0.41 &amp; 0.67 (34.4) &amp; 0.48 (24.8) &amp; 0.80 (40.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>256 &amp; 220 &amp; 1/2 &amp; 512 &amp; 2 &amp; 8.27 &amp; 1.59 &amp; 5.19 &amp; 1.43 &amp; 4.19 (50.2) &amp; 1.58 (18.9) &amp; 2.57 (30.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>256 &amp; 220 &amp; 1 &amp; 1024 &amp; 4 &amp; 6.39 &amp; 2.77 &amp; 8.34 &amp; 1.31 &amp; 2.23 (46.5) &amp; 1.11 (23.2) &amp; 1.46 (30.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>256 &amp; 220 &amp; 2 &amp; 2048 &amp; 8 &amp; 2.33 &amp; 1.49 &amp; 2.23 &amp; 1.13 &amp; 1.26 (38.0) &amp; 0.76 (22.9) &amp; 1.30 (39.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>256 &amp; 220 &amp; 4 &amp; 4096 &amp; 16 &amp; 2.41 &amp; 1.52 &amp; 1.59 &amp; 1.08 &amp; 0.77 (31.9) &amp; 0.82 (34.2) &amp; 0.82 (33.9)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1: Profiling data collected from the 12 short CPMD simulations of liquid water systems (each simulation is triplicated to gauge the reproducibility). The system size and level of MPI parallelization are controlled by $N_{\text{water}}$ and $\zeta$, respectively. In each calculation, QE partitions the $N_o$ bands to $N_{\text{tg}}$ parallel 3D FFT executions, each of which is performed with $N_z$ MPI processes if $N_z \times N_{\text{tg}} \leq N_{\text{proc}}$. When $N_z > N_{\text{proc}}$, $N_{\text{tg}}$ is set to 1 (as depicted in the case with an asterisk). The resulting timings include the average walltime costs spent in the GGA ($t_{\text{GGA}}$), Wannierlization ($t_{\text{MLWF}}$), and exx ($t_{\text{exx}}$) modules. The exx timings are further broken down to fractions coming from computation ($f_{\text{comp}}$), communication ($f_{\text{comm}}$), and idle ($f_{\text{idle}}$); note that these fractions are in percents. The fluctuations in the timings (among the triplicates) is smaller than the reported precision; hence, error bar is neglected.
The initial structures are generated systematically in the following three steps: (i) randomly packing the given number of water molecules into a cubic unit cell whose cell dimension satisfies the targeted density; (ii) equilibrating the randomly packed structure via MD simulation using TIP4P2005 force field with fixed simulation cell at 300 K for 1.0 ns; (iii) further equilibrating the TIP4P2005 liquid via CPMD simulation using PBE0 xc functional with fixed cell at 330 K. In the CPMD simulation, the ionic temperatures were controlled with the massive Nose-Hoover chain thermostats, each with a chain length of 4. The nuclear and electronic degrees of freedom were integrated using the standard Verlet algorithm and a time step of 2.0 a.u. (∼0.05 fs). For each system size, the cell is kept fixed throughout the simulation. To ensure an adiabatic separation between the electronic and nuclear degrees of freedom in the CP dynamics, we used a fictitious electronic mass of 100 a.u. and the nuclear mass of deuterium for each hydrogen atom. The interactions between the valence electrons and the ions (consisting of the nuclei and their corresponding frozen-core electrons) were treated with Hamann-Schlüter-Chiang-Vanderbilt type norm-conserving pseudopotentials distributed with the Qbox package. The valence electronic pseudo-wavefunctions were expanded using a planewave basis set up to a planewave kinetic energy cutoff at 85 Ry. Mass preconditioning was applied to all Fourier components of the electronic wavefunctions having a kinetic energy greater than 25 Ry.

The average timings (triplicated) of the 12 CPMD simulations (with the same setting as the 500 PBE0 equilibration steps) were broken down to contributions from the GGA part (including pseudopotential calculations), Wannierlization procedure, and the exx module as listed in Table 4.1. For the GGA part in QE, the FFT calculation—the algorithmic bottleneck—employs two levels of parallelization. The first level splits the 3D FFT into 2D (z-planes along xy directions) and 1D (z sticks) FFTs, where the \( N_z \) 2D FFTs (\( N_z \) determined by planewave cutoff) are distributed and evaluated in parallel by the pool of available MPI processes. The second level on top of the
first one further uses $N_{tg}$ “taskgroups” to parallelize the $N_o$ 3D FFTs. As such, the GGA part can effectively utilize $N_z \times N_{tg}$ MPI processes. Wannierlization takes (on average) merely 3–4 SODD steps per CPMD step to reach maximal localization and only provides a minor contribution to the walltime cost. On the other hand, The exx module remains the overall bottleneck of the hybrid DFT simulation, however, with a comparable cost to the GGA part (within a factor of 3 as shown in the $t_{exx}/t_{GGA}$ column). We note in passing that the timings of the exx module (Fig. 3.1) does not include the cost of invFFT step used to generate real-space MLWFs. We also note that the partitions of the available pool of processors are optimized for the exx module—the bottleneck of these hybrid DFT calculations; as such, these settings may not be most efficient for the GGA part (our development version is compatible with that of QE-5.0.2, in which the FFT implementation does not scale well with OpenMP parallelization). We expect an enhanced scalability with respect to OpenMP in the GGA part in newer QE versions where threaded FFT support is implemented.

Based on above set of calculations, we obtained the strong-scaling behavior of the average walltime cost spent in the exx module—as a function of $\zeta$, i.e., $t_{exx}[\zeta]$—for each system size (see Fig. 4.3). For each system size (i.e., $N_{\text{water}}$), we chose $\zeta = 1/2$ (a reasonable/realistic setup for GGA calculation on these systems) as the baseline for computing the ideal scaling in the definition of the strong-scaling efficiency as

$$\eta_{MPI}^{\text{exx}}(\zeta) \equiv \frac{t_{exx}[\zeta = 1/2]}{2 \zeta t_{exx}[\zeta]}.$$ (4.2)

Using this definition, the strong-scaling efficiency (averaged over $(\text{H}_2\text{O})_{64}$, $(\text{H}_2\text{O})_{128}$, and $(\text{H}_2\text{O})_{256}$) gradually decreases with higher level of parallelization from $\sim 93\%$ ($\zeta = 1$), to $\sim 66\%$ ($\zeta = 2$), and to $\sim 50\%$ ($\zeta = 4$). We did not go beyond $\zeta = 4$ since liquid water systems run short of PEs to be solved (per process wise).
This gradual degradation of the strong-scaling efficiency originates from a combined effects from the computation, the communication, and the idle time (due to workload imbalance) as shown in Fig. 4.3 with detailed breakdown listed in Table 4.1. For the computation, all steps scale nearly ideally (> 90% efficiency) except for the serial Construction of Pair List step (i.e., the static load-balancing algorithm). This serial step is often quite cheap for medium sized systems (e.g., (H$_2$O)$_{64}$ and (H$_2$O)$_{128}$); however, it dominates in larger system (e.g., (H$_2$O)$_{256}$) due to its non-linear scaling behavior with system size. This non-linear scaling cost can be mitigated with a Verlet list algorithm\textsuperscript{143} to avoid unnecessary consideration of distant pairs in every

![Figure 4.3: The strong-scaling test via MPI on the exx module. The ideal scaling (based on $\zeta = 1/2$ timings) for each test is given as dashed lines. The small pie plots demonstrate the breakdown of the fractions of CPU time spent in computation ($f_{\text{comp}}$, colored), communication ($f_{\text{comm}}$, black), and idle due to load imbalance ($f_{\text{idle}}$, white).](image-url)
MD step. In addition to computation, the communication also contributes to the decay of efficiency towards the strong-scaling limit. This non-ideal behavior in communication originates from 

(i) the sending/receiving of MLWFs and the wavefunction forces and

(ii) the redistribution back and forth between the ORBITAL and GRID data distribution schemes. The former is relevant for smaller systems \(((H_2O)_{64}\text{ and } (H_2O)_{128})\), while the latter dominates in larger scales \(((H_2O)_{256})\) due to its associated all-to-all communication. To reduce the cost of communication and its deleterious effects on the strong-scaling efficiency, a scheme with localization-based redistribution and asynchronous (non-blocking) sending/receiving operations that overlap with the PE solution will be discussed in a forthcoming paper.

The last and the most critical issue that limits the strong-scaling efficiency in this algorithm is the idle time due to workload imbalance. This imbalance arises from the imperfect spread of the available number of PEs (weighted by the variation of required CG steps to solution) across the available pool of processes using our static load-balancing algorithm (Sec. 3.3.2). This static load-balancing algorithm is subject to the following three constraints: 

(i) lack of dynamical scheduling, 

(ii) based on the assumption that each MLWF introduce identical amount of workload, and

(iii) limiting the solution of EXX contributions from a pair of MLWFs \(\tilde{\phi}_i\) and \(\tilde{\phi}_j\) on either \(P_i\) or \(P_j\). Hence, the current \texttt{exx} module faces challenge to balance workloads when there is an imbalance in the number of pairs due to the physical nature of the problem, \(i.e.,\) the presence of (instantaneous) heterogeneity such as in liquids or near interfaces. One simple route to obtain a reasonable load-balancing for heterogeneous system is to introduce different levels of parallelization \(\zeta\) for different types of MLWFs, which lifts the second assumption. For example, consider a surface where each MLWF in the bulk and at the surface have four and two overlapping orbitals, respectively. The workload originates from a bulk MLWF is twice the amount of a surface MLWF. To obtain a manageable load-balancing for such surface, we can employed
Figure 4.4: Weak-scaling ($\zeta$ fixed along each line) test via MPI on the exx module with respect to $N_{\text{water}}$ water liquid. The ideal scaling (based on $(\text{H}_2\text{O})_{64}$ timings) for each test is given as dashed lines.

\[ \zeta_{\text{bulk}} = \frac{N_{\text{bulk}}}{N_{\text{o}}^{\text{bulk}}} = 2 \zeta_{\text{surf}} = \frac{2N_{\text{surf}}}{N_{\text{o}}^{\text{surf}}} \]

where $\zeta_{\text{bulk}}$ and $\zeta_{\text{surf}}$ are the levels of parallelization for bulk and surface, respectively, determined based on their corresponding numbers of processes ($N_{\text{bulk}}^{\text{proc}}$ and $N_{\text{surf}}^{\text{proc}}$) and orbitals ($N_{\text{o}}^{\text{bulk}}$ and $N_{\text{o}}^{\text{surf}}$). As such, load-balancing is accomplished by compensating the twice heavier workload for MLWFs in the bulk than at the surface with correspondingly doubled parallelization level. Other more sophisticated load-balancing algorithms that mitigate this challenge by removing one or more of these constraints will be addressed in future work.

The same set of data can also be used for a weak-scaling analysis (i.e., vary problem size with fixed ratio of the problem size to the number of processing elements) via MPI as shown in Fig. 4.4. Within a highly parallel setup ($\zeta = 4$), the exx module is
able to run at a walltime cost lower than 3 seconds per MD step even for the largest system used herein, i.e., (H$_2$O)$_{256}$ liquid. This walltime cost makes it possible to perform long (e.g., 10–100 ps) CPMD simulations within an affordable human time. The weak-scaling efficiency (for each fixed $\zeta$) is defined as:

$$\eta_{w}\text{MPI}(N_{\text{water}}) \equiv \frac{t_{\text{exx}}[N_{\text{water}} = 64]}{t_{\text{exx}}[N_{\text{water}}]}, \quad (4.3)$$

in which $t_{\text{exx}}[N_{\text{water}}]$ is the average walltime cost over 50 CPMD steps for a system of $N_{\text{water}}$ water molecules. As shown in Fig. 4.4 the exx module is quite scalable to the large system size limit. In this case, $\eta_{w}\text{MPI}$ (average over $\zeta = 1/2, 1, 2, 4$) for (H$_2$O)$_{128}$ and (H$_2$O)$_{256}$ liquids retain within 89% and 81%, respectively. For lower level of parallelization ($\zeta = 1/2$), the main obstacle for scalability to larger system sizes (e.g., (H$_2$O)$_{256}$) is the workload imbalance. For higher level of parallelization ($\zeta \geq 1$), the communication (both the send/receive and the redistribution processes) impairs scaling efficiency at large-scale. The upward tilt (faster decaying $\eta_{\text{MPI}}^{rmw}$) in the (H$_2$O)$_{256}$ weak-scaling line in Fig. 4.4 is due to heavier communication in larger systems which reduces the scalability. Aside from the aforementioned limitations on the load-balancing and communication left to be addressed in future studies, the exx module provide an accurate and efficient general purpose framework to perform large-scale hybrid DFT-based AIMD simulations for condensed-phase systems with walltime cost comparable to the less-scalable semi-local functionals when provided an appropriate amount of processors.

### 4.2.2 Intranode Parallelization via OpenMP

For the intranode OpenMP parallelization, threading is widely applied for operations within the exx module to speedup the loading/off-loading of variables, calculation of the ME, and the CG-based Poisson solver, which is the computational bottleneck of
this algorithm. In this case, we assess the performance of the OpenMP parallelization via a strong-scaling test (i.e., fixed problem size with varying number of processing elements) based on $t_{\text{Step IV}}[N_{\text{thread}}]$—the averaged walltime cost for solving Step IV (in Fig. 3.1) over 50 CPMD steps of an (H$_2$O)$_{64}$ liquid when $N_{\text{thread}}$ OpenMP threads is activated. Here, we measure the OpenMP scalability as a function of the number of

Figure 4.5: Strong-scaling test on the average walltime cost for solving the PEs with respect to the number of OpenMP threads based on a 50-step CPMD trajectory of an (H$_2$O)$_{64}$ liquid. The test were performed at planewave cutoffs at both 85 Ry (circles) and 150 Ry (triangles) using different number of physical cores ($N_{\text{core}}$, in blue). Additional speedup can be achieved by hyperthreading the 16 physical cores in each node into 64 threads (in red). The ideal scaling using physical cores is shown in black dashed line.
threads per process \((N_{\text{thread}})\) with the strong scaling efficiency defined as:

\[
\eta_{\text{OpenMP}}^{s}(N_{\text{thread}}) \equiv \frac{t_{\text{Step IV}}[N_{\text{thread}} = 1]}{N_{\text{thread}}t_{\text{Step IV}}[N_{\text{thread}}]}, \tag{4.4}
\]

To keep a consistent internode communication pattern, we choose one MPI process per node with \(\zeta = 1\) parallelization, and five different levels of OpenMP parallelization level (by activating 1, 2, 4, 8, and 16 physical cores per node) to perform the strong-scaling test on the OpenMP threads. This test shows that the OpenMP parallelization is close to ideal \((\eta_{\text{OpenMP}}^{s} > 80\%, \text{when all 16 physical cores are activated})\) as shown in Fig. 4.5. With heavier workload from a denser real-space grid (planewave cutoff at 150 Ry, appropriate for constant-pressure simulations), this algorithm preserves 92% of the strong-scaling efficiency when using all 16 physical cores; for lighter workload (85 Ry, suitable for simulations with fixed cells), a 84% efficiency is observed. An additional 30%–40% speedup can be achieved by hyperthreading physical cores into more threads.

4.3 Crystalline Symmetry and Performance

Another relevant factor affecting the computational cost of the exx module is the symmetry of the unit cell. As discussed in Sec. 3.3.4, the size of the Laplacian stencil \((6n + 2nN_{\text{off}} + 1)\) depends on the number of non-zero off-diagonal elements \((N_{\text{off}})\) in the upper (or lower) triangle of \(F\) \((cf.\ Eqs. (3.6)–(3.7))\). Since the cost of solving PE (the computational bottleneck of the exx module) grows proportionally with this size of the stencil, it is advantageous to minimize \(N_{\text{off}}\). In this section, we discuss the relationship between \(h\) and \(N_{\text{off}}\).

Before discussing how the symmetry of \(h\) translates to \(N_{\text{off}}\), we note that \(F\) is invariant to the orientation of \(h\) \((i.e., \text{when } h \text{ is subject to an arbitrary 3D orthogonal})\)

\[\text{In IBM Blue Gene/Q system, each physical core can be split into 4 hyperthreads}\]
rotation \widehat{O} to give new \( \vec{h} = \widehat{O} \vec{h} \). This effect can be easily verified by transforming \( \mathbf{F} \) (cf. Eqs. (3.6)–(3.7)),

\[
F_{\alpha\beta} = \sum_a J_{\alpha a} J_{\alpha \beta} = |L_\alpha| |L_\beta| \sum_a (h^{-1})_{\alpha a} (h^{-1})_{\beta a}
= |L_\alpha| |L_\beta| (h^{-1}(h^{-1})^T)_{\alpha\beta},
\]

(4.5)

and identifying that \( h^{-1}(h^{-1})^T = h^{-1}\widehat{O}^{-1}(\widehat{O}^{-1})^T(h^{-1})^T = h^{-1}(\vec{h})^{-1})^T \), where \( \vec{h}^{-1} = h^{-1}\widehat{O}^{-1} \) and \( (\widehat{O}^{-1})^T = \widehat{O} \). Given the invariance of the \( \mathbf{F} \) with the orientation of the cell, it is convenient to assume that \( \vec{h} \) is upper triangular. Based on the six independent variables from this cell tensor, off-diagonal terms of the \( \mathbf{F} \) matrix can be discussed for different Bravais lattices. By straightforward algebraic operations, we found that \( N_{\text{off}} = 0 \) when the cell is orthorhombic (no non-zero off-diagonal term in the upper triangular cell), \( N_{\text{off}} = 1 \) when monoclinic (one non-zero off-diagonal term in the cell), and \( N_{\text{off}} = 3 \) when triclinic (three non-zero off-diagonal terms in the cell). However, it is interesting to note that when there are two non-zero off-diagonal terms in the cell tensor, \( N_{\text{off}} \) could be 2 (if \( h_{23} = 0 \)) or 3 (otherwise). In other word, additional saving can be obtained for triclinic cells if \( L_3 \) falls in the \( xz \) plane when \( L_1 \) aligns to \( \hat{x} \) and \( L_2 \) is in the \( xy \) plane.

In view of efficient setup for the \texttt{exx} module, it is ideal to use orthorhombic cell whenever possible. Similarly, for an \( NpT \) simulation, it is more efficient to enforce unit cell symmetry for orthorhombic or monoclinic systems. Moreover, judicious selection of supercells from certain non-orthorhombic cells can be made orthorhombic, \emph{e.g.}, for hexagonal or trigonal crystal systems. For a hexagonal cell (let the angle between \( L_1 \) and \( L_2 \) to be the only non-orthogonal angle), one can choose an orthorhombic supercell with new lattice vectors \( \{L'_1, L'_2, L'_3\} \) such that \( L'_1 = L_1, L'_2 = 2L_2 + L_1, \) and \( L'_3 = L_3 \). This supercell contains twice as many particles as in the unit cell. In case of ice \( \text{Ih} \) example in Chapter 5, we use the above transformation on an \((\text{H}_2\text{O})_{12}\)
cell and double both $L'_1$ and $L'_3$ to generate a $(\text{H}_2\text{O})_{96}$ supercell to allow freedom for the introduction of higher level of proton disorder and to allow sufficient sampling of the Brillouin zone by only the $\Gamma$ point. Similar transformation can be done to transform other trigonal crystals such as rhombohedral cells (e.g., ice II) into (three times larger) hexagonal and then (six times larger) orthorhombic supercells. However, such transformation is not used in the triple-point simulation in Chapter 5 due to its high anisotropy in the size of the resulting orthorhombic cell (with largest lattice vector of $\sim 22.5$ Å and shortest $\sim 6.3$ Å). Such orthorhombic cell requires a supercell of $(\text{H}_2\text{O})_{144}$ to be suitable for $\Gamma$-point only algorithm (typically requires $|L_\alpha| > 10$ Å $\forall \alpha$) as it is computationally very demanding; hence, we choose to use rhombohedral cell in case of ice II.

### 4.4 Conclusions and Future Outlook

In this chapter, I have provided a systematic performance optimization of the exx module while simultaneously keeping a formally exact evaluation of the EXX related quantities. To achieve higher computational efficiency, we employ distinct the proto-subdomains for self- and pair-exchange calculations, respectively. The proto-subdomains can be chosen smaller to speed up the pair-exchange calculation (which is the main computational bottleneck) than the self-exchange case since the pair-exchange case has smaller orbital overlap and lacks monopole term (most long-range term) in the ME. For this purpose, we introduced a set of user-defined parameters to fine tune the balance between computational cost and accuracy and demonstrated a systematic optimization of these parameters using a $(\text{H}_2\text{O})_{64}$ liquid snapshot. Based on this optimized parameter set, we investigate the parallel scaling of the exx module with respect to MPI and OpenMP schemes using IBM Blue Gene/Q supercomputer Mira. In this test, we found exx module is quite scalable toward the strong-scaling
limit even with a highly parallel setting ($\zeta = 4$). In a more detailed breakdown of the \texttt{exx} timings, we found that the scalability is affected by serial computation step (generation of $\mathcal{L}$), interprocess communication, and idle due to load imbalance. In the weak-scaling analysis, the scalability is close to ideal (efficiency greater than 80%) even for realistic system size ($\text{(H}_2\text{O})_{256}$). In this case, an EXX calculation of $\text{(H}_2\text{O})_{256}$ can be performed within 3 seconds (walltime) per MD step making integration of long AIMD trajectory a feasible task. On the other hand, the \texttt{exx} module is also highly scalable with respect to \texttt{OpenMP} threading (in terms of the computational bottleneck—the solution for $\bar{v}_{ij}(r)$). On Mira more than 80% of strong-scaling efficiency is preserved when all 16 physical cores on each node is activated; additional, 30%–40% speedup can even be achieved with the use of hyperthreading technology.

Our scheme bares limitations to simulate metallic systems; in this case, hybrid DFT with a screened exchange should be used.\textsuperscript{63} For an inhomogeneous system, additional local grids are required to capture the different level of localizations; application of the recursive subspace bisection method\textsuperscript{78,89} may be more convenient. We expect potential improvement of computational efficiency if one apply the adaptively compressed exchange (ACE) operator.\textsuperscript{72} Our algorithm can also be applied to hybridize the highly accurate strongly-constrained-and-appropriately-normed meta-GGA functional.\textsuperscript{46,144} The further information can be found on the online manual of \texttt{QE}.\textsuperscript{†}

Although the \texttt{exx} module is quite efficient and has enabled numerous findings mentioned in this thesis, there is large room for improvement. Future improvements on the \texttt{exx} module include computation, communication, and workload balancing. For the most critical issue—workload balancing, we plan to develop a dynamic and graph-theoretic scheduling scheme to mitigate the idle time in our current static load-balancing algorithm. For the communication overhead, we will asynchronously

overlap the interprocess send/receive operations with computation as well as exploit
the localization of MLWFs to restrict the real-space domain for the redistribution
between GRID and ORBITAL schemes. For the computation, we will employ more
advance PE solvers (e.g., preconditioned conjugate gradient method and multigrid
approach) as well as state-of-the-art high-performance computing technology (e.g.,
graphical accelerators, advanced vectorization, and hybrid field-programmable gate
array architectures) to boost the exx module. In addition to the performance of exx
module, we also plan to generalize this approach to range-separated hybrid DFT and
other post-Hartree-Fock methods.
Part II

Understanding Molecular Crystals and Liquids
Chapter 5

Application to Ice Polymorphs at
the Ice Ih–II–III Triple Point

As discussed in the first part of this thesis, the current massively parallel implementation of our MLWF-based exx module enables AIMD simulations of large-scale condensed-phase materials (for long timescales) at the hybrid DFT level of theory. Initial versions of this algorithm have already been successfully used to perform a number of applications, which include highly accurate hybrid DFT based investigations into the electronic structure of various semi-conducting solids,\textsuperscript{119,120} the structural properties of ambient liquid water\textsuperscript{8,121} and aqueous ionic solutions,\textsuperscript{122} the structure and dynamics of the solvated $\text{H}_3\text{O}^+$ and $\text{OH}^-$ ions,\textsuperscript{123} as well as crystal structure prediction of molecular crystals.\textsuperscript{125} However, all of these simulations were performed either with given configurations or in the $NVT$ ensemble, in which the simulation cells (and hence the number density of these systems) were kept fixed. Here, we report our findings from the simulations performed in the $NpT$ ensemble using hybrid xc functionals to determine the equilibrium densities and phase transition properties of various ice polymorphs. Water ice, the most common molecular solid in nature, exhibits a rich and complex phase diagram that includes 16 distinct crystalline poly-
morphs in which water molecules possess a number of distinct arrangements, H-bond
(H-bond) networks, and densities. Recent studies have pointed out that an ac-
curate assessment of the phase transition properties among these ice phases is quite
challenging since their stabilities hinge upon a delicate balance between H bonding
and van der Waals (vdW) interactions. It was shown that vdW inclusive hybrid xc functionals can predict phase transition properties among ice Ih and many
other high-density phases of ice with much greater accuracy than that obtained from
GGA xc functionals. All of these studies were focused on the ice structures and
phase transition properties at very low temperatures where proton ordered phases
of ice are predominant. Here, we focus on a specific triple point in the water
phase diagram, which occurs between ice Ih, II, and III at 238 K and 2.1 kbar, in
order to demonstrate the accuracy of vdW inclusive hybrid xc calculations at realistic
temperatures and pressures as well as the performance of our algorithm for \( NpT \)
simulations.

Figure 5.1: Structures of the Ice Ih, Ice II, and Ice III phases at the experimental
triple point (238 K and 2.1 kbar) obtained at the PBE0+vdW\textsuperscript{TS} level of theory. The
blue boxes refer to the simulation cells, each of which contains 96 water molecules.
5.1 Methodology

The initial structures of these ices were constructed by employing the experimentally known symmetry of the structures. In order to accurately capture the effects of proton disorder, large supercell geometries containing 96 water molecules were used in the simulations as shown in Fig. 5.1. We have constructed an orthorhombic \(2 \times 2 \times 2\) supercell of ice Ih (subject to the transformation discussed in Sec. 4.3), a trigonal \(2 \times 2 \times 2\) supercell of ice II, and a tetragonal \(2 \times 2 \times 2\) supercell of ice III. Then proton disorder is introduced with an algorithm which combines the Bernal-Fowler ice rules\(^{153}\) and the constraint of vanishing net dipole moment in the simulation cell.\(^{154,155}\) The distinctive structures of these ices possess quite different topology of the H-bond network which can be quantified using the statistics of closed rings\(^*\) as shown in Fig. 5.2(a). The hollow structure of Ice Ih has only 6-membered rings, whereas, more compact ice II has only 23% six-membered rings and a large number of eight-membered (29%) and ten-membered (48%) rings. On the contrary, ice III has no six-membered H-bonded rings and has an equal distribution of five-, seven-, and eight-membered rings.

We have performed simulations in the \(NpT\) ensemble at the experimental triple point (238 K and 2.1 kbar) in order to compute the equilibrium volumes and enthalpies of the three ice phases using PBE0+vdW\(^{TS}\)\(_{SC}\)\(^{34,36}\) functional and compared with the corresponding experimental data. Also the effects of exact exchange and vdW interactions are demonstrated \textit{via} comparative analysis with simulations performed with PBE and PBE+vdW\(^{TS}\)\(_{SC}\) functionals. Each simulation was run for at least 6 ps and the \(NpT\) sampling was collected from the last 5 ps, which is sufficient to converge equilibrium properties in these crystalline phases. The cell dynamics at a constant pressure were controlled \textit{via} Parrinello-Rahman method.\(^{13,15}\) The ionic temperatures were controlled with “massive” Nosé-Hoover chain thermostats, each with

\(^{*}\)Rings with O atoms as vertices and H bonds as edges.
Figure 5.2: (a) Probability distributions, $P(N)$, of the $N$-membered H-bonded rings contained in the PBE0+vdW$_{SC}$ structures of Ice Ih, Ice II, and Ice III. All $P(N)$ were normalized to unity and do not reflect the total number of rings of a given size. (b) Oxygen-oxygen radial distribution functions, $g_{OO}(r)$, obtained from PBE0+vdW$_{SC}$ simulations of these ice phases at the corresponding experimental triple point (238 K and 2.1 kbar).
a chain length of 4. The Car-Parrinello (CP) equations of motion for the nuclear and electronic degrees of freedom were integrated using the standard Verlet algorithm and a time step of 2.0 a.u. (∼ 0.05 fs). To ensure an adiabatic separation between the electronic and nuclear degrees of freedom in the CP dynamics, we used a fictitious electronic mass of 100 a.u. and the nuclear mass of deuterium for each H atom. Mass preconditioning was applied to all Fourier components of the electronic wavefunctions having a kinetic energy greater than 25 Ry. The interactions between the valence electrons and the ions (consisting of the nuclei and their corresponding frozen-core electrons) were treated with Hamann-Schlüter-Chiang-Vanderbilt type norm-conserving pseudopotentials distributed with the Qbox package. The electronic wavefunctions were expanded using a planewave basis set and a planewave kinetic energy cutoff of 130 Ry is used to converge stress tensors for a reliable cell dynamics. To enforce the evolution of the dynamics with a constant planewave kinetic energy (instead of constant number of planewaves) we have employed the method of Bernasconi et al. in conjunction with additional buffer planewaves corresponding to a planewave kinetic energy cutoff of 150 Ry. Given the large supercell, we sample the Brillouin zone only at the Γ point. For the EXX calculations we choose the converged cutoff parameters for water as described in Chapter 4, i.e., \( R_{\text{pair}} = 8.0 \text{ bohr}, \) \( R_{\text{PE}}^p = 6.0 \text{ bohr}, R_{\text{PE}}^b = 5.0 \text{ bohr}, R_{\text{ME}}^s = 10.0 \text{ bohr}, \) and \( R_{\text{ME}}^p = 7.0 \text{ bohr}. \)

---

†As proposed in reference, in order to maintain a constant kinetic energy of \( E_0 \), a smooth step function of height \( A \) and width \( \sigma \) is added to the kinetic factor \( G^2 \) as: \( G^2 \rightarrow G^2 + A \left[ 1 + \text{erf} \left( \frac{2G^2 - E_0}{\sigma} \right) \right] \). We have used \( A = 200 \text{ Ry}, \sigma = 15 \text{ Ry}, \) and \( E_0 = 130 \text{ Ry} \). In QE, the corresponding input parameters are: \( \text{qcutz}=200, \text{q2sigma}=15, \text{ecfixed}=130, \) and \( \text{ecutwfc}=150 \). Furthermore, we have used reference cell parameters which are at least 4.6% larger in each lattice direction than the corresponding equilibrium lattice parameters. The reference cell parameters were chosen to be large enough such that the fluctuations in the simulation cell parameters never exceeded the reference cell parameters.
5.2 Density and Enthalpy at the Triple Point

We begin to assess the performance of various theoretical methods by comparing the predicted equilibrium volumes of the ice phases with the corresponding experimental values. As shown in Table 5.1, the experimentally measured volumes of these ice phases decreases in the following sequence: ice Ih, III, and II; and all functionals employed here are able to reproduce this trend. However, the accuracy of PBE is rather unpredictable in determining the equilibrium absolute volumes ($V$) of these ices. For the low-density phase ice Ih PBE predicts only 1.4% smaller volume than experiment, however, for the high-density phase ice III the predicted volume is 7.4% larger than experiment. This larger overestimation of the volumes of the higher density phases is consistent with previous study, where vdW interactions were shown to become progressively more important in higher density phases.\cite{148,149} When vdW interactions are accounted for, we find that the equilibrium volumes of all ice phases decrease from PBE to PBE+vdW$_{\text{TS}}^{\text{SC}}$, which for ice Ih, II, and III are 4.1%, 11.7%, and 7.1%, respectively. This clearly indicates that vdW interactions have a larger influence on higher density ice phases. Unlike PBE, PBE+vdW$_{\text{TS}}^{\text{SC}}$ predicts the equilibrium volumes of these ice phases with more consistency by underestimating the experimental volumes in the range 3.1%–5.5%. On the contrary the inclusion of EXX effect weakens the H-bond strength\cite{143} and hence PBE0+vdW$_{\text{TS}}^{\text{SC}}$ effectively increases the volume in comparison to PBE+vdW$_{\text{TS}}^{\text{SC}}$ and comes in a better agreement with experiment. The influence of EXX effect is larger in ice Ih, which is strongly H bonded compared to higher density phases,\cite{143} and effectively increases the volume of ice Ih by 1.3%, followed by ice II (0.9%) and III (0.7%). As a result, the overall agreement with the experimental volumes is superior with PBE0+vdW$_{\text{TS}}^{\text{SC}}$ conceding errors ranging from 2.2%–4.2% compared to experiment. We should also note that NQEs are not accounted for in our simulations, which will affect the ice structures. In ice, NQEs tends to weaken the H-bond strength and results in an expansion of ...
the volume compared to the corresponding classical simulations, therefore, a better agreement with experiment is expected if NQEs is accounted for in conjunction with PBE0+vdW_{\text{SC}}^{\text{TS}} functional.
Table 5.1: Summary of the computed and experimental structural and energetic properties of ice phases at the experimental triple point (238 K and 2.1 kbar). From left to right: the specific method employed, equilibrium volume ($V$), difference in equilibrium volume ($\Delta V$), and difference in enthalpy ($\Delta H$). Relative percentage errors in the computed $V$ with respect to experiment are given in parentheses.

<table>
<thead>
<tr>
<th>Method</th>
<th>$V$ (Å³/H₂O)</th>
<th>$\Delta V$ (Å³/H₂O)</th>
<th>$\Delta H$ (meV/H₂O)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ih</td>
<td>III</td>
<td>Ih-III</td>
</tr>
<tr>
<td>PBE</td>
<td>31.26 (−1.4)</td>
<td>27.77 (+7.4)</td>
<td>26.21 (+4.0)</td>
</tr>
<tr>
<td>PBE+vdW_{TS}SC</td>
<td>29.98 (−5.5)</td>
<td>24.73 (−4.3)</td>
<td>24.42 (−3.1)</td>
</tr>
<tr>
<td>PBE0+vdW_{TS}SC</td>
<td>30.38 (−4.2)</td>
<td>24.93 (−3.6)</td>
<td>24.65 (−2.2)</td>
</tr>
<tr>
<td>Expt.</td>
<td>31.71</td>
<td>25.85</td>
<td>25.20</td>
</tr>
</tbody>
</table>

*From the equations of state provided in Ref.\textsuperscript{157}a Ref.\textsuperscript{158}b Ref.\textsuperscript{159}c
We have also computed oxygen-oxygen radial distribution functions (RDFs) from the ice structures obtained from PBE0+vdW_{TS} simulations. The details of the RDFs of ice phases at the experimental triple-point condition are rather scarce in the literature, hence the RDFs computed here with a highly accurate PBE0+vdW_{TS} potential will provide a reasonably accurate reference for the structures of these ice phases. As shown in Fig. 5.2, the positions of the first peak change marginally among these ice phases, which are located at 2.70 Å, 2.71 Å, and 2.73 Å in ice Ih, III, and II, respectively. However, the intensity of the first peak is significantly reduced going from low-density phase ice Ih (5.18) to higher density phases, ice III (4.05) and II (3.57). The reduction in the peak intensity is compensated with a broadening and the overall description of the first shell is reasonably similar in these ice phases. Whereas, the striking differences in the RDFs appear beyond the first shell in the 3–5 Å distance range. In the low-density ice Ih the second shell is well separated from the first shell and the second peak is situated at 4.41 Å. In ice III, the second shell is shifted towards much shorter distances having the second peak at 4.14 Å and the second shell becomes broader extending up to the first shell. Ice II being the most dense phase among the three phases shows the second peak at 3.60 Å and overlap between the second and first shell gets even stronger. The RDFs clearly demonstrate that the densification of ice III and II is a consequence of increased number density of the molecules in the distance range of 3–5 Å, essentially due to the collapse of the second shell.

Now moving on to the phase transition properties which are computed here by means of differences in the volume (∆V) and the enthalpy (∆H) among the ice phases. In Table 5.1, it can be seen that for both ∆V and ∆H, PBE0+vdW_{TS} provides the best agreement with experimental data. In comparison to experiment, PBE severely overestimates the ∆V for ice Ih–III and Ih–II transitions and underestimates the same for the III–II transition. With PBE+vdW_{TS} and PBE0+vdW_{TS} there are systematic improvements in the description of ∆V. With PBE0+vdW_{TS} the ∆V differs only by
0.4–0.8 Å³/H₂O compared to experiments. The experimental data shows that ice III has the lowest enthalpy among these three phases at this thermodynamic conditions; being only 1.8 meV/H₂O more stable than ice Ih. Obtaining such a small enthalpy difference is challenging and all of the functionals employed herein fail in this regard by incorrectly predicting ice Ih as the most stable phase. However, the ΔH of the Ih–III transition predicted with PBE0+vdW^{TS}_{SC} (3.1 meV/H₂O) is much closer to experiment than that predicted with PBE (46.7 meV/H₂O). Similarly, for the Ih–II and the III–II transitions, PBE overestimates ΔH severely, whereas PBE0+vdW^{TS}_{SC} predicts the ΔH within 3 meV/H₂O of the experimental data. Overall, we have shown that with PBE0+vdW^{TS}_{SC} the properties of ice phases at the experimentally known triple point can be described with high accuracy. The remaining deficiencies can possibly be circumvented by including NQEs and/or vdW interactions beyond pairwise approximations.

5.3 Conclusion

In this chapter, the exx module has allowed us to perform AIMD simulations of ice Ih, II, and III at the experimental triple point using state-of-the-art dispersion-inclusive hybrid DFT functional PBE0+vdW^{TS}_{SC}. The capability of the exx module on handling large-scale condensed-phase systems has enabled the use of large simulation cells that contains (H₂O)₉₆ in each phases. These large cells are ideal to capture the proto-disorder nature in these ice phases. We also provided a systematic comparison between PBE, PBE+vdW^{TS}_{SC}, and PBE0+vdW^{TS}_{SC} xc functionals on the equilibrium density and enthalpies; this comparison provides understanding on the individual and collective effects of dispersion and exact-exchange interactions in these ice phases at their triple point. Based on these simulations, we report the structural properties of these phases including the H-bond ring statistics and OO RDFs which bring us one
step closer to comprehensive understanding of these proto-disordered ice phases. It is worth noting that our simulations on the ice phases in this chapter assumes the nuclei as classical particles; however, the nuclear quantum fluctuations of both H and O atoms are known to have sizable influence\textsuperscript{[160]} on these systems. As such, we will apply the \texttt{exx} module to study these nuclear quantum effects in liquid water and ice Ih in Chapter 6.
Chapter 6

Probing Nuclear Quantum Effects in Liquid Water and Ice Ih

Understanding the structures of liquid water and ice Ih experimentally has been a long-standing challenge due to the presence of structural disorder (in liquid) and proton disorder (in ice Ih, $^{15}$) as mentioned in Chapter 5. Although pioneering works have provided useful insight into the radial distribution functions (RDFs) in liquid water$^{161-164}$ and structure factors in ice Ih$^{161}$, a comprehensive and atomistic understanding on the structures of these water systems is largely lacking to date. An alternative approach to bypass the experimental difficulties is via the use of theoretical simulations. Previous studies using theoretical simulations have identified that accurate determination of the structures of liquid water and ice Ih requires sophisticated treatment of the electronic xc interactions$^{8,9,48,148,149}$ and the quantum nature of the nuclei$^{165-168}$. Among the electronic xc effects, the H-bond network and the non-bonded dispersion interactions are particularly delicate and critical in accurate modeling of liquid water and ice Ih$^{8,148,148,149}$. Besides the requirement for a sophisticated electronic structure theory, the sampling of the proton and structural disorder as well as the nuclear quantum effects (NQEs) in water systems also stress
the need for extensive sampling in the configuration space with large length scale and long time scale. In view of both needs, the electronic structure method of choice is the dispersion-inclusive hybrid DFT—PBE0+vdW\textsubscript{TS}\textsubscript{SC}—that accounts for the H-bond network and the non-bonded dispersion interactions with good accuracy; the sampling method for the NQEs, on the other hand, is \textit{via} Feynman path-integral (PI) approach\textsuperscript{[169,174]} in conjunction with the AIMD method, \textit{i.e.}, PI-AIMD simulation, which provides a formally exact mapping for a quantum nuclei into a classical ring polymer. The use of PI-AIMD simulation at PBE0+vdW\textsubscript{TS}\textsubscript{SC} level of theory with large unit cells to accommodate proton/structural disorder creates a grand computational challenge—a challenge that is overcome in this chapter \textit{via} the efficiency and scalability of our \texttt{exx} module along with the power of massively parallel supercomputer architectures. In this chapter, I describe the computational details to deal with this challenge (Sec. 6.1), investigate the resulting structural (Sec. 6.2) and electronic (Sec. 6.3) properties of these water systems, and present a correlation between the orbital localization and the autoprotolysis-type configurations observed during these simulations (Sec. 6.4).

### 6.1 Methodology

Many techniques are necessary to overcome the computational challenge of performing PI-AIMD simulations at PBE0+vdW\textsubscript{TS}\textsubscript{SC} level with large-scale unit cell to accommodate structural and proton disorder under realistic \textit{NpT} conditions. In this section, we describe the methodology used in this study. The electronic structure calculations to evaluate the ionic and cell forces required in the propagation of the PI-AIMD simulations were computed using the \texttt{CP} module (along with our \texttt{exx} implementation) in the \texttt{QE} package. The core electrons were treated with the Hamann-Schl"uter-Chiang-Vanderbilt (HSCV) norm-conserving pseudopotentials\textsuperscript{[140,141]} distributed with \texttt{Qbox},\textsuperscript{[142]}
while the valence (pseudo-)wavefunctions were represented explicitly with a planewave basis with (kinetic energy) cutoff of 115 Ry. Within a constant-pressure simulation, the fluctuating cell incurs a spurious “Pulay-like” stress. To remove this stress, we followed Bernasconi et al.\textsuperscript{156} to modify the planewave kinetic energy as follows:

\[ G^2 \rightarrow G^2 + A \left[ 1 + \text{erf} \left( \frac{G^2 - E_0}{\sigma} \right) \right], \quad (6.1) \]

in which \( A = 200 \text{ Ry}, \, \sigma = 15 \text{ Ry}, \) and \( E_0 = 115 \text{ Ry}. \) In practice, we selected Miller indices within an (unmodified) kinetic energy cutoff of 125 Ry based on a reference cell that is sufficiently large to cover the fluctuations along each lattice vector during a given PI-AIMD simulation. With this modification in Eq. (7.1), a large kinetic energy penalty is smoothly added to planewaves with energy beyond the targeted 115 Ry cutoff. As such, the valence wavefunctions will have vanishing coefficients on these planewaves, which are effectively removed from the basis set. In doing so, the constant-pressure simulation mimics a constant-planewave-cutoff dynamics.

Within each PI-AIMD step, the electronic ground state wavefunctions were obtained using a second-order damped CP dynamics (fixing the ionic and cell degrees of freedom) with a fictitious electron mass \( \mu = 200 \text{ au} \) and time step \( \Delta \tau = 5 \text{ au}. \) In this case, we considered convergence to be reached via both an energy threshold of \( 1.7 \times 10^{-8} \text{ au}/\text{H}_2\text{O} \) and an ionic force threshold of \( 3.0 \times 10^{-4} \text{ au}. \)

The Quantum nuclei were modeled with 8-bead (Trotter dimension) supplemented with a colored-noise generalized Langevin equation thermostat—the PIGLET method.\textsuperscript{175,176} The cell is treated classically using the Raiteri-Gale-Bussi approach\textsuperscript{177} and is thermostatted using another generalized Langevin equation thermostat\textsuperscript{178} with a cell mass consistent with a 200 fs timescale using i-PI package.\textsuperscript{174}

For liquid water, we used a cubic simulation cell containing \((\text{H}_2\text{O})_{64}\) to account for the structural disorder. The simulation was performed under ambient thermody-
namically condition (300 K and 1.0 bar). For preliminary equilibration, we performed a 10-ps $NVT$ PI-AIMD simulation at the experimental density (1.00 g/cm$^3$) followed by a 11-ps $NpT$ PI-AIMD simulation. Then, an 8-ps $NpT$ production calculation was performed. For ice $Ih$, we chose an orthorhombic cell containing $(\text{H}_2\text{O})_{96}$ (same as what was used in Sec. 5.1) with a thermodynamic condition of (273 K and 1.0 bar). In this case, 0.5-ps $NVT$ and another 0.5-ps $NpT$ PI-AIMD simulation were run consecutively for equilibration; then, another 1.0-ps $NpT$ simulation was performed for production. The cell symmetries were fixed during both simulations, i.e., the cell in liquid water trajectory were kept cubic and that of the ice $Ih$ orthorhombic; as such, the $\text{exx}$ module can perform at its maximal efficiency for the evaluation of hybrid DFT related quantities—wavefunctions and cell forces—as discussed in Sec. 4.3.

### 6.2 Structural Properties

In this section, we analyzed the PI-AIMD trajectories obtained via the aforementioned methodology. Based on the $NpT$ sampling at PBE0+vdW$^{TS}_{SC}$ level, we found that the averaged density of liquid water (300 K, 1.0 bar) is 1.01 g/cm$^3$ and that of ice $Ih$ (273 K, 1.0 bar) is 0.97 g/cm$^3$. While the density of liquid water agrees quite well with the experimental value at 1.00 g/cm$^3$, the ice $Ih$ density is overestimated by 0.05 g/cm$^3$ compared to the experiment (at 0.92 g/cm$^3$). The 5% overestimation in ice $Ih$ density smaller but consistent with the previous 7% estimation, which shows the need for more sophisticated treatment of the electronic structures such as diffusion quantum Monte Carlo (DMC) method or coupled-cluster theory. Despite the remaining challenge to further refine the density for ice $Ih$, our PI-AIMD trajectories show promising improvements in the predicted structural properties such as radial distribution functions (RDFs) of the liquid water (Sec. 6.2.1) and structure factors of ice $Ih$ (Sec. 6.2.2).
6.2.1 Radial Distribution Functions of Liquid Water

Based on the \textit{NpT} PI-AIMD trajectory of liquid water at PBE0+vdW$^{TS}_{SC}$ level, we compute the radial distribution functions (RDFs), including \( g_{OO}(r) \) for O–O pairs, \( g_{OH}(r) \) for O–H pairs, and \( g_{HH}(r) \) for H–H pairs. To begin with, we compare the \( g_{OO}(r) \) obtained from our PI-AIMD simulation under \textit{NpT} condition with more commonly used \textit{NVT} method—AIMD simulation with classical nuclei in which NQEs are modeled by a 30 K temperature elevation. In this comparison (Fig. 6.1), our new

![PBE0+vdW](image)

Figure 6.1: Comparison of O–O RDFs obtained from simulations (solid lines) and X-ray scattering experiment (dashed lines) at 300 K and 1.0 bar. The simulations were performed at PBE0+vdW$^{TS}_{SC}$ level of theory under \textit{NVT} (green, AIMD simulation using classical nuclei along with a 30 K temperature elevation to mimic NQEs) and \textit{NpT} conditions (red, PI-AIMD simulation for explicit account for NQEs).

PI-AIMD simulation improves the height of the first peak compared to the experiment
and shows no appreciable change beyond the first peak compared with the commonly used $NVT$ approach. As such, this difference (by around 0.3) in the height of the first peak arises mainly from the differential convergence in the finite-size effects between $NpT$ and $NVT$ samplings. Specifically, when an $NVT$ simulation with large cell containing $(\text{H}_2\text{O})_{128}$ the height of the first peak reduces to slightly below 2.6, which is in quite good agreement with our $NpT$ simulation. The faster convergence in the finite-size effect in the $NpT$ simulation is due to the presence of cell fluctuation.

Despite the improvement on the $g_{\text{OO}}(r)$ in our PI-AIMD simulation compared to experiment, the result of PI-AIMD simulation still appears over structured. An interesting observation of this over structured picture in the PI-AIMD simulation is via directly comparing our simulated $g_{\text{OO}}(r)$ to the available experimental assignments at various temperatures. In this comparison, we found that our simulated $g_{\text{OO}}(r)$ aligns well with experiment at 268 K as shown in Fig. 6.2. In this case, we shifted $g_{\text{OO}}(r)$ of the PI-AIMD simulation by 0.06 Å to align the first peak—size of the first coordination shell. In doing so, we observed very good agreement between the PI-AIMD simulation and the experiment. This observation is indicative of that the melting point of PBE0+vdW\textsubscript{TS} water is slightly above 300 K.

The hypothesis that liquid water using PBE0+vdW\textsubscript{TS} level of theory is under supercooled condition at (300 K, 1.0 bar) is further supported by comparing mean kinetic energy of the H nuclei between the simulation and a deep inelastic neutron scattering (DINS) measurement as shown in Table 6.1. In this comparison, the quantum kinetic energy of H agrees well in ice Ih case (within 4 meV/atom). In the liquid water comparison, the PI-AIMD result differs significantly from the experimental measurement of ambient liquid water by 12 meV/atom but agrees with supercooled water at 271 K instead (within 2 meV/atom). Both comparisons—via $g_{\text{OO}}(r)$ and $\langle E_k \rangle$—suggest that liquid water using PBE0+vdW\textsubscript{TS} level of theory is under supercooled condition at (300 K, 1.0 bar) and the undercooling in this case is
Figure 6.2: Comparison of O–O RDFs obtained from PI-AIMD simulation (red line) to experiment (black dashed line) performed under supercooled condition.

Table 6.1: Comparison of computed (PI-AIMD) and experimental quantum kinetic energies in liquid water and ice Ih. All experimental values are obtained from Ref.113

<table>
<thead>
<tr>
<th>Phase</th>
<th>Method</th>
<th>$T$ (K)</th>
<th>$p$ (bar)</th>
<th>$\langle E_k \rangle$ (meV/atom) $^H$</th>
<th>$\langle E_k \rangle$ (meV/atom) $^O$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ice Ih</td>
<td>PI-AIMD</td>
<td>273</td>
<td>1.0</td>
<td>161</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>Expt.</td>
<td>271</td>
<td>1.0</td>
<td>157 ± 2</td>
<td>N.A.</td>
</tr>
<tr>
<td>Liquid</td>
<td>PI-AIMD</td>
<td>300</td>
<td>1.0</td>
<td>158</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>Expt.</td>
<td>300</td>
<td>1.0</td>
<td>146 ± 3</td>
<td>N.A.</td>
</tr>
<tr>
<td></td>
<td>Expt.$^a$</td>
<td>271</td>
<td>1.0</td>
<td>156 ± 2</td>
<td>N.A.</td>
</tr>
</tbody>
</table>

$^a$ Supercooled liquid.

Such undercooling at (300 K, 1.0 bar) is likely due to the imperfection of the PBE0+vdW$^{TS\ SC}$ to approximate the exact xc functional; however, the actual melting point of water at PBE0+vdW$^{TS\ SC}$ level requires free energy calculations on the solid and liquid phases.
the $\approx 30$ K undercooling is rather small in terms of energy which manifests that PBE0+vdW$_{TS}^{SC}$ provides a highly accurate description for liquid water systems.

While the commonly used approach—AIMD with classical nuclei and 30 K temperature elevation to mimic the NQEs in liquid water—works well for O–O correlations, it is inappropriate to model intramolecular NQEs (e.g., in the O–H RDF). In Fig. 6.3, we compare the $g_{\text{OH}}(r)$ from previous classical AIMD simulation, our PI-AIMD simulation, and the experiment. In this case, the 30 K elevation largely underestimates the intramolecular NQEs and gives rise to sharp peaks (and wells) in $g_{\text{OH}}(r)$. On the contrary, these quantum fluctuations are properly captured in the PI-AIMD simulation. As such, the resulting $g_{\text{OH}}(r)$ agrees quite well with the experimental assignment. Again, the PI-AIMD result is slightly over structured due to the underlying xc functional approximation. Despite the limitation from the xc functional in our simulation, there remains room to improve the experimental assignment. For example, the first peak of in the experimentally assigned $g_{\text{OH}}(r)$ has a nearly symmetric Gaussian shape. This Gaussian shape is due to the use of a harmonic intramolecular empirical potential for structural refinement during the experimental assignment. In this regard, our PI-AIMD simulation (despite the imperfection of the xc functional) provides even
a more realistic asymmetric—anharmonic—picture for the first peak in \( g_{\text{OH}}(r) \), which could aid the design of the next generation empirical potential used in the refinement of the experimental assignment.

As in case of \( g_{\text{OH}}(r) \), AIMD simulation with classical nuclei and 30 K elevation in temperature is also inappropriate to sample the \( g_{\text{HH}}(r) \) due to large intramolecular NQEs in liquid water (Fig. 6.4). In this case, our PI-AIMD simulation also agrees quite well to a recent joint X-ray/neutron scattering experiment,\(^{180}\) in which intermolecular \( g_{\text{HH}}(r) \) is reported. We also plot the earlier assignment\(^{161}\) which contains the intramolecular information; however, these two experimental assignments have large deviation among themselves. In this case, our PI-AIMD simulation provides a quite reasonable \( g_{\text{HH}}(r) \) that falls within the deviation of the two experimental as-

![Figure 6.4: Comparison of H–H RDFs from classical AIMD and PI-AIMD simulations to experiment.](image)

\[ g_{\text{HH}}(r) \]

\[ r \text{ [Å]} \]
signments and agrees quite well with the with the Soper 2013 result along with the first (intramolecular) peak that is inaccessible in the newer experimental assignment.

6.2.2 Structure Factors of Ice $I_h$

Compared with in liquid water, the crystalline symmetry in ice $I_h$ makes experimental assignment of RDFs more challenging and has quite significant uncertainty. As such, instead of comparing the RDFs, we compute from our PI-AIMD trajectory the partial structure factors, which can be probed directly from experiment. The result of the comparisons on the partial structure factors are shown in Fig. 6.5. The quite significant deviation in the partial structure factors between the experiment and the EPSR fit indicates that the ensemble generated using EPSR method is unrepresentative to the experimental observation. This difficulty arises from the crystalline symmetry in ice (in contrast with the isotropic symmetry in liquid water). On the other hand, the PI-AIMD simulation provides a more realistic description of the intramolecular and intermolecular interactions compared with the empirical potential used in the EPSR fit. As such, the resulting partial structure factors obtained from PI-AIMD simulation aligns well (despite the 53 K difference in temperature) with the experiment. We expect more accurate partial structure factors can be obtained from a PI-AIMD simulation at PBE0+vdW$_{SC}^TS$ level under 220 K. Taken together the density, RDFs, and partial structure factors obtained using the exx module at PBE0+vdW$_{SC}^TS$ level of theory, we have obtained a highly accurate ensemble with realistic system size and reasonable simulation time, which can be further analyzed to provide more insight into the structural properties of liquid water and ice $I_h$.

This could be done at the cost of an increased Trotter dimensions
Figure 6.5: Comparison of O–O, O–H, and H–H partial structure factors in ice I_h between simulated (red), empirical potential structure refinement (EPSR, blue) fit, and experiment (black). All partial structure factors uses the Faber-Ziman (FZ) definition. The experiment and the EPSR were performed under 220 K and our PI-AIMD simulation was under 273 K.
6.3 Electronic Properties

The PI-AIMD approach also gained us access to the electronic properties of liquid water and ice Ih. As the exx module requires the MLWFs, we have recorded the MLWF centers. The MLWF centers can provide useful information such as the molecular dipole moment. In this section, we use the MLWFs required by the exx module to perform analysis on the average strength of H bonds and the molecular dipole moment in liquid water as demonstrated in Sec. 6.3.1. In addition to the MLWFs, we also investigate the canonical KS orbitals (at GGA level) to understand how the NQEs affect the edge of the valence band maximum in Sec. 6.3.2.

6.3.1 Wannier Functions in Liquid Water

With the MLWFs in hand, the NQEs on the strength of the H bond in liquid water can be understood using the distribution function of the distance between O and its associated MLWF centers as shown in Fig. 6.6. In this analysis, two peaks are observed, which correspond to the bond pair (more distant from O) and lone pair (closer to O) electrons. The distance between the lone pair from the O atom can be used to qualitatively analyze the strength of the H bond—a stronger H bond pulls the lone pair away from the O atom. In Fig. 6.6, the lone pair is closer to the O atom in the PI-AIMD result compared to the AIMD case, i.e., the H bond strength is slightly weaker at the presence of the NQEs, which also contributes to the less structured first peak in $g_{OO}(r)$ (Fig. 6.1). In addition, the larger spread of the peaks also demonstrate larger variation in the H-bond environment. Both of these observations are consistent with the fact that significant quantum fluctuation is present on the H atoms when PI-AIMD simulation is performed.

On the other hand, the MLWF centers also provides the possibility of performing molecular dipole analysis. In Fig. 6.7, we analysis the molecular dipoles of each water
molecule along the AIMD and the PI-AIMD trajectories. We found that the NQEs in PI-AIMD simulation results in a slightly higher the dipole moment compared with the classical AIMD case—both in the mean value and in the spread.

6.3.2 Valence Band Maximum

In addition to the MLWF representation, we also investigate the eigenvalues of the canonical KS orbitals at the PBE level\footnote{In this case, the dispersion correction leads to essentially a constant shift to the orbital energy.}. In practice, we take a snapshot every 100
MD steps to compute the eigenvalues and the density of states (DOS) for liquid water (Fig. 6.8) and ice Ih (Fig. 6.9). For both systems, we found that NQEs significantly broadens the highest valence band and leads to a long tail toward the conduction band direction (which is $\approx 0.5$–1.0 eV). On the other hand, the conduction band is less affected by these fluctuations. As such, the overall band gap is reduced by around 1 eV due to the quantum fluctuations on top of the classical thermal fluctuations.\footnote{In this calculation, $\Gamma$ point is used to sample the Brillouin zone, which is appropriate for valence bands but requires further discretization when it comes to conduction bands. In this case, we will focus on the information of the valence bands; further refinement of the conduction bands will left for future study.}

The elongated tail in the valence band (in the PI-AIMD simulation compared with the AIMD simulation) originates from the large distortion promoted by the quantum

---

**Figure 6.7**: Molecular dipole distribution in the PI-AIMD trajectory of liquid water at PBE0+vdW$^{TS}_{SC}$ level.
fluctuations (see Sec. 6.4). At the presence of such fluctuations, the KS orbitals becomes localized on the most distorted region. To identify such localized KS orbitals, we used the inverse participation ratio (IPR)\textsuperscript{181} defined as (within a volume $V$):

$$\text{IPR}[\phi_i(r)] = \frac{\int d\mathbf{r} |\phi_i(r)|^2}{V \int d\mathbf{r} |\phi_i(r)|^4},$$  \hspace{1cm} (6.2)
Figure 6.9: Density of state computed at the PBE level of theory using a subset of the PI-AIMD ice Ih trajectory sampled at the PBE0+vdW\textsuperscript{TS}\textsubscript{SC} level.

in which $\phi_i(r)$ is the $i$-th canonical KS orbital with eigenvalue $\varepsilon_i$. The definition of Eq. (6.2) is a functional of the KS orbital that measures its extent of delocalization. For example, for a completely delocalized orbital,

$$\phi(r) = \frac{1}{\sqrt{V}},$$  \hspace{1cm}  (6.3)
the resulting IPR value is \( V \) which is the volume itself. On the contrary, for a completely localized limit, \( i.e., \) a delta function:

\[
\phi(r) = \delta(r);
\]

its IPR takes on value \( 1/V \), which vanishes at the limit of the large \( V \). In other words, the IPR value provides a measure for the volume of the support on a given orbital and, hence, is a convenient quantity to characterize orbital localization of a given energy level \( \varepsilon_i \) (in this case canonical KS orbitals have well-define eigenvalues unlike the MLWFs).

Based on the IPR analysis, we perform an analysis on KS orbital localization due to distortion on top of the DOS (also shown in Fig. 6.8 for liquid water and Fig. 6.9 for ice \( \text{I}_h \). In this IPR analysis, we use an energy window of 0.2 eV to obtain blocked mean and standard deviation to simplify the picture. We found that the IPR values is around an order of magnitude smaller near the edge of the valence-band maximum in both liquid water and ice \( \text{I}_h \) cases. Such localization is accessible with both classical thermal and quantum fluctuations as shown in the right panels of Figs. 6.8 and 6.9.

### 6.4 Autoprotolysis

To gain further understanding on the localization, we investigate this effect by visualizing the orbitals of two liquid water snapshot—in this case, one has a highest occupied orbital (HOMO) at the valence-band edge (near \( -0.15 \) eV in Fig. 6.10) while the snapshot with HOMO at the peak of the DOS (near \( -0.35 \) eV in Fig. 6.10). In this case, the orbital with IPR value around \( 1 \times 10^{-2} \) is delocalized over multiple water molecules (with lone-pair characteristics as shown in the lower left panel of Fig. 6.10), and is the most frequently observed characteristics among the available snapshots.
Figure 6.10: Localization of canonical KS orbitals near the valence band maximum in the PI-AIMD trajectory of liquid water at 300 K: (left panel) near the peak of the density of state; (right panel) near the upper edge.

For the other case (a rare event throughout the PI-AIMD or AIMD simulation), the IPR value is around $1 \times 10^{-3}$, in which the orbitals at the valence-band edge is
almost completely localized on a single water molecule as shown in the lower right panel of Fig. 6.10. In this panel, such localization corresponds to a proton-transfer-type distortion in which an OH bond (of the center water molecule) is elongated (from its typical length 1.0 Å to 1.3 Å) along an H-bond direction (to the lower left water molecule) such that the H atom is almost at the midpoint of the two O atoms. If the process continues, it will ultimately lead to autoprotolysis where the center molecule transforms to an OH\(^-\) and the lower left molecule to H\(_3\)O\(^+\). In this autoprotolysis limit, the localized HOMO on OH\(^-\) becomes one of the lone-pair orbitals that bares the negative charge.

Interestingly, such autoprotolysis-type configuration is also observed in case of ice Ih at the valence-band edge. Using a similar analysis as in the liquid water case, we visualizing the orbital at the valence-band edge as shown in Fig. 6.11. As in liquid water, our ice Ih simulations also visited the autoprotolysis-type configurations with localized HOMOs whose energies locate at the edge of the valence band. The observation in ice Ih even has a higher extent of proton transfer near the autoprotolysis limit. In the lower panel of Fig. 6.11, the water molecule with largest lobe of the HOMO has a very extended OH bond (1.5 Å that could already be assigned as an H-bond); this H atom now is closer (with a distance of 1.2 Å) to the upper right water molecule (which then becomes H\(_3\)O\(^+\). The larger extent of the autoprotolysis-type configuration in ice Ih is likely due to the presence of larger nuclear quantum fluctuation (as the temperature is lower than in case of the liquid water simulation) and the almost perfect H-bond structure in ice Ih (while water has broken H-bonds). The larger tendency for proton transfer in ice Ih than in liquid water suggests that ice Ih might have a larger proton conductance than liquid water, which agrees the experimental observation. In this experiment, the proton diffusion constant (for \(T > 235\) K) is \(\sim 10^{-3}\) cm\(^2\)/s, which is around 10 times larger than that in water at 295 K, and the activation energy is around \(\sim 0.01\) eV.
In this chapter, we applied the exx module developed in the first part of this thesis in conjunction with supercomputer architectures to perform PI-AIMD simulations at PBE0+vdW\textsuperscript{TS} level under realistic \emph{NpT} conditions in liquid water and ice Ih. These simulations simultaneously account for sophisticated electronic xc interactions, nu-

---

Figure 6.11: Localization of canonical KS orbitals near the valence band maximum in the PI-AIMD trajectory of ice Ih at 273 K: (left panel) near the peak of the density of state; (right panel) near the upper edge.

6.5 Conclusion and Future Outlook

In this chapter, we applied the exx module developed in the first part of this thesis in conjunction with supercomputer architectures to perform PI-AIMD simulations at PBE0+vdW\textsuperscript{TS} level under realistic \emph{NpT} conditions in liquid water and ice Ih. These simulations simultaneously account for sophisticated electronic xc interactions, nu-
clear quantum fluctuations, and experimental thermodynamic conditions. Based on these simulations, we found that the density of liquid water is reproduced quite well at 1.01 g/cm$^3$ (+1% error) while the ice Ih density is less accurate at 0.97 g/cm$^3$ (+5% error). Despite the remaining challenge in predicting the ice density, our PI-AIMD trajectories predict the experimental RDFs in liquid water and partial structure factors in ice Ih with good accuracy. We also identify that the ensemble generated by our PBE0+vdW$^{TS}_{SC}$ PI-AIMD liquid water simulation (performed at 300 K) resembles a supercooled water at 268 K in experiment in terms of available O–O RDF and the quantum kinetic energy of H atom measured in DINS experiment. This observation suggests that PBE0+vdW$^{TS}_{SC}$ water is supercooled by of the order of 30 K at (300 K, 1.0 bar). In addition, our PI-AIMD simulation also provides refined structural properties of liquid water such as O–H and H–H RDFs, in which experimental values are limited by the EPSR fit based on empirical potentials. Besides the structural properties, we also found that the presence of nuclear quantum fluctuations slightly reduces the H-bond strength in liquid water and increases the strength of molecular dipole moment as well as its spread. These nuclear quantum fluctuations also result in an elongated exponential tail (compared with the classical fluctuations) in the DOS in both liquid water and ice Ih, which reduces the band gap by around 1 eV. At the edge of the valence band maximum, large distortion gives rise to localization of KS orbitals which is almost completely localized on a single water molecule (in the lone-pair-type molecular orbital). This type of localization can be easily measured using IPR, which provides a measure of volume of support of a given KS orbital. By visualizing such localized KS orbitals, we found that the distortion associated with such orbital localization corresponds to the autoprotolysis process. These configurations can be promoted both by classical thermal and quantum fluctuations. We also found that the ice Ih has a larger tendency to form such autoprotolysis-type configurations, which can be a possible cause of the larger proton mobility in ice Ih than in liquid
water. Last but not least, our PI-AIMD simulation also provide an accurate data set for training machine-learning based force fields which will mitigate the prohibitive computational cost for the application to large length and time scale.
Chapter 7

Application to Organic Molecular Crystals

Organic molecular crystals are versatile materials with widespread use across many fields, e.g., pharmaceuticals, explosives, and nonlinear optics. These properties—e.g., biological activity of a drug, energy density of an explosive, and optical response of a nonlinear medium—are all governed by their underlying molecular crystal structures as well as their (often numerous) polymorphs. These examples stress the need for a reliable theoretical methods for crystal structure prediction (CSP)—both to provide key physical insight into such structure-property relationships and to offer the promise of rational design of molecular crystals with targeted properties. In this chapter, we study the prediction of structural and thermal properties of dispersion-bound molecular crystals in constrast with H-bond molecular crystals such as ice phases studied in Chapter 5. Such prediction of structural and thermal properties of dispersion-bound molecular crystals relies on a complex interplay between many effects. These effects include anharmonicity, nuclear quantum effects (NQEs), many-body dispersion interactions, and Pauli

---

repulsion. The roles of these effects are analyzed by successively comparing structural properties due to (i) classical thermal (Sec. 7.3) and quantum (Sec. 7.4) fluctuations, (ii) van der Waals (vdW) corrections at effective-pairwise and sophisticated many-body dispersion (Sec. 7.5.1) levels, and (iii) density functional theory (DFT) at generalized-gradient approximations (GGA) and hybrid level (Sec. 7.5.2) using the exx module developed in the first part of this thesis.

7.1 Molecular Crystals under Finite Temperatures

Despite that all applications of molecular crystals occur at finite temperatures ($T$) and pressures ($p$), most CSP methods focus on determining structural properties (e.g., lattice parameters and cell volumes) at 0 K. While such athermal prediction is suitable for many covalent and ionic solids, this approach is unlikely to provide accurate structural properties for non-covalently bound systems such as molecular crystals. The failure of this approach for molecular crystals is due to the often large thermal expansivities originating from relatively weak and highly anharmonic intermolecular forces. This phenomenon can be seen in a benzene crystal—a prototypical molecular crystal—which expands by 2.7% from 4 K–138 K\textsuperscript{190,191} while such thermal expansion is at least one order of magnitude smaller for similar temperatures for covalent solid such as Si\textsuperscript{192}.

As a first step, we investigate the influence of anharmonicity on the structural properties over a set of pyridine-like molecular crystals (PLMCs) in Sec. 7.3. This set comprises the following N-heterocyclic aromatic compounds: pyridine (with two polymorphs),\textsuperscript{193,194} pyrrole,\textsuperscript{195} pyridazine (two different thermodynamic conditions),\textsuperscript{196} and bipyridine.\textsuperscript{197} These molecules are pervasive throughout chemistry, biology, and
agriculture that play relevant roles as ligands, solvents, pharmacophores, and herbicide precursors.\textsuperscript{178}

### 7.2 Simulation Methodology

To study how finite $T$ and $p$ influence structural properties in molecular crystals, we utilize \textit{ab initio} molecular dynamics (AIMD)\textsuperscript{177} simulations in the isobaric-isothermal ($NpT$) ensemble. With a quite favorable ratio of cost to accuracy, DFT\textsuperscript{175,179} based on GGA is often used to treat the electrons and has become the \textit{de facto} standard in \textit{ab initio} simulations of condensed-phase systems. Despite the widespread success, semi-local functionals face two significant shortcomings—presence of self-interaction error (SIE) and lack of vdW interactions—as described in Sec. 1.3. In this case, the ability to account for the vdW interactions is crucial for even qualitatively describing non-covalently bound molecular crystals.\textsuperscript{179} As such, we employed the Perdew-Burke-Ernzerhof (PBE) GGA-based exchange-correlation (xc) functional\textsuperscript{28} in conjunction with a fully self-consistent implementation\textsuperscript{8,36} of the Tkatchenko-Scheffler (TS) dispersion correction,\textsuperscript{34} denoted by PBE+vdW\textsuperscript{TS}SC herein.

The vdW\textsuperscript{TS} method is an effective-pairwise ($C_6/R^6$) approach in which all atomic parameters (\textit{e.g.}, dipole polarizabilities, vdW radii, and dispersion coefficients) are functionals of the electron density. Hence, this model accounts for the local chemical environment of each atom and yields interatomic $C_6$ coefficients accurate to $\approx 5\%$.\textsuperscript{31,34} When compared with low-$T$ experiments, variable-cell (VC) optimizations with PBE+vdW\textsuperscript{TS} typically predict lattice parameters to $\approx 2\%$ in crystals containing small organic molecules like ammonia, benzene, urea, and naphthalene.\textsuperscript{200,201,202} In the self-consistent implementation, such non-local correlation effects are accounted for in the charge density through the dispersion contribution to the xc potential. Such self consistency in the evaluation of the PBE+vdW\textsuperscript{TS}SC energy and forces ensures appro-
 appropriated energy conservation during AIMD and can also sizably affect binding energies in highly polarizable molecules and materials as well as work functions in coinage metals. As the aim of this chapter is to quantify the influence of anharmonicity on the PLMC cell parameters under experimental conditions \( (T_{\text{expt}}, p_{\text{expt}}) \), we compare results of VC optimizations at \( (0 \text{ K}, p_{\text{expt}}) \) and \( NpT \)-based AIMD simulations at \( (T_{\text{expt}}, p_{\text{expt}}) \).

In this chapter, all AIMD simulations were performed in the isobaric-isothermal \((NpT)\) ensemble using the Car-Parrinello (CP) scheme and the planewave-pseudopotential Quantum ESPRESSO (QE) package. Core electrons were treated with the Hamann-Schlüter-Chiang-Vanderbilt (HSCV) norm-conserving pseudopotentials distributed with Qbox, while the valence electronic wavefunctions were computed explicitly with a planewave (kinetic energy) cutoff of 110 Ry. In the CPMD equations of motion (see Eqs. (2.6)-(2.8)), the valence electronic wavefunctions were assigned a fictitious mass of \( \mu = 300 \text{ au} \), with mass preconditioning on all Fourier components that have a kinetic energy greater than 3 Ry. Following Ref., we remove the “Pulay-like” stress resulting from cell fluctuations by modifying the planewave kinetic energy as:

\[
G^2 \rightarrow G^2 + A \left[ 1 + \text{erf} \left( \frac{G^2 - E_0}{\sigma} \right) \right],
\]

in which \( A = 200 \text{ Ry} \), \( \sigma = 15 \text{ Ry} \), and \( E_0 = 110 \text{ Ry} \). In practice, this setting includes Miller indices with a kinetic energy up to 130 Ry when using a reference cell that is big enough to contain the largest cell fluctuations during the simulation. With this modification, all \( NpT \)-based simulations mimic the dynamics at a constant planewave (kinetic energy) cutoff of 110 Ry.
Since the nuclei are treated classically in the AIMD framework, sampling of structural properties in the \( NpT \) ensemble is independent of isotope substitution. Hence, we replace each H atom in the system with D to allow for an integration time step of \( \Delta t = 4.0 \text{ au} \approx 0.1 \text{ fs} \). To maintain the temperature of the system (at the experimental value of \( T_{\text{expt}} \)), we attach massive Nosé-Hoover thermostat chains to each ionic degree of freedom with a chain length of 4 and a frequency of 60 THz. To maintain the pressure of the system (at the experimental value of \( p_{\text{expt}} \)), we utilized the Parrinello-Rahman barostat. All AIMD simulations were started from the experimental structure and equilibrated for 1 ps. All statistical quantities reported in this work were based on AIMD trajectories that were at least 10 ps in duration.

### 7.3 Classical Anharmonicity

Fig. 7.1 compares the predicted volumes from VC optimizations and AIMD simulations with experiment, clearly demonstrating that anharmonicity effects are indeed non-negligible in determining this structural property. VC optimizations always underestimate this quantity and the inclusion of anharmonicity via \( NpT \)-based AIMD systematically reduces the mean absolute error (MAE) from 4.7% to 1.2% in the predicted volumes. In fact, the influence of anharmonicity can be quite substantial in the PLMC set, as evidenced by the 6.4% change in \( V_{\text{err}} \) for pyridazine at \((295 \text{ K}, 2.7 \text{ kBar})\). We note that the extent to which anharmonicity will influence cell volume expansion depends on a complex interplay between \( p_{\text{expt}} \) and the cohesive forces at work in the crystal (which act together to suppress expansion) and \( T_{\text{expt}} \) (which provides thermal energy to explore the nuclear potential energy surface).

---

\(^1\)Nuclei were commonly assumed as classical particles for the study of dispersion-bound molecular crystal—as the molecules are heavy and bound by weak dispersion interactions. However, this research will show that this common assumption is not valid in general.
Figure 7.1: Predicted cell volumes ($V_{\text{pred}}$) from VC optimizations and AIMD simulations using PBE+vdW$_{\text{SC}}$ for the PLMC set. Errors are defined with respect to experiment ($V_{\text{expt}}$) at the indicated thermodynamic conditions as $V_{\text{err}} = (V_{\text{pred}} - V_{\text{expt}})/V_{\text{expt}}$. Inset: Overlay of predicted (blue) and experimental (red) pyridine-I structures.
<table>
<thead>
<tr>
<th>System</th>
<th>Method</th>
<th>Supercell</th>
<th>$a$ (Å)</th>
<th>$b$ (Å)</th>
<th>$c$ (Å)</th>
<th>$\alpha$ (°)</th>
<th>$\beta$ (°)</th>
<th>$\gamma$ (°)</th>
<th>$V$ (Å³)</th>
<th>$\Delta V_{err}$ (%)</th>
<th>RMSD (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pyridine-I (16 molec)</td>
<td>VC (0 K, 1 Bar)</td>
<td>$1 \times 1 \times 1$</td>
<td>17.25</td>
<td>8.88</td>
<td>11.14</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>1712</td>
<td>−4.0</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>AIMD (153 K, 1 Bar)</td>
<td>$1 \times 1 \times 1$</td>
<td>17.43(3)</td>
<td>8.92(2)</td>
<td>11.31(5)</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>1767(2)</td>
<td>−0.9(1)</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>PI-AIMD (153 K, 1 Bar)</td>
<td>$1 \times 1 \times 1$</td>
<td>17.51(4)</td>
<td>8.95(3)</td>
<td>11.44(6)</td>
<td>90.0(1)</td>
<td>89.5(3)</td>
<td>90.02(7)</td>
<td>1789(2)</td>
<td>+0.3(1)</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>Expt</td>
<td>$1 \times 1 \times 1$</td>
<td>17.52</td>
<td>8.97</td>
<td>11.35</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>1784</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Pyridine-II (4 molec)</td>
<td>VC (0 K, 11 kBar)</td>
<td>$2 \times 2 \times 1$</td>
<td>5.33</td>
<td>6.56</td>
<td>11.30</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>386.0</td>
<td>−4.3</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>AIMD (298 K, 11 kBar)</td>
<td>$2 \times 2 \times 1$</td>
<td>5.46(1)</td>
<td>6.72(4)</td>
<td>11.23(5)</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>412.8(5)</td>
<td>−0.4(1)</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>Expt</td>
<td>$2 \times 2 \times 1$</td>
<td>5.40</td>
<td>6.80</td>
<td>11.23</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>414.0</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Pyrrole (4 molec)</td>
<td>VC (0 K, 1 Bar)</td>
<td>$2 \times 1 \times 3$</td>
<td>7.23</td>
<td>10.10</td>
<td>4.96</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>361.9</td>
<td>−4.8</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>AIMD (103 K, 1 Bar)</td>
<td>$2 \times 1 \times 3$</td>
<td>7.35(2)</td>
<td>10.19(1)</td>
<td>4.99(1)</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>372.8(3)</td>
<td>−1.99(7)</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>Expt</td>
<td>$2 \times 1 \times 3$</td>
<td>7.29</td>
<td>10.29</td>
<td>5.07</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>380.2</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>Pyridazine (4 molec)</td>
<td>VC (0 K, 2.7 kBar)</td>
<td>$3 \times 1 \times 1$</td>
<td>3.675</td>
<td>10.62</td>
<td>9.70</td>
<td>90.1</td>
<td>91.4</td>
<td>90.0</td>
<td>378.4</td>
<td>−8.2</td>
</tr>
<tr>
<td></td>
<td>AIMD (295 K, 2.7 kBar)</td>
<td>$3 \times 1 \times 1$</td>
<td>3.809(5)</td>
<td>10.86(2)</td>
<td>9.83(1)</td>
<td>90.0(1)</td>
<td>90.6(3)</td>
<td>90.1(2)</td>
<td>405.1(5)</td>
<td>−1.7(1)</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>Expt</td>
<td>$3 \times 1 \times 1$</td>
<td>3.843</td>
<td>10.96</td>
<td>9.78</td>
<td>90.0</td>
<td>91.1</td>
<td>90.0</td>
<td>412.0</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>Pyridazine (4 molec)</td>
<td>VC (0 K, 6.1 kBar)</td>
<td>$3 \times 1 \times 1$</td>
<td>3.643</td>
<td>10.53</td>
<td>9.68</td>
<td>90.2</td>
<td>90.8</td>
<td>90.0</td>
<td>370.4</td>
<td>−4.3</td>
</tr>
<tr>
<td></td>
<td>AIMD (295 K, 6.1 kBar)</td>
<td>$3 \times 1 \times 1$</td>
<td>3.735(6)</td>
<td>10.76(3)</td>
<td>9.79(2)</td>
<td>89.9(2)</td>
<td>90.5(4)</td>
<td>89.9(2)</td>
<td>391.9(5)</td>
<td>1.3(1)</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>Expt</td>
<td>$3 \times 1 \times 1$</td>
<td>3.719</td>
<td>10.75</td>
<td>9.68</td>
<td>90.0</td>
<td>91.5</td>
<td>90.0</td>
<td>386.9</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Bipyridine</td>
<td>VC (0 K, 1 Bar)</td>
<td>$2 \times 2 \times 1$</td>
<td>5.529</td>
<td>5.98</td>
<td>11.58</td>
<td>90.00</td>
<td>96.46</td>
<td>90.00</td>
<td>380.7</td>
<td>−2.6</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>AIMD (123 K, 1 Bar)</td>
<td>$2 \times 2 \times 1$</td>
<td>5.576(6)</td>
<td>5.99(1)</td>
<td>11.67(2)</td>
<td>90.01(5)</td>
<td>95.91(9)</td>
<td>90.07(9)</td>
<td>387.6(4)</td>
<td>−0.9(1)</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>Expt</td>
<td>$2 \times 2 \times 1$</td>
<td>5.486</td>
<td>6.17</td>
<td>11.61</td>
<td>90.00</td>
<td>95.28</td>
<td>90.00</td>
<td>391.0</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>

Table 7.1: Predicted and experimental structural properties for the PLMC set. All simulations were performed using PBE+vdWTS and the numbers in parentheses denote uncertainties in the predicted values. Orthorhombic symmetry was enforced throughout the VC optimizations and AIMD simulations on pyridine-I, pyridine-II, and pyrrole. For all other molecular crystals (as well as the PI-AIMD simulation of pyridine-I) the full cell tensors were allowed to fluctuate. The number of molecules per unit cell and the chosen simulation supercell sizes are also listed along with the RMSD of the atomic positions with respect to experiment.
AIMD simulations also yield PLMC lattice parameters that agree remarkably well with experiment (Table 7.1). By accounting for anharmonicity, AIMD systematically reduce the MAE in the predicted lattice parameters from 2.0% to 1.3% with respect to experiment. As seen above, VC optimizations tend to underestimate PLMC lattice parameters; however, this trend does not always hold as evidenced by the slight negative linear thermal expansion observed along the c axis in pyridine-II. This predicted effect is consistent with the experimental data\textsuperscript{194} and reproduces the reference lattice parameter with extremely high fidelity. By considering the lattice parameter fluctuations throughout the AIMD trajectory, we found that the c axis was not the softest (most flexible) dimension in pyridine-II, hence the apparent negative linear thermal expansion in this molecular crystal has a distinctly different origin than that of methanol monohydrate.\textsuperscript{203} Since this effect is also observed during GGA-based AIMD (which do not account for dispersion interactions), this phenomenon is most likely electrostatic in nature for pyridine-II. In addition, the structure and orientation of the individual molecules inside the PLMC unit cells are also well described by AIMD with PBE+vdW\textsuperscript{TS SC} (Fig. 7.1 and Fig. 7.2), with associated root-mean-square deviations (RMSD) of 0.17 Å across this set of dispersion-bound molecular crystals.

To demonstrate the accuracy of the NpT-based AIMD simulations at the PBE+vdW\textsuperscript{TS SC} level for structural properties, we overlay the predicted and experimental equilibrium crystal structures for a set of pyridine-like molecular crystals (PLMCs) comprising pyridine (two polymorphs),\textsuperscript{193,194} pyrrole,\textsuperscript{195} pyridazine (two different thermodynamic conditions),\textsuperscript{196} and bipyridine\textsuperscript{197} (Fig. 7.2).

Based on these findings, we conclude that structural predictions are significantly improved when anharmonicity is accounted for via NpT-based AIMD simulations, yielding finite-temperature structural properties in dispersion-bound molecular crystals that are within 2% of experiment. However, the results reported herein still systematically underestimate the experimental PLMC cell volumes. For more accurate
Figure 7.2: Overlays of predicted (blue) and experimental (red) structures for the PLMC set. Shown here are all AIMD simulations (yellow background) for the PLMC set and the PI-AIMD simulation (green background) for pyridine-I.

And reliable predictions, we find that NQEs (such as zero-point fluctuations), many-body dispersion interactions, and Pauli repulsion all have a non-negligible influence over the structural and thermal properties of dispersion-bound molecular crystals. To demonstrate this, we now focus our attention on a detailed case study of the pyridine-I polymorph.

7.4 Nuclear Quantum Effects in Pyridine-I

While AIMD simulations are able to furnish accurate structural properties for the PLMCs across a range of thermodynamic conditions, the shape of the thermal expansion curve for deuterated pyridine-I from neutron powder diffraction experiments significantly differs from our theoretical predictions (Fig. 7.3). In this regard, the predicted $V(T)$ curve is linear across the entire $T$ range considered (i.e., 12 K–153 K at $p_{\text{expt}} = 1$ Bar), reflecting the use of classical mechanics for the nuclear motion. The experimental curve, on the other hand, shows non-linear behavior in this $T$ interval, with significant deviations from linearity at low temperatures, i.e., for $T \leq 50$ K. This observation strongly indicates that NQEs (in particular zero-point fluctuations)
Figure 7.3: Predicted and experimental thermal expansion curves for pyridine-I. Experimental data is included for pyridine-I (C₅H₅N, gold circle), from single-crystal X-ray diffraction, and fully deuterated pyridine-I (C₅D₅N, open black circles), from neutron powder diffraction. A fit of the experimental thermal expansion curve for deuterated pyridine-I using the Debye model for $V(T)$ is given by the purple line (Eq. (7.2)). Theoretical data is included for VC optimizations (blue circle), AIMD simulations (blue line), and PI-AIMD simulations (gold circle with error bar) at the PBE+vdW level; estimated PBE+MBD results (green line, Eq. (7.7)); estimated PBE0+MBD results (red line, see text for details).

play a non-negligible role in governing the structural and thermal properties of this dispersion-bound molecular crystal.
7.4.1 Experimental Behavior

To gain further insight into the thermal expansion behavior in this system, we utilize the Debye model, which is an isotropic acoustic approximation for the phonons in a solid. Within this framework, \( V(T) \) can be derived (see Appendix B.1) from the corresponding Gibbs free energy (at a given \( p \)) as:

\[
V(T) = V(0) + \left[ 3Nk_B \frac{\Theta'_D}{\Theta_D} D \left( \frac{\Theta_D}{T} \right) \right] T, 
\]

(7.2)
in which \( V(0) \) is the cell volume at 0 K, \( N \) is the number of atoms, \( \Theta_D = \Theta_D(p) \) is the Debye temperature, \( \Theta'_D = d \Theta_D(p)/dp \) is the pressure derivative of \( \Theta_D \) (which accounts for anharmonicity in the underlying PES), and \( D(\cdot) \) is the Debye function. Quite interestingly, we find that the experimental thermal expansion curve for \( C_5D_5N \) can be fit rather well with Eq. (7.2), as shown by the purple line in Fig. 7.3 (and Fig. B.1). A similarly good fit using the Debye interpolation formula was obtained for the methanol monohydrate molecular crystal. The validity of the Debye model for thermal expansion in pyridine-I is further supported by the physical value for the Debye temperature obtained from the fit, namely, \( \Theta_D = 235(5) \) K. This corresponds to an average sound velocity of 1710 m/s in this system, which falls within the experimentally determined range for the sound velocity of the closely related benzene molecular crystal.

Estimation of Sound Velocity in Pyridine-I

To demonstrate the validity of \( \Theta_D \), we take advantage of the relationship between the Debye temperature and the average sound velocity \( \bar{u} \), which is given as:

\[
\bar{u} = \frac{k_B \Theta_D}{\hbar} \left( \frac{V}{6\pi^2 N} \right)^{1/3}. 
\]

(7.3)
This average is computed from the transverse sound velocity $u_t$ and the longitudinal sound velocity $u_l$ as:

$$\frac{3}{\bar{u}^3} = \frac{2}{u_t^3} + \frac{1}{u_l^3}.$$  \hfill (7.4)

From Eq. (7.3) and our fitted value for $\Theta_D$, we find that $\bar{u} = 1710$ m/s in the pyridine-I polymorph. To justify this finding, we seek to compare the predicted $\bar{u}$ against the experimental value for this quantity. Due to the lack of direct experimental measurements of $\bar{u}$ in pyridine-I, we take the available measurements of this quantity in the closely related benzene molecular crystal at 170 K.\textsuperscript{205} In Ref.\textsuperscript{205} the longitudinal sound velocities were reported with anisotropy, and we take their averages to be $u_l = 3136$ m/s. Despite the lack of experimental measurements of the transverse sound velocities in benzene, this quantity can be estimated using the elastic properties of this molecular crystal, which leads to the inequality that $u_t > \sqrt{2}u_l$\textsuperscript{206} As such, we estimate the upper bound on the average sound velocity in benzene (and pyridine-I) as $\bar{u} = (3/(1 + 2\sqrt{8}))^{1/3}u_l = 2405$ m/s. To obtain a lower bound, we take advantage of the fact that sound typically travels slower in liquids than in solids, and estimate this quantity with the corresponding measurement in liquid pyridine at 293 K,\textsuperscript{207} giving $\bar{u} = 1437$ m/s. Although the estimated bounds are quite rough, this analysis provides us with a physical range for $\bar{u}$ in the pyridine-I polymorph. The fact that $\bar{u}$ obtained from the Debye model falls in this physical range of average sound velocity for pyridine-I reaffirms the validity of the fitted Debye temperature.

### 7.4.2 Feynman Path-Integral Study

The fact that $\Theta_D$ is slightly above the melting temperature of pyridine-I ($T_m = 232$ K) suggests that NQEs should have a sizable influence across the entire crystalline range of stability in this polymorph. To directly confirm the importance of NQEs in
determining the structure of pyridine-I, we performed a PI-AIMD simulation using PBE+vdW$^{TS}_{SC}$ at (153 K, 1 Bar).\textsuperscript{173,175} When compared to the 3\% volume expansion due to classical thermal fluctuations (\textit{cf.} the difference between the VC optimization at 0 K and AIMD simulation at 153 K, see Table \text{7.1}), we find that the inclusion of NQEs results in an additional 1.2\% expansion in the cell volume. This change is quite sizable ($\approx 40\%$ of the classical thermal expansion) and further reduces $V_{err}$ in pyridine-I to $+0.3\%$ with respect to experiment.

For the detailed case study of pyridine-I, path-integral (PI-AIMD) simulations in the $N_pT$ ensemble were performed using the $i$-PI package\textsuperscript{174} based on ionic forces and stress tensors generated at the PBE+vdW$^{TS}_{SC}$ level in \textsc{qe}. The valence electronic wavefunctions were converged to the Born-Oppenheimer (BO) surface using second-order damped dynamics (with a fictitious mass of $\mu = 400$ au and a time step of $\Delta \tau = 5$ au $\approx 0.12$ fs). Between two consecutive PI-AIMD steps, the energy convergence threshold was set to $6.0 \times 10^{-7}$ au and the force convergence threshold was set to $1.0 \times 10^{-3}$ au, with a time step of $\Delta t = 20$ au $\approx 0.48$ fs. For preliminary equilibration, we performed a 0.5 ps $NVT$ PI-AIMD simulation followed by a 0.5 ps $NpT$ PI-AIMD simulation. For the production run, we performed an 8 ps $NpT$ simulation with quantum nuclei, which was modeled with 8-beads (Trotter dimension) supplemented with the PIGLET method.\textsuperscript{175,176} To thermostat the cell, a classical GLE thermostat\textsuperscript{178} was employed for which we chose a fictitious mass consistent with a 200 fs timescale.

To investigate how NQEs lead to such an appreciable change in the pyridine-I cell volume, we first analyze how nuclear quantum fluctuations affect rigid molecular motions, \textit{i.e.}, translations and librations, in this molecular crystal. To quantify these effects, we compute the corresponding temperature correction ($\Delta T$) from the leading-order quantum correction to the momentum distribution for a Cartesian coordinate
(q) corresponding to the molecular center of mass:

\[ \Delta T = \frac{\hbar^2}{12MT^2} \langle F_q^2 \rangle, \quad (7.5) \]

in which \( M \) is the molecular mass and \( \langle F_q^2 \rangle \) is the mean-square force along \( q \) obtained by statistically averaging over the classical AIMD trajectory\[. In doing so, we find that \( \Delta T \approx 10 \) K for the rigid translational and librational modes in this system. However, the additional thermal expansion due to NQEs corresponds to a temperature elevation of \( \approx 50 \) K (assuming linear thermal expansion for \( T \geq 153 \) K), which is higher than the contributions from such rigid molecular motions and indicative of an additional mechanism for the observed expansion in pyridine-I.

**Role of Pauli Repulsion**

Figure 7.4: Comparison of the intermolecular HH, CH, and NH pair-correlation functions obtained from AIMD (blue) and PI-AIMD (red) simulations of pyridine-I using PBE+vdW\[TS\]. Dashed vertical lines represent the sum of the van der Waals (vdW) radii\[209\] for the respective pair of atoms.

To further understand the origin of this NQE-induced volume expansion, we computed a series of intermolecular pair-correlation functions involving the peripheral atoms on each pyridine molecule (\( g_{HH}(r), g_{CH}(r), g_{NH}(r) \)) based on AIMD and PI-AIMD simulations of this non-covalently bound molecular crystal (Fig. 7.4). From

\[ \Delta T = \frac{\hbar^2}{12I_jT^2} \langle \tau_j^2 \rangle, \]

in which \( I_j \) is the associated moment of inertia and \( \tau_j \) is the torque along the corresponding principal rotational axis, \( \phi_j \).
these plots, one can immediately see that the inclusion of NQEs—which cause individual pyridine molecules to fluctuate to a larger extent—lead to shorter intermolecular contacts (and hence more charge density overlap) among neighboring molecules in pyridine-I. For instance, the peripheral H atoms on neighboring pyridine molecules are closer by $\approx 0.2$ Å, and the probability of finding these two H atoms at a distance shorter than the sum of their vdW radii ($r = 2.4$ Å) is enhanced by 28% when accounting for NQEs. With atom pairs located within their respective vdW envelope, there will be an increase in the Pauli repulsion experienced by neighboring pyridine molecules, which in turn leads to a larger equilibrium cell volume in the molecular crystal. These findings hold for all atom pairs considered and demonstrate that intermolecular Pauli repulsion promoted by intramolecular quantum fluctuations is the dominant physical mechanism responsible for the observed cell volume increases in pyridine-I due to NQEs.

7.5 Thermal Expansion of Pyridine-I

Considering now the thermal expansivity (or thermal expansion coefficient),

$$\alpha(T) = \frac{1}{V(T)} \left( \frac{\partial V(T)}{\partial T} \right)_p,$$

we determined an experimental value of $\alpha = 3.5 \times 10^{-4}$ K$^{-1}$ for pyridine-I at (153 K, 1 Bar) based on the C$_5$D$_5$N thermal expansion curve. This value agrees quite well with the analytical finding from the Debye interpolation, i.e., $\alpha = 3.7 \times 10^{-4}$ K$^{-1}$, further illustrating the utility of this model in describing this system. However, the $\alpha$ value from classical AIMD simulations using PBE+vdW$^{TS\text{sc}}$ ($\alpha = 2.1 \times 10^{-4}$ K$^{-1}$) significantly underestimates the experimental value by $\approx 40\%$. Since cohesion in pyridine-I is dominated by dispersion interactions (Fig. B.2 and Table B.1 in Appendix B.2), this suggests that PBE+vdW$^{TS\text{sc}}$ overestimates the cohesive forces at
work in this non-covalently bound molecular crystal. This finding is consistent with other studies on molecular crystal lattice energies with this method. As such, we now investigate how a more comprehensive treatment of the beyond-pairwise many-body dispersion forces impacts our prediction of this thermal property in pyridine-I.

### 7.5.1 Role of Many-Body Dispersion

Beyond-pairwise dispersion interactions include terms such as the three-body Axilrod-Teller-Muto (ATM) contribution \( C_9/R^9 \), which is more short-ranged than the \( C_6/R^6 \) term in the effective-pairwise vd\( \text{W}^{\text{TS}} \) level and often provides a repulsive contribution to the binding energy. Since the inclusion of the ATM term alone is usually not sufficient to describe the full many-body expansion of the dispersion energy, we employ the many body dispersion (MBD) model to investigate how these higher-order non-bonded interactions affect the structural and thermal properties in pyridine-I. The MBD approach furnishes a description of all \( N \)-body dispersion energy contributions by mapping the atoms comprising the system onto a set of coupled quantum harmonic oscillators, and then computing the long-range correlation energy in the random-phase approximation (RPA). When coupled with DFT, MBD has been shown to provide an accurate and reliable description of the non-covalent

<table>
<thead>
<tr>
<th>Pyridine-I</th>
<th>( \alpha , (10^{-4} \text{ K}^{-1}) )</th>
<th>( \alpha_{\text{err}} , (%) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBE+vdW( ^{\text{TS}} ) ( \text{SC} )</td>
<td>2.1(3)</td>
<td>-40.0</td>
</tr>
<tr>
<td>est. PBE+MBD</td>
<td>3.7(5)</td>
<td>+5.7</td>
</tr>
<tr>
<td>est. PBE0+MBD</td>
<td>3.7(5)</td>
<td>+5.7</td>
</tr>
<tr>
<td>Debye Model</td>
<td>3.65(4)</td>
<td>+4.3</td>
</tr>
<tr>
<td>Expt. [104]</td>
<td>3.5(1)</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 7.2: Thermal expansivity (\( \alpha \)) values for pyridine-I at (153 K, 1 Bar) from theoretical simulations (at the PBE+vdW\( ^{\text{TS}} \) \( \text{SC} \), est. PBE+MBD, and est. PBE0+MBD levels), the Debye model, and experiment. Errors are reported with respect to the experimental value and the numbers in parentheses denote uncertainties in \( \alpha \).
interactions in molecules and materials,\textsuperscript{31} ranging from molecular crystals\textsuperscript{202,215,216} to complex polarizable nanostructures.\textsuperscript{217,218}

To account for many-body dispersion interactions, we estimated\textsuperscript{7} the average cell volume at the PBE+MBD level ($\langle V \rangle_{\text{MBD}}$) by Boltzmann reweighting the configurations from the PBE+vdW\textsubscript{SC} TS trajectory, \textit{i.e.},

$$\langle V \rangle_{\text{MBD}} = \frac{\langle V \exp[-\beta (U_{\text{MBD}} - U_{\text{TS}})] \rangle_{\text{TS}}}{\langle \exp[-\beta (U_{\text{MBD}} - U_{\text{TS}})] \rangle_{\text{TS}}},$$

(7.7)
in which $\beta$ is the inverse temperature, $U_{\text{TS}}$ and $U_{\text{MBD}}$ are the corresponding dispersion energies from these two methods, and $\langle \cdot \rangle_{\text{TS}}$ represents a statistical average over the PBE+vdW\textsubscript{SC} TS ensemble. The resulting estimates for $\langle V \rangle_{\text{MBD}}$ are shown in Fig. 7.3 and were used to determine that $\alpha = 3.7 \times 10^{-4}$ K$^{-1}$ at the PBE+MBD level, which is in significantly better agreement with the experimental value than PBE+vdW\textsubscript{SC} TS (Table 7.2). However, the estimated PBE+MBD cell volumes are noticeably larger than experiment, with predictions that are now less accurate than PBE+vdW\textsubscript{SC} TS. This may be due, in part, to the perturbative estimate of the PBE+MBD cell volumes using Eq. (7.7). Since MBD provides a more comprehensive treatment of dispersion interactions,\textsuperscript{29,31} this can also be indicative of other deficiencies present in the xc functional.

**Many-Body Dispersion Calculation**

For the MBD model, all calculations were performed using FHI-AIMS.\textsuperscript{219} In particular, We used the range-separated self-consistent screening (rsSCS) polarizability\textsuperscript{41} with a dipole field cutoff set to 120 Å for the calculation of the self-consistent screening (SCS) polarizability. The dipole field cutoff for the coupled-fluctuating dipole moment Hamiltonian was set to 100.0 Å. A supercell dimension of 50 Å was nec-

\textsuperscript{8}At the current time, analytical and on-the-fly evaluation of the ionic and cell forces at the MBD level were not available in the existing implementation of this scheme and these limitations will be addressed in future work.
necessary to converge the total MBD energy (for a 16-pyridine unit cell) to less than 1 meV. All calculations were run at the “really tight” setting.

### 7.5.2 Role of Exact Exchange via the exx Module

Hybrid functionals such as PBE0,\(^{60}\) which include a fraction of exact exchange, have been found to be more accurate overall than PBE in the treatment of molecular crystals.\(^{202}\) In the pyridine-I molecular crystal, we find that PBE0+vdW\(^{TS}\)\(_{SC}\)\(^{8,75,220}\) predicts a reduction in the 0 K cell volume by \(\Delta V = -1.02 \text{ Å}^3/\text{molec}\) when compared to PBE+vdW\(^{TS}\)\(_{SC}\). This effect likely originates from a combination of small changes in the molecular geometries as well as a better treatment of Pauli repulsion between neighboring molecules. Hence, we estimate the PBE0+MBD volume by adding this constant shift to the PBE+MBD results above (Fig. 7.3). This largely corrects the overestimation of the cell volume with PBE+MBD, and the resulting estimated PBE0+MBD values are now in better agreement with both the experimental volume (on an absolute scale) and the thermal expansivity. We stress here that an improved theoretical description of the Pauli repulsion might be of particular importance when simultaneously accounting for NQEs, which increase the amount of charge density overlap among neighboring molecules in pyridine-I.

### 7.6 Conclusions

In this chapter, we explored how a complex interplay between anharmonicity, NQE, many-body dispersion interactions, and Pauli repulsion influence the structural and thermal properties of dispersion-bound molecular crystals. By focusing on pyridine-I, we showed that the Debye model is well-suited to describe the thermal expansion behavior in this system across the range of available experimental temperatures. With a Debye temperature just above the melting point, we expect that NQE will be siz-
able across the entire crystalline range of stability in this polymorph. At low $T$, PI-AIMD simulations become computationally intractable (due to the steep increase in the required Trotter dimension) and it would be more efficient to include NQE via the quasiharmonic or self-consistent harmonic approximations.\textsuperscript{221–223} Based on our detailed case study of the pyridine-I molecular crystal, we expect that the qualitative trends outlined herein are robust and transferable to other dispersion-bound molecular crystals. In this regard, a logical extension of this work would include a fully self-consistent treatment of the pyridine-I molecular crystal (as well as other important non-covalently bound molecular crystals) that accounts for NQE as well as an improved description of the underlying electronic structure. Beyond the structural and thermal properties considered herein, the existence of thermodynamically relevant polymorphs further advocates for the determination of structures, stabilities, and properties of molecular crystals under $NpT$ conditions. Based on the findings presented in this chapter, free energy calculations that simultaneously account for nuclear quantum fluctuations and many-body dispersion interactions within a DFT scheme with reduced SIE will be required for an accurate and reliable description of dispersion-bound molecular crystals.
Part III

Appendices: Technical Details
Appendix A

Tight Convergence to the Electronic Ground State

In certain cases, tight convergence to the ground electronic state is necessary, e.g., in numerical phonon calculations. We found that in order to achieve such tight convergence, one has to pay particular attention to two factors in addition to the aforementioned five cutoffs: the convergence criteria of both the Marzari-Vanderbilt functional and the PE solver. For loosely converged Marzari-Vanderbilt functional, the MLWFs are not truly unique and the associated MLWF centers are subject to oscillations. As such, these non-unique orbitals also lead to oscillatory behavior in the orbital-dependent calculation of $E_{xx}$ that impairs tight convergence. On the other hand, PE solver with large error tolerance leads to inaccurate MLWF-product potentials, and spoils tight convergence of the electronic structure. In addition, while the choice of the aforementioned cutoffs do not introduce oscillatory behavior or noise that impair the convergence of $E_{xx}$, if these radii are not chosen large enough, they can lead to small upward drifts in the energy profile during the convergence. This small upward drift originates from slight inconsistency in the accuracies of $E_{xx}$ and $\{\tilde{D}^i_{xx}(r)\}$.
In QE, the default tolerances are set to for the change of the Marzari-Vanderbilt functional ($1.0 \times 10^{-8}$ au) for the second order damped dynamics. For the PE solution, we set the default tolerance of the residual error in the CG solver to $1.0 \times 10^{-6}$ au. Based on these default settings, the $(\text{H}_2\text{O})_{64}$ liquid example can be converged to around $1.0 \times 10^{-6}$ au accuracy. For tighter convergence in this example system (e.g., around $1.0 \times 10^{-12}$ au accuracy), one can set the tolerance of both the Marzari-Vanderbilt functional and the Poisson solver to $1.0 \times 10^{-12}$ au. With these tight tolerance and the set of cutoff radii mentioned in the Sec. 4.1, one may observe a slight upward drift of the order of $1.0 \times 10^{-7}$ au in the end of the convergence profile. Such upward drift can be avoided by using larger $R_{PE}$ and $R_{ME}$. 
Appendix B

Dispersion-Bound Molecular Crystals

B.1 Thermal Expansion of Pyridine-I in the Debye Model

For small $T$ corrections, the thermodynamic potential (Gibbs free energy) in the Debye model, $\Phi(p, T)$, can be expressed as:

$$\Phi(p, T) = \Phi(p, 0) + Nk_B T \left\{ 3 \log \left[ 1 - \exp \left( -\frac{\Theta_D(p)}{T} \right) \right] - D \left( \frac{\Theta_D(p)}{T} \right) \right\}, \quad (B.1)$$

where $\Phi(p, 0)$ is the temperature-independent thermodynamic potential at 0 K, $N$ is the number of atoms, $\Theta_D = \Theta_D(p)$ is the pressure-dependent Debye temperature, and $D(x)$ is the Debye function defined as:

$$D(x) = \frac{3}{x^3} \int_0^x \frac{z^3}{\exp(z) - 1} \, dz. \quad (B.2)$$
One can obtain the volume by taking the derivative of Eq. (B.1) with respect to \( p \) as follows:

\[
V(T) = \frac{\partial \Phi(p, T)}{\partial p} = \frac{d\Phi(p, 0)}{dp} + N k_B T \frac{\partial}{\partial p} \left\{ 3 \log \left[ 1 - \exp \left( - \frac{\Theta_D}{T} \right) \right] - D \left( \frac{\Theta_D}{T} \right) \right\}
\]

\[\text{(B.3)}\]

\[
= V(0) + N k_B T \left\{ 3 \frac{\partial}{\partial p} \log \left[ 1 - \exp \left( - \frac{\Theta_D}{T} \right) \right] - \frac{\partial}{\partial p} D \left( \frac{\Theta_D}{T} \right) \right\}, \quad \text{(B.4)}
\]

where \( V(0) \equiv \frac{d\Phi(p, 0)}{dp} \). Letting \( x = \frac{\Theta_D T}{T} \), Eq. (B.4) can then be simplified as:

\[
V(T) = V(0) + N k_B T \left\{ 3 \frac{\partial}{\partial x} \log \left[ 1 - \exp \left( - x \right) \right] - \frac{\partial}{\partial x} D \left( x \right) \right\} \left( \frac{\partial x}{\partial p} \right) \quad \text{(B.5)}
\]

\[
= V(0) + N k_B T \left\{ \frac{3 \exp (-x)}{1 - \exp (-x)} + \frac{3 D(x)}{x} \right\} \left( \frac{\partial x}{\partial p} \right) \quad \text{(B.6)}
\]

\[
= V(0) + N k_B T \left\{ \frac{3D(x)}{x} \right\} \left( \frac{\partial x}{\partial p} \right) = V(0) + N k_B T \left\{ \frac{3D(\Theta_D T)}{\Theta_D T} \right\} \left( \frac{\Theta_D T}{T} \right) \left( \frac{\partial \Theta_D T}{\partial p} \right), \quad \text{(B.7)}
\]

where \( \Theta'_D = d\Theta_D/dp \). Eq. (B.7) then simplifies to our working expression for \( V(T) \), i.e.,

\[
V(T) = V(0) + \left[ 3N k_B \frac{\Theta'_D}{\Theta_D} D \left( \frac{\Theta_D}{T} \right) \right] T, \quad \text{(B.8)}
\]

To utilize the Debye model for \( V(T) \) in Eq. (B.8), we performed a non-linear least square fitting via the Levenberg-Marquardt algorithm on the experimental thermal expansion curve for fully deuterated pyridine-I (C_5D_5N). From this fitting, we find that \( \Theta_D = 235(5) \) K, \( V(0) = 107.58(2) \text{ Å}^3/\text{molec} \), and \( k_B \Theta'_D = 0.33(1) \text{ Å}^3/\text{molec} \).

Here we note that \( \Theta'_D \) is a measure of the anharmonicity in the system since thermal expansion is present only when \( \Theta'_D \neq 0 \). The Debye model fits very well with the experimental data over the entire \( T \) range as shown in Fig. (B.1)
Figure B.1: Comparison of the thermal expansion curves of C$_5$D$_5$N from the Debye model (red) and experiment$^{[194]}$ (open black circles) up to the melting temperature ($T_m = 232$ K).

B.2 Cohesive Interactions in Pyridine-Like Molecular Crystals

In the set of PLMCs, all the comprising molecules have permanent dipole moments (with the exception of bipyridine due to the fact that its pyridine subunits are antiparallel to each other). In condensed-phase environments, these dipolements can be analyzed using maximally localized Wannier functions$^7$ (MLWFs) as described in Appendix B.3. For simplicity, we refer to classical electrostatic dipole–dipole interactions (due to permanent dipole moments) as dipole interactions, while quantum mechanical dispersion interactions (due to induced dipole moments from instantaneous charge fluctuations) are referred to as van der Waals (vdW) interactions. Without loss of generality, we will focus on pyridine-I using the structure obtained from a variable-cell (VC) optimization at the PBE+vdW$^{TS}$ level. We take the aforementioned molecular dipoles as point dipoles located at the center of mass of each pyridine molecule. Here, we find that the vdW interactions contribute to the majority of the cohesive energy in the condensed phase while the dipole interactions remain attractive and non-negligible (Table B.1). The relatively lower importance of the dipole interac-
tions originates from the anti-ferroelectric ordering of the molecular dipole moments in pyridine-I, which leads to quick saturation of the dipole interaction. As shown in Fig. B.2, the blue/red and purple/brown molecular pairs form two sets of anti-ferroelectric dipole arrangements. The rate of saturation in both the dipole and vdW interactions can be compared by considering their cumulative contributions to the cohesive energies, i.e., $E_{\text{dip}}(r)$ and $E_{\text{vdW}}(r)$ as a function of the pair distance $r$ as follows:

$$E_{\text{dip}}(r) = \sum_{\langle I,J \rangle} \varepsilon_{IJ}^{\text{dip}} \Theta \left( \frac{R_{IJ} - r}{\delta} \right)$$

$$E_{\text{vdW}}(r) = \sum_{\langle A,B \rangle} \varepsilon_{AB}^{\text{vdW}} \Theta \left( \frac{R_{AB} - r}{\delta} \right),$$

where $\langle I, J \rangle$ denotes summation over unique molecular pairs, $\varepsilon_{IJ}^{\text{dip}}$ is the dipole energy between point dipoles $I$ and $J$ separated by $R_{IJ}$, $\langle A, B \rangle$ denotes summation over unique atom pairs, $\varepsilon_{AB}^{\text{vdW}}$ is the vdW energy between atoms $A$ and $B$ separated by $R_{AB}$, and $\Theta(x) = \text{erfc}(x)/2$ is a smooth (Heaviside) step function that switches off the respective interaction within a distance range of $\delta$. The $\delta$ was chosen to be 1.1 Å, which is comparable to the size of a pyridine molecule. From the cumulative energies shown in Fig. B.2, the dipole interactions saturate much faster than the dominant vdW interactions in pyridine-I. From this analysis, vdW interactions are clearly the dominant cohesive interactions and are therefore crucial to predicting structural properties in the pyridine-I polymorph.

A similar analysis was performed on the remaining PLMCs and is shown in Table B.1. From this data, one can easily ascertain that the dipole interaction energies are minor when compared with the vdW interaction energies for all PLMCs.
Figure B.2: Comparison of $E_{\text{dip}}(r)$ and $E_{\text{vdW}}(r)$ using the pyridine-I structure obtained from a VC optimization at the PBE+vdW$_{\text{TS}}$ level of theory. The lower-left inset illustrates the MLWF centers (red) used in the molecular dipole analysis. The central insets show both the pyridine-I structure and the corresponding alignment of the molecular dipoles; molecules with the same color belong to the same generating asymmetric unit.

Table B.1: Comparison between molecular ring size, average molecular dipole moments ($\mu$), vdW interaction energies ($E_{\text{vdW}}$), dipole interaction energies ($E_{\text{dip}}$), and average thermal expansivities ($\alpha$) in the PLMCs. We computed $E_{\text{vdW}}$ and $E_{\text{dip}}$ based on structures obtained from VC optimizations at (0 K, $p_{\text{expt}}$). Thermal expansivities (thermal expansion coefficients) were computed via $\alpha(T) = \frac{1}{V(T)} \left( \frac{\partial V(T)}{\partial T} \right)$ at $T_{\text{expt}}$ using the predicted cell volumes from VC optimizations and AIMD simulations.

B.3 Molecular Dipole Analysis

To quantify the permanent dipole moments in the PLMCs, we performed a molecular dipole analysis using MLWFs, which provide a localized orthogonal representation of
the electronic wavefunctions. When each MLWF can be unambiguously assigned to a molecule (as illustrated in the lower left inset of Fig. B.2), one can compute the corresponding dipole (or higher multipole) moments of a condensed-phase molecule. Since this is the case for the PLMCs (and most molecular crystals), we compute the molecular dipole moment for a given molecule \( \mu_I \) using the locations and charges of the ions (nuclear contribution to \( \mu_I \)) as well as the centers and charges of the MLWFs (electronic contribution to \( \mu_I \)). For example, the dipole moment for an individual molecule in the pyridine molecular crystal would be computed via:

\[
\mu_I = \sum_{H \in I} R_H + 4 \sum_{C \in I} R_C + 5 \sum_{N \in I} R_N - 2 \sum_{W \in I} R_W,
\]

in which \( R_H, R_C, \) and \( R_N \), are the Cartesian coordinates of H, C, and N, respectively, while \( R_W \) corresponds to the MLWF centers (indexed by \( W \)) assigned to this pyridine molecule. Based on this molecular dipole analysis, the magnitude of the dipole moment of a gas-phase pyridine molecule computed was found to be 2.3 D at both the PBE and PBE+vdW\^[TS] levels (within a 15 Å cubic unit cell), which is quite consistent with the experimental value of 2.2 D. By analyzing the experimental structure of the pyridine-I molecular crystal, we find that the average magnitude of the molecular dipole moment increases to 3.3 D at the PBE+vdW\^[TS] level. This significant increase in the molecular dipole moment is evidence of the drastic difference between the gas- and condensed-phase environments. Average molecular dipole moments for each of the PLMCs are provided in Table B.1.
Bibliography


A.* 2012, 109, 14791–14795.

[40] DiStasio Jr., R. A.; Gobre, V. V.; Tkatchenko, A. Many-body van der Waals 
interactions in molecules and condensed matter. *J. Phys.: Condens. Matter 
2014, 26*, 213202.


[42] Blood-Forsythe, M. A.; Markovich, T.; DiStasio Jr., R. A.; Car, R.; Aspuru-
Guzik, A. Analytical nuclear gradients for the range-separated many-body dis-

[43] Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D. C.; Lundqvist, B. I. Van der 
246401.

[44] Vydrov, O. A.; Van Voorhis, T. Nonlocal van der Waals Density Functional 


[46] Sun, J.; Ruzsinszky, A.; Perdew, J. P. Strongly Constrained and Appropriately 

[47] Sun, J.; Remsing, R. C.; Zhang, Y.; Sun, Z.; Ruzsinszky, A.; Peng, H.; Yang, Z.; 
Paul, A.; Waghmare, U.; Wu, X.; Klein, M. L.; Perdew, J. P. Accurate first-
principles structures and energies of diversely bonded systems from an efficient 

Structural, electronic, and dynamical properties of liquid water by ab initio 
molecular dynamics based on SCAN functional within the canonical ensemble. 

[49] Calegari Andrade, M. F.; Ko, H.-Y.; Car, R.; Selloni, A. Structure, Polarization, 
and Sum Frequency Generation Spectrum of Interfacial Water on Anatase TiO₂. 

[50] Perdew, J. P.; Zunger, A. Self-interaction correction to density-functional ap-

[51] Cohen, A. J.; Mori-Sánchez, P.; Yang, W. Insights into Current Limitations of 


[210] A consistent picture is also provided by a Voronoi analysis of the AIMD and PI-AIMD trajectories to quantify the size of the fluctuating pyridine molecules. Here, we find that nuclear quantum fluctuations enlarge the average molecular size from 110.6 Å³ (AIMD) to 111.7 Å³ (PI-AIMD), with a corresponding increase in the associated standard deviation from 2.6 Å³ to 2.9 Å³.


