Title	Diffusion Monte Carlo Study of High-Pressure Solid Hydrogen Phases
Author(s)	Putu Bhargo Abhimana, Chrysnanda
Citation	
Issue Date	2024-09
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/19358
Rights	
Description	Supervisor: 前園 涼, 先端科学技術研究科, 修士(情報科学)

Master's Thesis

Diffusion Monte Carlo Study of High-Pressure Solid Hydrogen Phases

PUTU BHARGO ABHIMANA CHRYSNANDA

Supervisor: Ryo Maezono

Graduate School of Advanced Science and Technology
Japan Advanced Institute of Science and Technology

Information Science

September, 2024

Abstract

The high-pressure phase diagram of hydrogen has garnered much research interest ever since it came to light that hydrogen might transition into a metallic state at high pressure. Nowadays, the relevance of studying hydrogen under high pressure also comes from astrophysics as hydrogen is the largest constituent of most of the stars and gaseous planets, and is subject to extreme pressure and temperature in the interior of those celestial objects.

Hydrogen is believed to transform into its atomic metallic phase with a crystal structure belonging to the $I4_1/amd$ space group (also often called Cs-IV structure) at around 500 GPa. Said pressure is already near the limit of what is currently achievable by the experimental apparatus such as the diamond anvil cells. Consequently, one has to rely on computational methods for investigating the hydrogen phase diagram in the pressure regime beyond 500 GPa.

Previous computational studies employing a crystal structure search together with the density functional theory (DFT) method predicted that Cs-IV phase is stable up to 2000 GPa. Their structure search also uncovered some novel structures such as the $oC12$ and $cI16$ structures that are predicted to become more stable than Cs-IV above 2000 GPa.

In this thesis, the phase diagram of high-pressure solid hydrogen is investigated within the pressure range of 500 GPa to 2000 GPa. The diffusion quantum Monte Carlo (QMC) method is employed to evaluate the static enthalpy of candidate structures, while their zero-point energy is calculated via harmonic approximation from DFT forces. DMC is known to be more reliable in its treatment of electronic structure than DFT, at the cost of being more computationally heavy. Additionally, an evolutionary crystal structure search is employed to obtain more candidate structures that are energetically competitive.

This study revealed that the $Pnma$ structure discovered by the previous study is already more stable than Cs-IV at 2000 GPa. Two of the new candidate structures discovered in this work managed to become more stable than Cs-IV at some point within 500 GPa to 2000 GPa. The inclusion of the aforementioned structures modified the phase diagram of solid hydrogen in said pressure range. Hydrogen is predicted to undergo phase transitions in the order of $Cs\text{-IV} \rightarrow C2/c\text{-}6 \rightarrow Cs\text{-IV} \rightarrow C2/c\text{-}10 \rightarrow Pnma$, where the transitions happen at 800 GPa, 900 GPa, 1000 GPa, and 2000 GPa.

Keywords: First-principles calculation, High-pressure hydrogen phases, Crystal structure search, Diffusion Monte Carlo

Acknowledgments

I would like to express my gratitude to my thesis advisors, Prof. Ryo Maezono and Associate Prof. Kenta Hongo for their advice, support, and the opportunities they have given to me during my Master's year in JAIST.

This research topic on high-pressure solid hydrogen was brought to me by Assistant Prof. Ichiba Tomohiro, so I would also like to thank him for introducing me to this research field and for the guidance he provided. I am also thankful for his tutorial on running QMC simulations with QMCPACK and Nexus and many other things in other projects.

I would like to thank everyone in Maezono and Hongo laboratory, fellow students and the laboratory staff, for helping me out whenever I run into trouble in research or with the administrative documents.

And to all my friends in JAIST, thanks for making these past 2 years more epic and fun.

Lastly, I wish to thank my parents for their neverending support, motivation, and phone calls, and for helping me to stay connected with my family and friends at home.

Contents

List of Figures

Chapter 1

Introduction

1.1 Background

Hydrogen is the first element in the periodic table. Owing to its simplicity, the hydrogen atom has been central to the early development of quantum mechanics which was heavily focused on explaining its experimentally observed emission spectra. The interest in its high-pressure phase was sparked by a series of theoretical studies that hinted at the interesting properties dense hydrogen might possess. In 1935, E. Wigner and H. B. Huntington predicted that hydrogen could undergo metallization at a pressure of 25 GPa [3]. Thirty years later in 1968, N.W. Ashcroft predicted that metallic hydrogen might be a high-temperature superconductor [4]. These predictions fascinated many physicists to investigate the possible transitions between the insulating, conducting, and the speculated superconducting phase of hydrogen.

The study of high-pressure hydrogen phases also holds great significance in astrophysics. As the most abundant element in the universe, hydrogen, together with helium are the main components of many celestial objects such as gaseous planets and stars. These two elements alone comprise around 85% and 75% of the total mass of Jupiter and Saturn, respectively [5]. Much of the knowledge on these gas giants comes from the data collected by past spacecraft flyby missions. For Jupiter, a very accurate measurement of its gravitational field was obtained by the Juno spacecraft in 2016 [6]. Unfortunately, even with much available data on gravitational and magnetic fields, inferring the internal structure of gaseous planets is not possible without any accurate knowledge of the behavior of hydrogen, helium, and their mixture under high pressure and temperature that exist in the interior of such planets. For instance, the pressure-temperature profile of the planets depends heavily on the predicted compressibility and specific heat of those elements. The percentage by mass of heavier elements also needs to be inferred from the pressure-temperature chart under the constraints imposed by the recorded gravity data [5]. It is evident that gaining a more in-depth understanding of high-pressure hydrogen phases will have a significant impact on the modeling of many celestial objects.

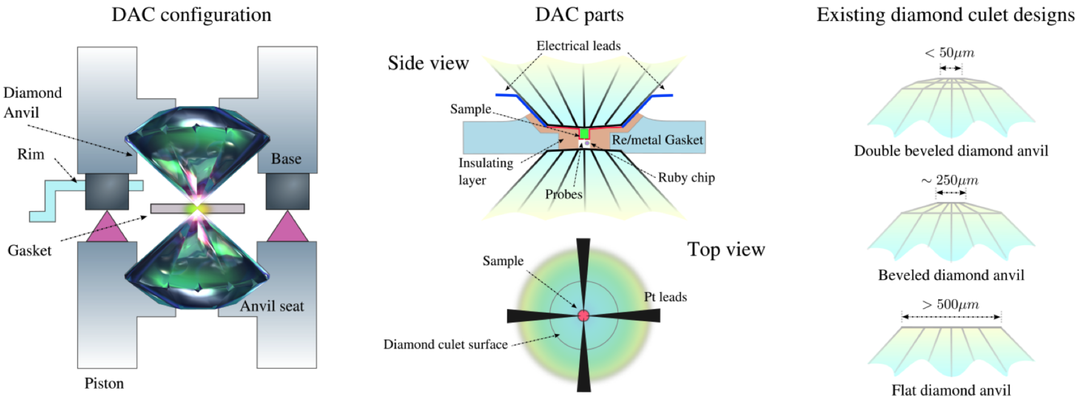

Figure 1.1: The left panel shows the schematics of the DAC while the center panel shows the more detailed parts around the metal casket encapsulating the studied material. The types of diamond cutlet designs are shown on the right panel. (Figure adapted from ref [1])

The need for high-pressure studies in general, not just for hydrogen, brought a lot of progress in the development of experimental apparatus to accommodate high-pressure material research in laboratory settings. Such a pressure regime is currently achievable with the aid of diamond anvil cells (DACs). In DACs, the sample to be studied is compressed between the sharp tips of two opposing diamonds (Fig. 1.1). Due to diamond being the hardest known material, a very high pressure could be achieved without breaking the anvil. Furthermore, the desirable optical properties of diamonds made DACs even more versatile. Diamond is transparent for a wide range of electromagnetic spectrum, not only towards visible light but also for X-rays. This accommodates the utilization of many spectroscopy techniques to measure the properties of the sample as the pressure is being exerted. X-ray diffraction, in particular, is a popular method to determine the structure of the materials. Raman spectroscopy and reflectance measurement are also often incorporated into experimental research with DACs to observe the change in the characteristics of the sample. Such measurements are crucial to detect the phase transition in materials, as it is often accompanied by a sudden change in optical properties.

The advancement in DACs technology has increased the achievable pressure up to 400 GPa for conventional DACs [7], while the recently developed toroidal diamond anvil cell (T-DAC) can reach a pressure up to 600 GPa [8]. Thus, extensive experimental studies of high-pressure hydrogen phases have been conducted up to almost 500 GPa. These studies revealed the richness and the complexity of the hydrogen phase diagram, while also showing that the existence of metallic hydrogen was at a much higher pressure than the previous theoretical prediction. Solid hydrogen is known to have 6 different phases which are labeled as phase I to VI (Fig. 1.2). In phase I, solid hydrogen takes the hexagonal close-packed structure with a freely rotating hydrogen molecule on its lattice points [9]. The transition to phase II (often called the broken symmetry phase) happens at around 110 GPa, at which each hydrogen molecule loses its rotational freedom and is uniformly aligned in the same orientation.

As the pressure accessible by DACs increased, the discovery of phase transition to phase II and phase III revealed the inadequacy of the characterization methods to determine the structure of hydrogen. The difficulty arises from the hydrogen's small scattering efficiency, which produces a weak diffraction intensity that can be easily obscured by the background noise. In fact, among all the discovered solid hydrogen phases, only the structure of phase I was able to be identified experimentally. For the others, it was necessary to consider the computational methods. These approaches work by considering some proposed structures or candidate structures of a given phase for evaluation. Such evaluations will produce some predictions regarding the candidate structure's properties that can be compared with the results from experiments. The investigation of phase III

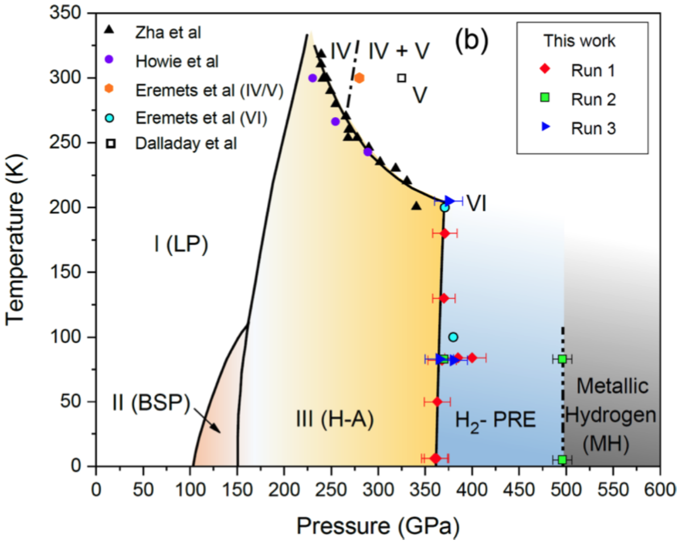

Figure 1.2: The known phase diagram of high-pressure hydrogen. In the figure, the low-temperature phase VI is labeled as H_2-PRE (Figure adapted from ref [2])

using computational methods of density functional theory (DFT) and a random crystal structure search yielded the $C2/c$-24 and $Cmca$-12 structures [10] which are still believed to be the most suitable candidates for phase III. At an even higher pressure, the transition into metallic hydrogen has been observed. An experiment conducted by Loubeyre, et al reported the transition to metallic hydrogen which they attributed to the transition from the $C2/c$-24 phase to the molecular metallic $Cmca$-12 phase at the pressure of 425 GPa [11]. Earlier in 2017, Dias, et al claimed to have finally realized the atomic metallic hydrogen predicted by Wigner and Huntington from the observed transition at 495 GPa in their experiment [12]. So far, the best candidate structure for the atomic metallic hydrogen phase is presumed to be the $I4_1/amd$ structure (Cs-IV) [13, 14].

Due to the limitations of current DACs capabilities, the pressure regime beyond 500 GPa has been much left untouched by researchers. This attainable pressure might be enough to crack the mystery of insulator-to-metal transition in high-pressure hydrogen. However, much less is understood about the hydrogen phase diagram above 500 GPa. There is still no consensus about how much higher in pressure the Cs-IV phase will still be the most stable, and what kind of phase it will transition into. Some computational studies using density functional theory, however, have predicted that Cs-IV is stable up to around 2000 GPa, at which pressure the DFT results also uncovered some more stable structures [15]. Consequently, the hydrogen phase diagram beyond 500 GPa still has much room to be explored.

1.2 Motivation

The technological advancement in high-pressure material research has no doubt shed some light on the high-pressure phase of hydrogen. However, as previously stated, the achievable pressure is still far from the condition that exists in the interior of gaseous planets and stars. Therefore, there is no choice but to rely on computational methods in investigating such pressure regimes. With no experimental data to confirm or deny the results, there is no way to verify the accuracy of the findings. Hence, it becomes very crucial to choose a reliable method in computational research. Up until now, the atomic metallic hydrogen (Cs-IV) phase is believed to be stable up until around 2 TPa. Further than that, previous DFT studies have uncovered some novel structures such as $oC12$, $R\bar{3}m$, and $cI16$, which are expected to be more stable than Cs-IV. However, DFT is known to be unreliable in its treatment of electronic structure and sometimes produces contradictory results due to the approximations in its implementation. As such, this research tries to asses the phase diagram of solid hydrogen between 500 GPa to 2000 GPa more accurately by employing the quantum Monte Carlo (QMC) method. QMC method could yield much more accurate calculations at the expense of larger computational costs.

Another challenge in the computational research of high-pressure hydrogen phases comes during the exploration of the candidate structures. Previous research has uncovered the fact that hydrogen possesses a complex energy landscape, with many kinds of structures differing so slightly in energy. As such, the crystal structure search algorithm has become vital to acquire competitive candidate structures. Due to the lack of previous research effort directed at the pressure range considered in this work, there is a high possibility of undiscovered structures. Therefore, this research will also employ a crystal structure search algorithm to find new candidate structures.

1.3 Chapter Organization

This master's thesis is organized as follows:

- Chapter 1 provides an introduction to high-pressure hydrogen research as well as the motivation for this master's thesis
- Chapter 2 explains the basic theories related to the methods utilized in this thesis, which mainly covers the crystal structure search based on the evolutionary algorithm and the electronic structure computational methods of DFT and QMC.
- Chapter 3 gives a more detailed description of the system considered in this thesis and the technical settings of the calculations.
- Chapter 4 presents the main results of this thesis followed by discussions.
- Chapter 5 concludes the discussion from the previous chapter and explores some possible future studies.

Chapter 2

Background Theories

2.1 Many-body Schrödinger Equation

The main problem of electronic structure calculations is to solve or find the approximate solution to the Schrödinger equation governing the time evolution of any quantum system:

$$i\hbar\frac{d}{dt}\left|\Psi(t)\right\rangle = \hat{H}\left|\Psi(t)\right\rangle \tag{2.1}$$

where $\Psi(t)$ is the wavefunction or the quantum state, $\hat{H}$ is the Hamiltonian which depends on the quantum system in question, and t is time. If the Hamiltonian does not explicitly depend on time, the solution to the above time-dependent Schrödinger equation can be separated into the spatial and temporal parts, that is, $\Psi(\mathbf{r}, t) = \psi(\mathbf{r})T(t)$. The temporal part takes the form of $T(t) = e^{-iEt/\hbar}$, while the spatial part has to satisfy the time-independent Schrödinger equation that is essentially an eigenvalue problem:

$$\hat{H}\psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{2.2}$$

where E is the energy of the system. The time-dependent Schrödinger equation admits many pairs of eigenvectors and eigenvalues $\{\psi_n(\mathbf{r}), E_n\}$. These eigenvectors are complete, in the sense that they can be used as a basis set to construct any solution of the time-dependent Schrödinger equation. In general, the state vector $|\Psi(t)\rangle$ can be written as

$$|\Psi(t)\rangle = \sum_n c_n e^{-iE_n t/\hbar} |\psi_n\rangle \tag{2.3}$$

When considering a many-body system made of multiple atoms and electrons, the Born-Oppenheimer approximation which treats the nuclei as stationary objects is often invoked. The motivation for such an assumption comes from the fact that the nucleus is around 1000 times heavier than electrons. The dynamics of electrons are much faster than the nucleus and from the electron's point of view, the nuclei are pretty much frozen. In the same way, the electrons appear to instantaneously relax to the equilibrium state from the perspective of the nuclei. This leads to the separation of the wavefunction into the ionic and electronic parts, as well as to the decoupling of the Schrödinger equation into 2 equations, one for the electrons and another one for the nuclei. Upon considering frozen nuclei, only the equation for the electrons needs to be considered and the electronic equation now takes the nuclei's positions as a set of parameters. Under this scheme, the Hamiltonian can be expanded as:

$$\hat{H} = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i,j} \frac{Z_i}{|\mathbf{r}_i - \mathbf{R}_j|} + \sum_{i<j} \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} \tag{2.4}$$

where $\mathbf{r_i}$ denotes the position of the i-th electron, while $\mathbf{R}_i$ and Z_i denote the position and charge of the i-th nucleus. The first term on the right-hand side of Equation (2.4) is the kinetic energy operator of the electrons. The second, third, and fourth terms are the electron-electron, electron-ion, and ion-ion Coulomb interaction terms, respectively. In Equation (2.4) above, the Hamiltonian is expressed in the atomic units, in which the mass of the electron, the reduced Planck constant, and the electric charge of the electron are equal to 1.

2.2 Density Functional Theory

Density functional theory has gained much popularity as an ab initio method for assessing the properties of materials. Its good predictive capabilities across various materials and low computational cost have made DFT a standard technique in computational material science and chemistry. As the name suggests, DFT revolves around finding the electron density $\rho(\mathbf{r})$ as opposed to finding the electronic wavefunction $\Psi(\mathbf{r}, t)$ directly by solving the many-body Schrödinger equation. This approach simplifies the task greatly as density is a three-dimensional quantity and the wavefunction is $3N$-dimensional, where N is the number of electrons in the system.

Hohenberg-Kohn Theorem

The Hohenberg-Kohn (HK) theorem [16] serves as a theoretical basis of DFT, which consists of two parts:

1. The first part of the theorem states that the external potential $v_{ext}(\mathbf{r})$ is a functional of the ground-state electron density $\rho(\mathbf{r})$ (up to a constant). Since this external potential fixes the form of the Hamiltonian, any property of the ground state such as its energy is also a functional of the ground-state electron density. In terms of $\rho(\mathbf{r})$, the functional of ground-state energy $E[\rho(\mathbf{r})]$ can be written as:

   $$E[\rho(\mathbf{r})] = T[\rho(\mathbf{r})] + V_{ee}[\rho(\mathbf{r})] + V_{ext}[\rho(\mathbf{r})] \tag{2.5}$$

   where $T[\rho(\mathbf{r})]$ is the kinetic energy functional, $V_{ee}[\rho(\mathbf{r})]$ is the electron-electron interaction energy functional, and $V_{ext}[\rho(\mathbf{r})]$ is the functional of the energy from an external potential.

2. The second part of the theorem states that the ground-state energy functional will evaluate to its minimum value if and only if the correct ground-state density is inserted. This is essentially the variational principle obeyed by the energy functional

with respect to electron density. Mathematically, for any trial electron density $\rho'(\mathbf{r})$, the energy functional satisfies:

$$E[\rho(\mathbf{r})] \leq E[\rho'(\mathbf{r})] \tag{2.6}$$

The equality will only be satisfied when $\rho'(\mathbf{r}) = \rho(\mathbf{r})$. The trial electron density should follow the same constraint imposed by the ground-state density, that is, the trial density is always positive and integrates to the correct number of electrons N.

These two statements solidify the theoretical foundation of the modern density functional theory with electronic density as the principal quantity. At first glance, it seems as if solving the Schrödinger equation can now be achieved by applying variational calculus to find the electronic density $\rho(\mathbf{r})$ that minimizes the energy functional $E[\rho(\mathbf{r})]$. However, this approach is hindered by the very fact that the exact form of the functionals, $T[\rho(\mathbf{r})]$ and $V_{ee}[\rho(\mathbf{r})]$, are still unknown. These two functionals are "universal" in the sense that their form does not depend on the system. Hence, they are often lumped together as a single universal functional $F[\rho(\mathbf{r})] = T[\rho(\mathbf{r})] + V_{ee}[\rho(\mathbf{r})]$. On the other hand, the $V_{ext}[\rho(\mathbf{r})]$ can be expressed as

$$V_{ext}[\rho(\mathbf{r})] = \int v_{ext}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} \tag{2.7}$$

which depends on the system in question. For instance, if the external potential is due to the electron-nuclei interaction, then $V_{ext}[\rho(\mathbf{r})]$ would depend on the positions of the nuclei and their charges.

Kohn-Sham Equation

The problem now turned to finding an accurate approximation of the universal functional. In practice, however, it was very difficult to formulate a satisfactory approximation for the kinetic energy functional $T[\rho(\mathbf{r})]$ of interacting electrons. In 1965, Kohn and Sham proposed a way to side-step such a problem by mapping the system of interacting particles (called the reference system) into a fictitious system of non-interacting particles with the same density as the reference system [17]. The idea is to separate $F[\rho(\mathbf{r})]$ into 3 parts:

$$F[\rho(\mathbf{r})] = T_s[\rho(\mathbf{r})] + V_H[\rho(\mathbf{r})] + E_{XC}[\rho(\mathbf{r})] \tag{2.8}$$

where $T_s[\rho(\mathbf{r})]$ is the kinetic energy of a non-interacting system with density $\rho(\mathbf{r})$ and $V_H[\rho(\mathbf{r})]$ is classical electrostatic energy:

$$V_H[\rho(\mathbf{r})] = \frac{1}{2}\int d\mathbf{r}d\mathbf{r}'\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{2.9}$$

The third term, $E_{XC}[\rho(\mathbf{r})]$, is called the exchange-correlation energy which encompasses the energy contribution that is not covered by $T_s[\rho(\mathbf{r})]$ and $V_H[\rho(\mathbf{r})]$. By minimizing Equation (2.8) with respect to electronic density, one can obtain the Kohn-Sham (KS) equation:

$$\left[-\frac{1}{2}\nabla^2 + v_{KS}(\mathbf{r})\right]\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}) \tag{2.10}$$

where $\psi_i(\mathbf{r})$ and ϵ_i are the single-particle orbital and its eigenvalue. The Equation (2.10) above is exactly the time-independent Schrödinger equation of a particle moving under the external potential $v_{KS}(\mathbf{r})$ called the Kohn-Sham potential, which is given by

$$v_{KS}(\mathbf{r}) = \int d\mathbf{r}'\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + v_{XC}(\mathbf{r}) + v_{ext}(\mathbf{r}) \tag{2.11}$$

By definition, the exchange-correlation potential $v_{XC}(\mathbf{r})$ can be written as:

$$v_{XC}(\mathbf{r}) = \frac{\delta E_{XC}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} \tag{2.12}$$

Solving the Kohn-Sham equation yields a set of single orbital wavefunctions that reproduces the electronic density $\rho(\mathbf{r})$ of the reference system:

$$\sum_{i=1}^{N}|\psi_i(\mathbf{r})|^2 = \rho(\mathbf{r}) \tag{2.13}$$

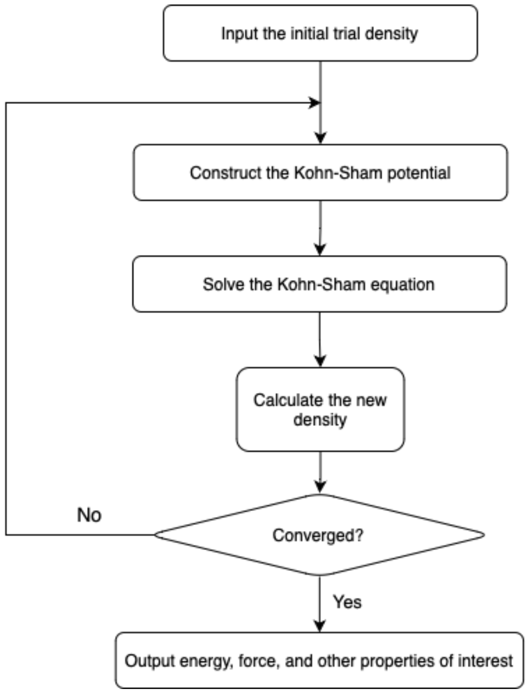

Figure 2.1: The flowchart of self-consistent calculation for solving the Kohn-Sham equation

The complexity of the problem has been yet again reduced to tackling a set of equations of a single non-interacting particle. In calculating the electronic density, the Kohn-Sham equation needs to be solved. However, to construct the Kohn-Sham equation, $v_{KS}(\mathbf{r})$ needs to be calculated first from the electronic density. Consequently, the Kohn-Sham equation is solved in a self-consistent manner. By first using a trial density, one can begin the process of using the newly obtained density as the input to the Kohn-Sham equation repeatedly until the results converge within the desired criteria. This iterative method is often called the self-consistent field (SCF) calculation, and the algorithm is shown in Fig. 2.1.

Exchange-Correlation Functional

Among the 3 functionals appearing on the right-hand side of Equation (2.8), only the exact form of exchange-correlation functional is unknown and needs to be approximated. The accuracy of DFT calculations depends heavily on the form of exchange-correlation functional employed. The simplest approximation of exchange-correlation functional is the local-density approximation (LDA):

$$E_{XC}^{LDA}[\rho(\mathbf{r})] = \int \epsilon_{xc}^{hom}[\rho(\mathbf{r})]\rho(\mathbf{r})d\mathbf{r} \tag{2.14}$$

where $\epsilon_{xc}^{hom}[\rho(\mathbf{r})]$ is the exchange-correlation energy per particle of an electron gas with uniform density ρ. As such, LDA is expected to be accurate only when describing a system with slowly varying electronic density. Despite this assumption, LDA has been found to be surprisingly reliable for a wide range of systems. Further improvement in the exchange-correlation functional can be achieved by including the gradient of the density. This approach is called the generalized gradient approximation (GGA), for which the functional takes the form:

$$E_{XC}^{GGA}[\rho(\mathbf{r})] = \int f(\rho(\mathbf{r}), \nabla\rho(\mathbf{r}))d\mathbf{r} \tag{2.15}$$

In practice, the functions ϵ_{xc}^{hom} and f inside the LDA and GGA functionals need to be parametrized by either experimental data or by very accurate ab initio calculations. An example of such parametrized functionals is the Perdew and Zunger LDA [18]. Over the years, many exchange-correlation functionals have been proposed as further improvements from the GGA functionals. As one example, the so-called Meta-GGA improves upon GGA functional by including higher derivatives of electron density and kinetic energy density [19].

2.3 Variational Quantum Monte Carlo

The variational quantum Monte Carlo (VMC) is a method based on the variational principle and the Monte Carlo integration. In the following section, the vector $\mathbf{R}$ is a $3N$-dimensional vector containing the positions of all the electrons in the system:

$$\mathbf{R} = \big(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3 ....., \mathbf{r}_N\big) \tag{2.16}$$

where $\mathbf{r}_i$ is the position of the i-th electron. In the context of quantum Monte Carlo, the vector $\mathbf{R}$ is often referred to as a walker.

Monte Carlo Integration

The Monte Carlo integration is a method to approximate integrals via random sampling. Suppose the following integral of multivariate function $g(\mathbf{R})$ needs to be evaluated:

$$I = \int g(\mathbf{R}) d\mathbf{R} \tag{2.17}$$

The above integral can be transformed by introducing an importance function $\mathcal{P}(\mathbf{R})$ so that it becomes

$$I = \int f(\mathbf{R}) \mathcal{P}(\mathbf{R}) d\mathbf{R} \tag{2.18}$$

where $f(\mathbf{R}) = g(\mathbf{R})/\mathcal{P}(\mathbf{R})$. If $\mathcal{P}(\mathbf{R})$ satisfies the constraints of a probability density function, that is, $\mathcal{P}(\mathbf{R}) \geq 0$ and $\int \mathcal{P}(\mathbf{R}) d\mathbf{R} = 1$, then I can be interpreted as the expectation value of $f(\mathbf{R})$. Such expectation value can be approximated by calculating the average of a finite number of samples:

$$I \approx \frac{1}{M} \sum_{m=1}^{M} f(\mathbf{R}_m) \tag{2.19}$$

From the central limit theorem, as the sample size increases, the distribution of the sample mean will get arbitrarily close to a normal distribution with mean μ_f and variance σ_f^2/M, where μ_f and σ_f^2 are the mean and variance of $f(\mathbf{R})$, respectively. This means that the error bar of the approximation decays as $1/\sqrt{M}$ regardless of the dimension of the integral.

In principle, reducing the variance of $f(\mathbf{R}_m)$ would lead to a smaller error bar for a given sample size. Therefore, it is important the choose a suitable importance function $\mathcal{P}(\mathbf{R})$ to make the Monte Carlo integration converge faster. It can be shown that the smallest variance is obtained when $\mathcal{P}(\mathbf{R}) = |g(\mathbf{R})/I|$, for which $f(\mathbf{R})$ becomes a constant and σ_f^2 becomes zero. However, this is meaningless as I is the quantity that needs to be evaluated. Hence another method is needed to make an importance function $\mathcal{P}(\mathbf{R})$ that is close to $g(\mathbf{R})/I$.

Variational Principle

The variational principle states that the expectation value of an observable $\langle O \rangle$ corresponding to the operator $\hat{O}$ with respect to a trial wavefunction $\Psi_T(\mathbf{R})$ is greater or equal to that evaluated by using the exact ground state wavefunction. Taking the energy as the observable with Hamiltonian as the corresponding operator, the variational principle can be expressed as:

$$E_0 \leq E_V = \frac{\displaystyle\int \Psi_T^*(\mathbf{R})\hat{H}\Psi_T(\mathbf{R})d\mathbf{R}}{\displaystyle\int \Psi_T^*(\mathbf{R})\Psi_T(\mathbf{R})d\mathbf{R}} \tag{2.20}$$

where E_0 is the ground-state energy. The above equation can be rewritten into a form suitable for Monte Carlo integration by multiplying the numerator with $\Psi_T(\mathbf{R})/\Psi_T(\mathbf{R})$ which leads to:

$$E_V = \frac{\displaystyle\int |\Psi_T(\mathbf{R})|^2 \frac{\hat{H}\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} d\mathbf{R}}{\displaystyle\int |\Psi_T(\mathbf{R})|^2 d\mathbf{R}} = \int \mathcal{P}(\mathbf{R}) E_L(\mathbf{R}) d\mathbf{R} \tag{2.21}$$

The energy can now be interpreted as the expectation value of the "local energy" $E_L(\mathbf{R})$:

$$E_L(\mathbf{R}) = \frac{\hat{H}\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} \tag{2.22}$$

from the probability density $\mathcal{P}(\mathbf{R})$ shown below:

$$\mathcal{P}(\mathbf{R}) = \frac{|\Psi_T(\mathbf{R})|^2}{\displaystyle\int |\Psi_T(\mathbf{R})|^2 d\mathbf{R}} \tag{2.23}$$

The Monte Carlo integration can finally be applied:

$$E_V \approx \frac{1}{M} \sum_{m=1}^{M} E_L(\mathbf{R}_m) \tag{2.24}$$

Within the VMC framework, the variational principle is also used to optimize the initial wavefunction over a set of parameters to make it closer to the actual ground-state wavefunction. From the perspective of variational principle, the energy E_V in Equation (2.24) is a suitable choice of cost function to be minimized for optimizing the wavefunction. From the perspective of Monte Carlo integration, however, the quality of the wavefunction can be evaluated by the variance of the local energy (σ_E^2). Choosing said variance to be minimized will also improve the convergence of the Monte Carlo integration. Alternatively, one can also choose to optimize a combination them both, such as σ_E^2/E_V^2.

Metropolis Algorithm

To calculate the expectation value of $E_L(\mathbf{R})$, a set of points $\{\mathbf{R}_m\}$ needs to be sampled from the high-dimensional probability density function. Direct sampling is rendered impossible due to the complexity of said function. In such cases, the metropolis algorithm is often used to provide an accurate sampling of any probability function without knowing its normalization factor [20]. The algorithm generates a series of sampling points by propagating a walker according to the following procedure:

1. Initialize a walker to position $\mathbf{R}$.
2. Propose a trial move of the walker to a new position $\mathbf{R}'$. The probability of proposing a move is taken from a probability density function $T(\mathbf{R}' \leftarrow \mathbf{R})$.
3. Accept the trial move to the new position $\mathbf{R}'$ with the probability

$$A(\mathbf{R}' \leftarrow \mathbf{R}) = \min\left(1, \frac{T(\mathbf{R} \leftarrow \mathbf{R}')\mathcal{P}(\mathbf{R}')}{T(\mathbf{R}' \leftarrow \mathbf{R})\mathcal{P}(\mathbf{R})}\right) \tag{2.25}$$

If the trial move is accepted, $\mathbf{R}'$ becomes the next sampling point. If the trial move is rejected, **R** becomes the next sampling point.

4. Return to step 2 and repeat the process over again.

The acceptance probability in Equation (2.25) is high whenever the trial move is proposed to a region with a higher probability density than the initial position. As such, the walker will tend to aggregate in the region with high probability density. Initially, the random walk undergoes an equilibration period where the initial points of the walk still depend on the starting point. However, the random walk will eventually settle down to a unique probability distribution mimicking the probability density to be sampled.

2.4 Diffusion Quantum Monte Carlo

The diffusion quantum Monte Carlo (DMC) method is based on the imaginary time Schrödinger equation, which can be obtained by rewriting the time-dependent Schrödinger equation using imaginary time $\tau = it$:

$$-\frac{\partial \left|\Phi\right\rangle}{\partial \tau} = (\hat{H} - E_T)\left|\Phi\right\rangle \tag{2.26}$$

where E_T is the energy offset. The addition of energy offset to the imaginary time Schrodinger equation allows the extraction of the ground-state wavefunction $\left|\Psi_0\right\rangle$ from $\left|\Phi\right\rangle$. To see this, consider the effect of applying (imaginary) time evolution operator on the initial wavefunction $\left|\Phi(\tau = 0)\right\rangle$:

$$\left|\Phi(\tau)\right\rangle = e^{-(\hat{H}-E_T)\tau}\left|\Phi(\tau = 0)\right\rangle \tag{2.27}$$

Decomposing $\left|\Phi\right\rangle$ using the eigenvectors of $\hat{H}$ (labeled as $\{\Psi_i\}$), and taking the long imaginary time limit give:

$$\lim_{\tau\to\infty} e^{-(\hat{H}-E_T)\tau}\left|\Phi(\tau = 0)\right\rangle = \lim_{\tau\to\infty}\sum_i c_i\left|\Psi_i\right\rangle e^{-(E_i-E_T)\tau} \tag{2.28}$$

Where E_i is the eigenvalue of the eigenvector $|\Psi_i\rangle$. It is evident that by setting $E_T = E_0$, every contribution other than that from the ground-state wavefunction $|\Psi_0\rangle$ decays exponentially because their energies are higher than the ground-state energy. Hence, as $\tau \to \infty$, the operator $e^{-(\hat{H}-E_T)\tau}$ effectively projects out the ground state wavefunction $|\Psi_0\rangle$. This is in contrast with the wavefunction optimization in VMC where the excited states do not disappear completely and still contribute to the energy.

Diffusion Equation

Equation (2.26) in position basis can be expressed as:

$$-\frac{\partial \Phi(\mathbf{R},\tau)}{\partial \tau} = \left[-\frac{1}{2}\sum_i \nabla_i^2 + V(\mathbf{R}) - E_T\right]\Phi(\mathbf{R},\tau) \tag{2.29}$$

Without the potential term $[V(\mathbf{R}) - E_T]$, the above equation becomes the diffusion equation in a $3N$-dimensional space, and $\Phi(\mathbf{R},\tau)$ can be interpreted as the particle density. Such an equation can be simulated by an ensemble of particles performing a random walk. In contrast, if the laplacian term was absent, the Equation (2.29) reduces to a rate equation that describes the growth or decay of the particle density. Hence, the evolution of $\Phi(\mathbf{R},\tau)$ governed by Equation (2.29) can be simulated by a random walk of particles combined with a branching process to account for the growth and decay in the number of particles. To propagate the density into a later imaginary time, it is essential to transform Equation (2.29) into its integral form. This can be done directly by taking the position representation of Equation (2.27):

$$\Phi(\mathbf{R},\tau_{\mathbf{0}}+\tau) = \int G(\mathbf{R}\leftarrow\mathbf{R}',\tau)\Phi(\mathbf{R}',\tau_{\mathbf{0}})d\mathbf{R}' \tag{2.30}$$

where the Green's function $G(\mathbf{R}\leftarrow\mathbf{R}',\tau)$ is:

$$G(\mathbf{R}\leftarrow\mathbf{R}',\tau) = \langle\mathbf{R}|\, e^{-(\hat{H}-E_T)\tau}\, |\mathbf{R}'\rangle \tag{2.31}$$

The exact form of the Green's function is unknown. However, if the time step τ is small, it can be approximated using the Trotter-Suzuki formula which leads to its factorization into the diffusion and branching term [21]:

$$G(\mathbf{R} \leftarrow \mathbf{R}', \tau) \approx G_d(\mathbf{R} \leftarrow \mathbf{R}', \tau) G_b(\mathbf{R} \leftarrow \mathbf{R}', \tau) \tag{2.32}$$

with an error proportional to τ^3. The first factor on the right-hand side of Equation (2.32) is the diffusion term:

$$G_d(\mathbf{R} \leftarrow \mathbf{R}', \tau) = (2\pi\tau)^{-3N/2} \exp\left[-\frac{(\mathbf{R} - \mathbf{R}')^2}{2\tau}\right] \tag{2.33}$$

which takes the form of a Gaussian function. It is the exact Green's function for the diffusion equation and is taken to be the transition probability of a walker to move from $\mathbf{R}'$ to $\mathbf{R}$ within a timespan of τ. The second exponential factor:

$$G_b(\mathbf{R} \leftarrow \mathbf{R}', \tau) = \exp\left[-\tau\left(V(\mathbf{R}) + V(\mathbf{R}') - 2E_T\right)/2\right] \tag{2.34}$$

is called the branching term and it governs the death or birth of a walker according to the following procedure:

1. if $G_b(\mathbf{R} \leftarrow \mathbf{R}', \tau) < 1$ then the branching term serves as the probability of the walker continuing its walk.

2. if $G_b(\mathbf{R} \leftarrow \mathbf{R}', \tau) \geq 1$ then the walker continues its walk, and a new walker is created in the same position with a probability of $G_b(\mathbf{R} \leftarrow \mathbf{R}', \tau) - 1$.

Since only the short-time approximation of Green's function is known, propagating the initial wavefunction into a large enough imaginary time is done by repeatedly applying Equation (2.30) over a small time step τ. The realization of such a process using a set of diffusing walkers is done by a repeated operation of moving each walker based on transition probability $G_d(\mathbf{R} \leftarrow \mathbf{R}', \tau)$ and incorporating the branching process governed by $G_b(\mathbf{R} \leftarrow \mathbf{R}', \tau)$ after each step.

Importance Sampling

The branching term (Equation 2.34) is controlled by $V(\mathbf{R})$, which could blow up whenever the electrons are too close to each other or to the nucleus. This causes the simulation to be inefficient and impractical as the number of walkers in the simulation will fluctuate rapidly. This problem is mitigated by performing an importance-sampling transformation. The transformation is done by multiplying Equation (2.29) with a "guiding" or "trial" wavefunction $\Psi_T(\mathbf{R})$:

$$-\Psi_T(\mathbf{R})\frac{\partial\Phi(\mathbf{R},\tau)}{\partial\tau} = \Psi_T(\mathbf{R})\left[-\frac{1}{2}\nabla^2 + V(\mathbf{R}) - E_T\right]\Phi(\mathbf{R},\tau) \tag{2.35}$$

Upon rearranging the terms on the above equation, one ends up with:

$$-\frac{\partial f(\mathbf{R},\tau)}{\partial\tau} = -\frac{1}{2}\nabla^2 f(\mathbf{R},\tau) + \nabla\cdot[\mathbf{v_D}(\mathbf{R})f(\mathbf{R},\tau)] + [E_L(\mathbf{R}) - E_T]\,f(\mathbf{R},\tau) \tag{2.36}$$

where $\nabla = (\nabla_1, \nabla_2, \ldots, \nabla_N)$ is the $3N$-dimensional gradient operator, $f(\mathbf{R},\tau) = \Phi(\mathbf{R},\tau)\Psi_T(\mathbf{R})$ is the importance-sampled or mixed wavefunction, and E_L is the local energy computed by using $\Psi_T(\mathbf{R})$ only. The term $\mathbf{v_D}(\mathbf{R})$ is called the drift velocity, and is defined as:

$$\mathbf{v_D}(\mathbf{R}) = \nabla\ln|\Psi_T(\mathbf{R})| = \Psi_T^{-1}(\mathbf{R})\nabla\Psi_T(\mathbf{R}) \tag{2.37}$$

After performing the importance-sampling transformation, $E_L(\mathbf{R})$ appears in the place of $V(\mathbf{R})$. As long as the guiding wavefunction is close enough to the actual ground state wavefunction, the variance of the local energy is guaranteed to be minimal. In the infinity time limit, $f(\mathbf{R},\tau)$ becomes the mixed distribution $\Psi_0(\mathbf{R})\Psi_T(\mathbf{R})$. The integral form of Equation (2.36) is

$$f(\mathbf{R},\tau_\mathbf{0}+\tau) = \int\tilde{G}(\mathbf{R}\leftarrow\mathbf{R}',\tau)f(\mathbf{R}',\tau_\mathbf{0})d\mathbf{R}' \tag{2.38}$$

where the importance-sample Green's function can also be factorized into the diffusion and branching parts $\tilde{G}(\mathbf{R}\leftarrow\mathbf{R}',\tau) \approx \tilde{G}_d(\mathbf{R}\leftarrow\mathbf{R}',\tau)\tilde{G}_b(\mathbf{R}\leftarrow\mathbf{R}',\tau)$, that can be expressed as:

$$\tilde{G}_d(\mathbf{R} \leftarrow \mathbf{R}', \tau) = (2\pi\tau)^{-3N/2} \exp\left[-\frac{(\mathbf{R} - \mathbf{R}' - \tau\mathbf{v_D}(\mathbf{R}'))^2}{2\tau}\right] \tag{2.39}$$

$$\tilde{G}_b(\mathbf{R} \leftarrow \mathbf{R}', \tau) = \exp\left[-\tau\left(E_L(\mathbf{R}) + E_L(\mathbf{R}') - 2E_T\right)/2\right] \tag{2.40}$$

The approximation assumes that the drift velocity is constant between $\mathbf{R}$ and $\mathbf{R}'$. This is a reasonable assumption except in the case where the guiding wavefunction $\Psi_T(\mathbf{R})$ varies rapidly. It is preferable to incorporate a rejection step similar to the Metropolis algorithm to reduce the time step error. When a move is proposed for the walker, the acceptance probability of that move is formulated as:

$$p_{accept}(\mathbf{R} \leftarrow \mathbf{R}') = \min\left(1, \frac{\tilde{G}_d(\mathbf{R}' \leftarrow \mathbf{R}, \tau)\Psi_T(\mathbf{R})^2}{\tilde{G}_d(\mathbf{R} \leftarrow \mathbf{R}', \tau)\Psi_T(\mathbf{R}')^2}\right) \tag{2.41}$$

The diffusion quantum Monte Carlo method described so far relies on the assumption that the wavefunction is positive definite and thus can be interpreted as a particle density. However, this is not the case when dealing with fermions for which the wavefunction can take both negative and positive values. To circumvent this problem, the fixed-node approximation is often utilized. The approximation works by placing absorbing barriers at the location of the nodes of the trial wavefunction $\Psi_T(\mathbf{R}, \tau)$. Whenever a walker tries to cross the barrier, the walker is either deleted or the move is rejected. Hence, each nodal pocket is isolated from each other and the DMC simulation can be run on each nodal pocket separately. Since the trial wavefunction is not exactly the ground-state wavefunction, their nodes might not completely coincide with each other. This introduces another source of error in the DMC calculations.

2.5 Evolutionary Crystal Structure Search

Crystal structure search or prediction tackles the problem of finding the most stable crystal structure for a given chemical composition and environment, such as temperature or pressure. It has been critical for the computational investigation and discovery of novel materials. In recent years, many efforts have been directed at applying known optimization algorithms for the crystal structure search. One of such attempts combined the well-known evolutionary algorithms into the context of structure search. This evolutionary crystal structure prediction still retains the general procedure of the evolutionary algorithm as follows:

1. Generate initial candidate structures to initialize the population.
2. Evaluate each structure in the population using some fitness criteria.
3. Select a subset of the population based on their fitness value to create new structures.
4. Generate new candidate structures from the selected population in step (3) via evolutionary operators.
5. Replace some or all of the population with the newly generated structures from step (4).
6. Repeat steps (2) to (5) until the convergence threshold is reached.

The evolutionary algorithm mimics the process of natural evolution where individuals with better qualities have more chance to pass their genes to the next generation. The later generations will be populated with individuals of high fitness value. Within the context of crystal structure prediction, the fitness value is usually the energy or enthalpy of the structures, which are evaluated by ab initio methods such as DFT. The evolutionary operators such as the hereditary (crossover) and mutation operators are also modified to be more suitable for crystal structure prediction. The crossover process can be implemented by combining matching slab cuts from parent structures in a meaningful way to create a child structure. Meanwhile, the mutation can be introduced by random distortion in the cell dimension and atomic position of the structure, or by permutations of different atomic species within the cell.

2.6 Phonon Calculations

Phonon describes the collective vibration of atoms inside a crystal structure. Calculating phonon properties is relevant in the study of crystals as it allows the investigation of many crystal properties such as thermal capacity, thermal expansion, and zero-point energy. Most importantly, the dynamical stability of the crystal can also be inspected via such calculations, which is crucial for studies utilizing crystal structure search techniques. Any dynamical instability exhibited by a crystal structure would make it impossible to exist as it would transform into another structure.

The theory behind phonon calculations starts from the expression of the energy of the system as a function of atomic positions $E(\{r_{j,l,\alpha}\})$, where $r_{j,l,\alpha}$ denotes the α cartesian component (x, y, or z direction) of the position vector of j-th atom in the l-th unit cell. If the atoms are displaced by a small amount around their equilibrium positions, the resulting energy can be written using the Taylor expansion as:

$$E = E_0 + \sum_{j,l,\alpha} \frac{\partial E}{\partial u_{j,l,\alpha}} u_{j,l,\alpha} + \sum_{j,l,\alpha} \sum_{j',l',\beta} \frac{1}{2} \frac{\partial^2 E}{\partial u_{j,l,\alpha} \partial u_{j',l',\beta}} u_{j,l,\alpha} u_{j',l',\beta} + \dots \tag{2.42}$$

where E_0 is the energy when each atom is at its equilibrium position and $u_{j,l,\alpha}$ is the displacement of an atom indexed by j, l in α cartesian direction. Using the harmonic approximation, only up to the second-order terms of the Taylor expansion are considered. The first-order derivative of E with respect to u in Equation (2.42) is the negative of atomic force, which vanishes when every atom is at its equilibrium position. The second-order derivative on said equation is called the force constant $\Phi_{\alpha,j,l;\beta,j'l'}$, and can be approximated by using the finite displacement method:

$$\Phi_{\alpha,j,l;\beta,j'l'} = \frac{\partial^2 E}{\partial u_{j,l,\alpha} \partial u_{j',l',\beta}} \approx -\frac{F_{j',l',\beta}(u_{j,l,\alpha}) - F_{j',l',\beta}}{u_{j,l,\alpha}} \tag{2.43}$$

where $F_{j',l',\beta}(u_{j,l,\alpha})$ is the force acting on j'-th atom of l'-th unit cell when an atom with index j, l is displaced by u in α direction. From the force constant matrix, the dynamical matrix $D_{\alpha,j;\beta,j'}(\mathbf{q})$ can be obtained by taking its Fourier transform:

$$D_{\alpha,j;\beta,j'}(\mathbf{q}) = \frac{1}{\sqrt{m_j m_{j'}}} \sum_{l'} \Phi_{\alpha,j,0;\beta,j'l'} \exp\left(i\mathbf{q} \cdot \mathbf{r}_{l'}\right) \tag{2.44}$$

where m is the atomic mass, $\mathbf{r}_l$ is the vector defining the position of l-th unit cell, and $\mathbf{q}$ is the wavevector. The dynamical matrix satisfies an eigenvalue equation that yields a set of phonon frequencies $\omega(\mathbf{q})$ and phonon eigenvectors $\eta(\mathbf{q})$. Once the atomic forces for a crystal with displaced atoms have been calculated from DFT or other methods, the dynamical matrix can readily be constructed and its eigenvalue equation can be solved. From the obtained phonon frequencies, the phonon density of states (DOS) can be constructed and any dynamical stability in the crystal is signified by the existence of imaginary phonon frequencies.

Chapter 3

Methods and Computational Details

3.1 Outline of Methods

3.1.1 General Workflow

The main equation used to assess the stability of a structure relative to the others is the enthalpy (H), which is defined as

$$H = E + PV + E_{ZPE} \tag{3.1}$$

where E is the energy to be evaluated by the electronic structure calculation, P is the pressure, V is the volume, and E_{ZPE} is the zero-point energy (ZPE) of the system. The above equation contains the E_{ZPE} term which is not often included in the definition of enthalpy. When studying a hydrogen system, its small atomic mass makes the ZPE large enough to disrupt the relative stability of the structures and thus needs to be included. We will strictly refer to the enthalpy without the inclusion of ZPE as static enthalpy, while that with the inclusion of ZPE as dynamic enthalpy. The most stable structure at a given pressure corresponds to the structure with the lowest enthalpy. Whenever the most stable structure changes due to a change in pressure, a phase transition happens.

The nature of the stability evaluation makes it necessary to consider many candidate structures and rank them in terms of their enthalpies. We considered the already-known candidate structures discovered in previous research for evaluation. Additional candidate structures were obtained by employing a crystal structure search. All these structures were optimized using DFT. We chose to carry out this optimization step with DFT since doing so using QMC would be too computationally heavy and therefore impractical. The QMC was only used to calculate the static enthalpy from the DFT-optimized structures. Finally, the zero-point energy was calculated using the harmonic approximation from the DFT force to obtain the dynamic enthalpy.

3.1.2 Known Candidate Structures

From the literature, we have obtained known candidate structures from the low-pressure side around 500 GPa and from the high-pressure side around 2000 GPa. The low-pressure structures considered in this work are $C2/c$-24 [10], $C2/c$-12 [10], $Cmca$-4 [10], $Cmca$-12 [10], and $mC24$ [15]. Meanwhile, for the high-pressure side, the following were considered: $oC12$ [15], $cI16$ [15], $R\bar{3}m$ [14], $R3m$ [14], $C2/m$-3 [14], and $Pnma$ [15]. The $mC24$, $oC12$, and $cI16$ structures belong to $C2/c$, $Cmcm$, and $I\bar{4}3d$ space groups, respectively. Some of these candidate structures are shown in Fig. 3.1. All structures including Cs-IV were optimized in the pressure range of 500 GPa to 2000 GPa. Upon optimization, some of the structures transformed into previously undiscovered structures. For those with transformation, we added "-low" and "-high" suffixes in their name to distinguish the low-pressure and high-pressure counterparts of the structure.

3.2 Computational Details

3.2.1 Crystal Structure Search

In this thesis, the USPEX code [22–24] was employed for the evolutionary crystal structure search. The main objective of the search was to minimize the static enthalpy. The structure search was run by constraining the number of atoms in the unit cell to be within 12 to 48 atoms and limiting the population size of each generation to contain

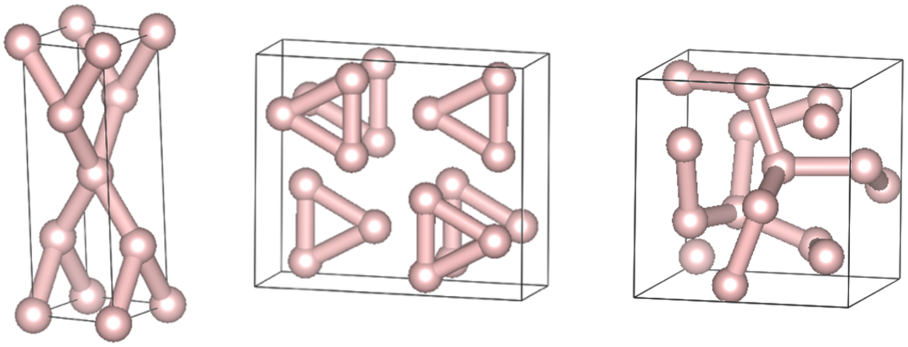

Figure 3.1: Some of the known candidate structures, from left to right: Cs-IV, $oC12$, and $cI16$. Fictitious bonds were added for clarity.

only 100 structures. We chose the top 50% of the current generation to create the next generation by applying the genetic operator. The heredity operation, random structure generation, lattice mutation, and atomic position mutation were responsible for creating 30%, 50%, 10%, and 10% of the new generation. The process was stopped until the 40th generation was achieved or if no improvement was observed for more than 10 generations.

This crystal structure search works in conjunction with DFT to optimize each structure within each generation. The structural optimization was carried out using the Perdew-Burke-Ernzerhof (PBE) [25] DFT method implemented within the Vienna *ab initio* simulation package (VASP) [26–29]. A three-stage relaxation was employed, where each stage utilized finer calculation settings than the previous stage to steadily increase the accuracy. In the final stage, the plane wave basis cutoff energy was set to 600 eV and the structural optimizations were carried out until the energy converged within 1×10^{-4} eV and the SCF calculation converged within 1×10^{-5} eV. For all stages, the k-mesh was set with a default spacing of 0.5 Å^{-1}.

3.2.2 Density Functional Theory Calculations

The obtained candidate structures from the literature and from the crystal structure search were re-optimized in varying pressure with more accurate DFT calculations. This section does not describe the DFT calculations used in the creation of DMC trial wavefunction. Here, the DFT calculations were also performed using the VASP package. The exchange-correlation functional was replaced by Van der Waals density functional (vdW-DF) [30–32], which is known to be able to accurately estimate the hydrogen molecular bond length and its pressure dependence that ultimately affects the calculated value of ZPE. The plane wave basis energy cutoff was set to 1500 eV and the spacing of k-point mesh was set to 0.17 Å^{-1}. This value of k-spacing was enough to assure the SCF convergence within 1 meV/atom for the Cs-IV, $mC24$-low, and two random $P1$ structures from the structure search conducted at 500 GPa. The structure optimizations were carried out until the atomic force was converged within 1×10^{-3} eV/Å and each SCF iteration converged within 1×10^{-6} eV. We used the hard projector-augmented wave (PAW) pseudopotential to represent the ionic core.

We have also investigated the validity of using pseudopotential on such a high-pressure region. There was a concern about the hydrogen atoms becoming too close to each other under a pressure as high as 2000 GPa, which would break the validity of using a pseudopotential. Investigation of all known candidate structures revealed that the shortest bond length was around 0.71 Å, which was only a 4% decrease from the molecular hydrogen bond length of 0.74 Å.

3.2.3 Quantum Monte Carlo Calculations

The quantum Monte Carlo calculations were performed by using the QMCPACK simulation code [33, 34] with Nexus workflow management system [35]. We used the fixed-node trial wavefunction in the form of Slater-Jastow type. The orbitals used to construct the Slater determinant were supplied from DFT calculations implemented in Quantum Espresso [36, 37]. In the following text, unless specified otherwise, every convergence test was done for all known candidate structures aside from $C2/c-24$, $R\bar{3}m$, $R3m$, $C2/m$-3-low, and $Pnma$. The plane wave basis cutoff energy was set to 350 eV, which was sufficient to make the energy converge within 1 meV/atom. For the Jastrow function, we included one-, two-, and three-body terms and the parameters were optimized using the linear optimization algorithm. The twist-averaged boundary condition with a 5×5×5 twist grid size was employed for systems with 200 atoms. This was enough to make the VMC energy converge within 2.7 meV/atom (0.1 mHa/atom) for Cs-IV, $mC24$-low, and $mC24$-high structures. The timestep was set to 0.005 Ha^{-1}, which was enough to make the energy converge within 1.36 meV/atom (0.05 mHa/atom), except in $cI16$-high which was only converged within 1.6 meV/atom. To reduce the finite size effect, we ran 2 QMC calculations for each structure using a system size of around 200 atoms and 400 atoms. Unfortunately, the extrapolation of the energy to an infinite system size resulted in an undesirable relative static enthalpy graph. Hence, we ignored the extrapolated energy and reported the DMC static enthalpy from the 400 atoms system only.

3.2.4 Phonon Calculations

The phonon calculations were carried out under the harmonic approximation implemented in the Phonopy package [38, 39], which used the force from DFT calculations executed with VASP. The SCF-only DFT calculations were done under the same settings described in section 3.2.2. We employed a supercell for the structures so that the cell lattice vectors were at least 10 Å in magnitude. This made the number of atoms in the simulations vary from around 1000 atoms to 2000 atoms.

Chapter 4

Results and Discussion

4.1 DFT Static Enthalpy of Known Structures

We have optimized the candidate structures in the whole pressure range considered in this work, and their DFT static enthalpies have been calculated (Fig. 4.1). As stated in Chapter 3, we found that some of these structures transformed into previously undiscovered structures. Two of the low-pressure known structures, the $mC24$-low and $Cmca$-12-low, transformed at 1300 GPa and 1800 GPa into $mC24$-high and $Cmca$-12-high, respectively. Similarly, the $cI16$-high and $C2/m$-3-high from the high-pressure side both transformed when the pressure was lowered to 900 GPa. Two other structures with transformation were $Cmca$-4-low and $C2/c$-24. For the former, its high and low-pressure counterparts had already been discovered in previous studies [10]. While for the latter, it underwent transformation 8 times between 1000 GPa and 2000 GPa. Those 8 new structures from $C2/c$-24 were excluded in this thesis.

Among these new structures, $mC24$-high has the lowest static enthalpy that is relatively constant at 6 to 7 meV/atom above Cs-IV within 1100 GPa to 2000 GPa. In contrast, the static enthalpies of known high-pressure structures at 2000 GPa are in the range of 26 to 33 meV/atom higher than Cs-IV, but they possess a steeper downward trend with increasing pressure compared to $mC24$-high. On the other hand, the low-pressure counterparts of the high-pressure known structures are not energetically competitive with the low-pressure known structures. At DFT static enthalpy level, hydrogen undergoes phase transitions in the order of $Cmca$-12-low → $Cmca$-4-low → $mC24$-low → Cs-IV, which happen at 520 GPa, 744 GPa, and 950 GPa, respectively. The Cs-IV phase is predicted to be stable up to 2000 GPa.

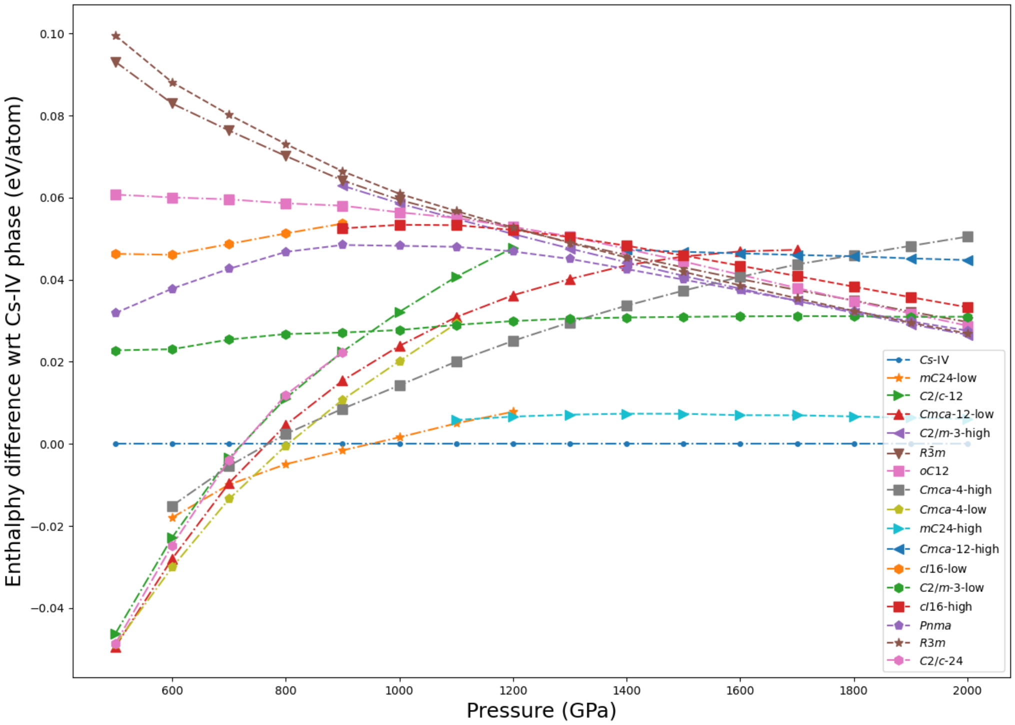

Figure 4.1: The relative DFT static enthalpies per atom of known candidate structures with respect to Cs-IV phase. Note the sudden termination on the enthalpy line for some structures due to structural transformation.

The relative DFT static enthalpy of $C2/c$-24, Cs-IV, $Cmca$-12-low, and $Cmca$-4-low and their transition pressure are consistent with the DFT results of previous work by McMinis, et al [40], which used the same vdW-DF functional as the one considered in this work. However, the static enthalpy difference between Cs-IV and the high-pressure known structures in this work is larger than what was reported in the previous DFT study by Liu, et al [15]. The reason for this discrepancy might be the difference in the DFT exchange-correlation functional employed, as said reference used the PBE functional. To confirm this, the $oC12$, $R\bar{3}m$, Cs-IV, and $Pnma$ structures were optimized using the PBE functional at 2000 GPa. Upon recalculating their relative DFT static enthalpy with respect to Cs-IV, the $oC12$, $Pnma$, and $R\bar{3}m$, were 4.8 meV/atom, 7.91 meV/atom, and 9.5 meV/atom above Cs-IV. This result is more consistent with the previous work, however, $oC12$, and $Pnma$ were still a few meV/atom higher in static enthalpy than the reference.

4.2 Crystal Structure Search Results

The structure search was conducted at five different pressures: 500 GPa, 800 GPa, 1100 GPa, 1400 GPa, and 1700 GPa. After re-optimization using DFT, some of the found structures underwent noticeable changes in their space group. Most of the time, the Phonopy package did not deem the resulting structure as that from a different symmetry due to the strict default symmetry criteria. Hence, we will report the symmetry of the structure after the symmetry tolerance was loosened to 1.e-3 or 1.e-2. For the naming, we adopted the common convention of using its space group number and the number of atoms in the primitive cell, separated by a dash (-).

Even though the maximum number of generations in the structure search was set to 40, every structure search found its best structure quite early and the searches were terminated within 12 to 18 generations. This led to around 1 thousand discovered structures for each search. We discarded those with the lowest symmetry of $P1$ and $P\bar{1}$ space groups as every single known candidate structure belongs to a space group with higher symmetry. This reduced the number of structures from each structure search to around 100 to 200 structures. In the discussion that follows, only the enthalpy graph for a few structures with the lowest static enthalpy will be shown.

4.2.1 Structure Search at 500 GPa

The 500 GPa structure search found many structures that were more stable than Cs-IV in terms of static enthalpy. This was expected as 500 GPa is at the border of the stability region of the Cs-IV phase. The $Cmca$-4-low, $Cmca$-4-high, $mC24$-low, and Cs-IV structures were rediscovered by this structure search. However, nothing resembling the $Cmca$-12-low, $C2/c$-12 and $C2/c$-24 structures were found. By excluding the already known structure, the 13 lowest enthalpy found structures at 500 GPa were in the $Pbam$, $Fmmm$, $P2_1/m$, $Fddd$, $P2_1$, $C2/m$, $C2$, and $C2/c$ space groups, whose relative static enthalpies varies between 45 meV/atom to 23 meV/atom lower than Cs-IV. Their enthalpy is shown in Fig. 4.2 below. Among them, the lowest one in static enthalpy at 500 GPa is the $Pbam$-8. However, this structure has been previously found as the slightly dynamically unstable version of the candidate structure for H_2-PRE phase [41].

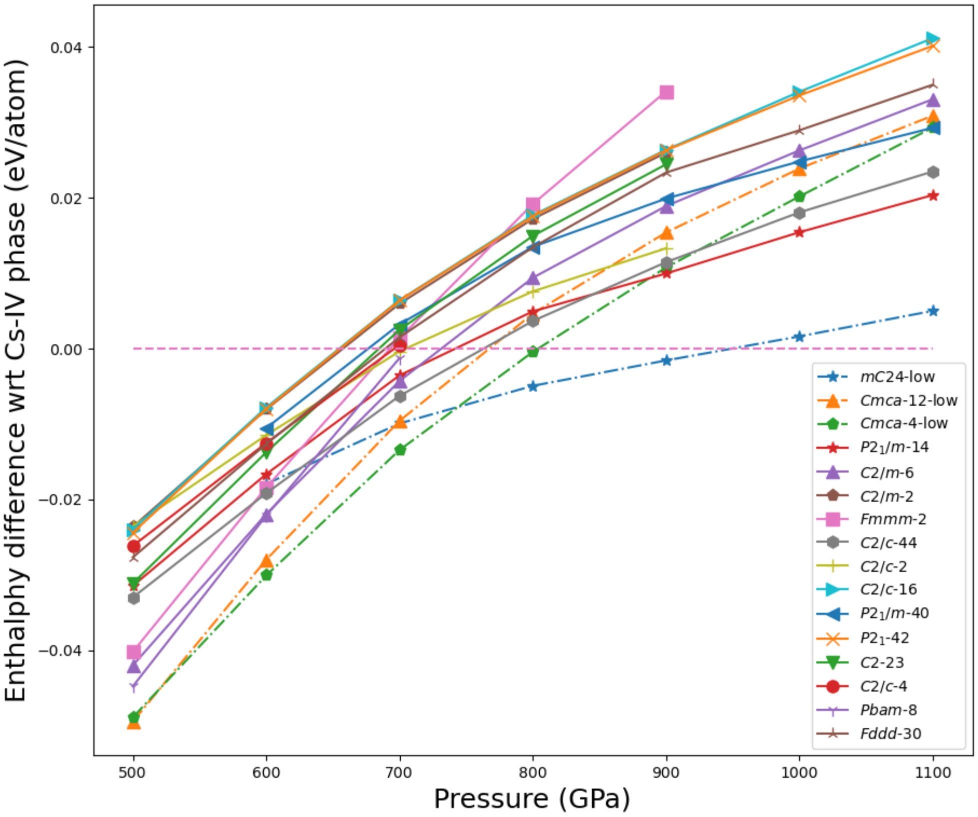

Figure 4.2: The relative static enthalpies per atom of found candidate structures from 500 GPa structure search with respect to Cs-IV phase. The relative static enthalpy of $mC24$-low, $Cmca$-4-low, and $Cmca$-12-low are also shown.

As can be seen in Fig. 4.2, within 500 GPa to 1100 GPa, none of the found structures managed to get under the static enthalpy of the most stable known structure at any given pressure. Hence, the phase transitions of hydrogen remain unchanged from the conclusion of section 4.1.

4.2.2 Structure Search at 800 GPa

The 800 GPa structure search yielded the $mC24$-low, $mC24$-high and Cs-IV structures. Furthermore, this structure search rediscovered the $Cmca$-12-low structure that was not discovered by the 500 GPa structure search. However, no $Cmca$-4-low and $Cmca$-4-high structures were found from the search, even though their static enthalpies lie between that of $mC24$-low and $Cmca$-12-low at 800 GPa. The enthalpy graphs of the found structures are shown in Fig. 4.3.

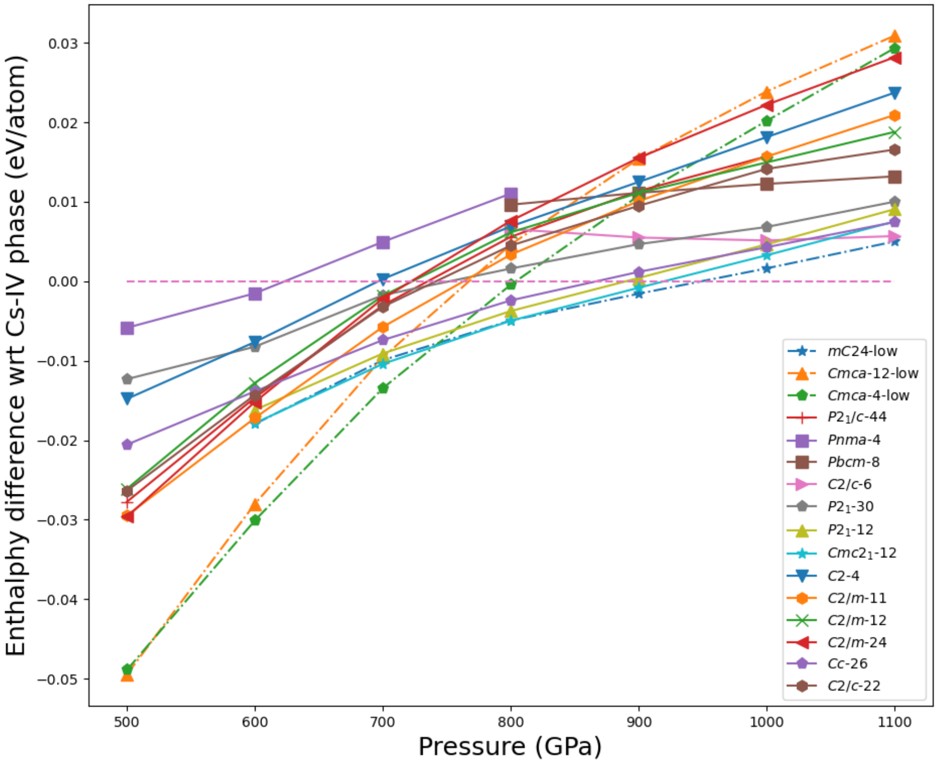

Figure 4.3: The relative static enthalpies per atom of found candidate structures from 800 GPa structure search with respect to Cs-IV phase. The static relative enthalpy of $mC24$-low, $Cmca$-4-low, and $Cmca$-12-low are also shown.

The structures with the lowest static enthalpy at 800 GPa are found to be in the $Cmc2_1$, $P2_1$, Cc, $C2/m$, $C2/c$, $C2$, $Pbcm$, and $Pnma$ space groups. The static enthalpy of the $Cmc2_1$-12 structure is nearly identical to that of $mC24$-low. At 700 GPa, it is only 0.5 meV/atom lower in static enthalpy than $mC24$-low, but becomes 0.8 meV/atom higher at 900 GPa. They both are also structurally similar, with some slight differences in the orientation of the H_2 molecules. The inclusion of $Cmc2_1$-12 structure slightly modifies the phase diagram of solid hydrogen, making it the most stable phase between $Cmca$-4-low and $mC24$-low phase with a region of stability from 740 GPa to 800 GPa.

By inspecting the results of 500 GPa and 800 GPa structure searches, it can be concluded that the structure searches conducted in this thesis were not exhaustive. For one thing, both structure searches failed to discover the $C2/c$-12 and $C2/c$-24 structures. Furthermore, the 500 GPa structure missed the most stable structure at 500 GPa, the $Cmca$-12-low, which was fortunately discovered by the 800 GPa search. The situation is reversed for the $Cmca$-4-low and $Cmca$-4-high, where they were not found by the 800 GPa search.

4.2.3 Structure Search at 1100 GPa to 1700 GPa

Unlike the previous two structure searches, the searches at 1100 GPa, 1400 GPa, and 1700 GPa yielded no structures with lower static enthalpy than Cs-IV. While it initially appeared as if those searches found some structures with static enthalpies below Cs-IV, those structures were found to be identical to Cs-IV upon inspection, up to a little difference in cell dimensions.

1100 GPa Structure Search

The static enthalpies of the found structures are shown in Fig 4.4. As can be seen on the graph, the lowest enthalpy found structure after Cs-IV is the $C2/c$-6 structure, whose relative static enthalpy with respect to Cs-IV is around 6 to 7 meV/atom. Such a value is very similar to the relative static enthalpy of $mC24$-high. This $C2/c$-6 structure is identical to the one found in the 800 GPa structure search. The rest of the most stable found structures are around 10 meV/atom to 20 meV/atom higher in static enthalpy than Cs-IV, belonging to $P2_1$, $C2$, $P2/c$, and $C2/c$ space groups.

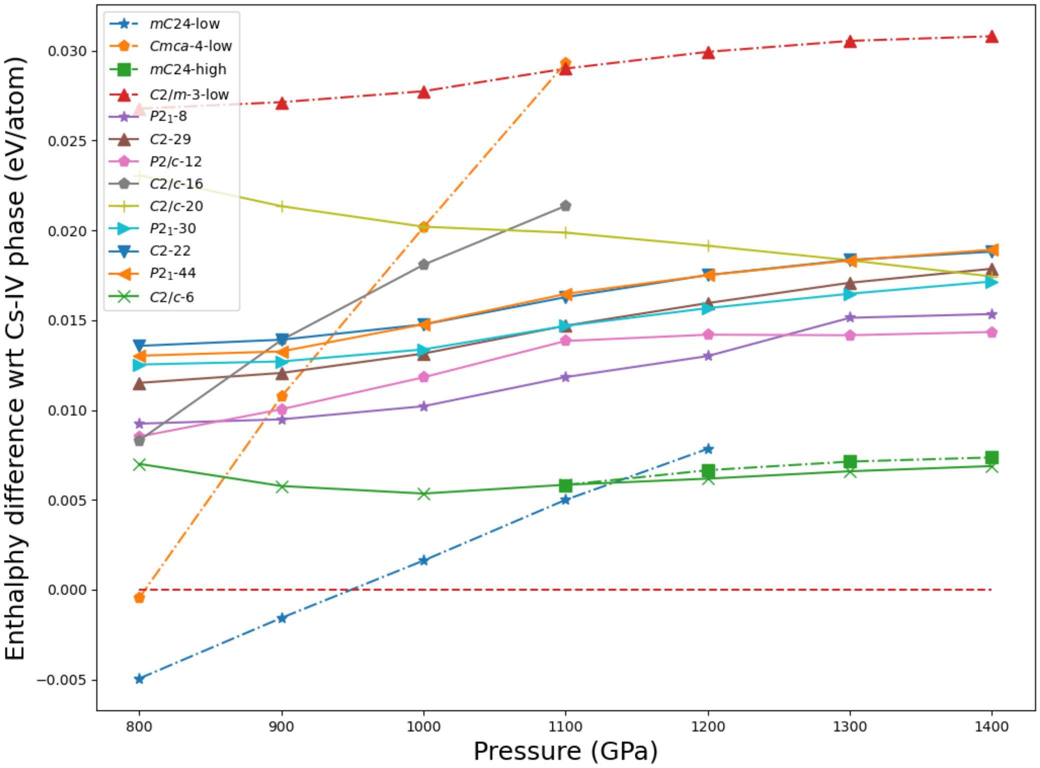

Figure 4.4: The relative static enthalpies per atom of found candidate structures from 1100 GPa structure search with respect to Cs-IV phase. The relative static enthalpy of $mC24$-low, $Cmca$-4-low, and $mC24$-high, and $C2/m$-3-low are also shown.

1400 GPa Structure Search

The 1400 GPa structure search rediscovered the $C2/c$-6 structure, identical to the one found in 1100 GPa and 800 GPa searches, as the most stable structure after Cs-IV. This search also found the $I422$-10 structure that was discovered by the 1700 GPa structure search. In the graph, the enthalpy plot of $I422$-10 is not shown as it is considerably more stable in higher pressure. However, the $C2/c$-6 structure is still shown as its static enthalpy value is relatively constant from 800 GPa to 1700 GPa. Other than those, the lowest-enthalpy found structures were in the $P2/c$, $P2_1$, $C2/c$ and $C2$ space groups. The plot of their enthalpies is shown in Fig 4.5.

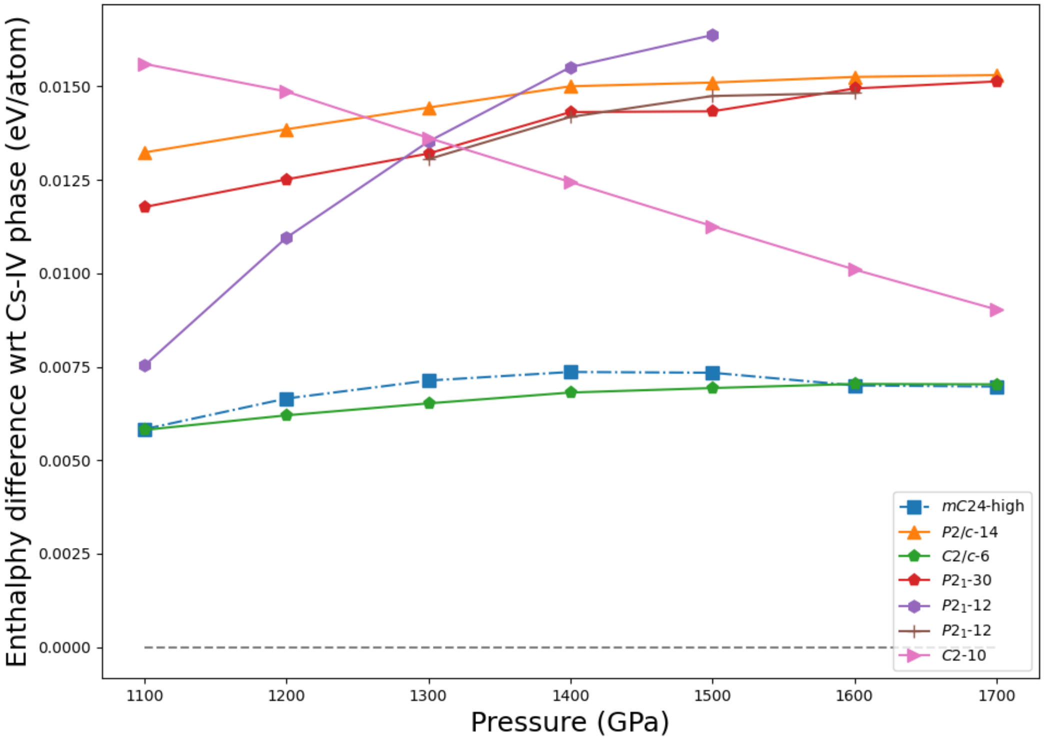

Figure 4.5: The relative static enthalpies per atom of found candidate structures from 1400 GPa structure search with respect to Cs-IV phase. The relative static enthalpy of $mC24$-high is also shown.

Structure Search at 1700 GPa

At 1700 GPa, we found structures with $P2_1/m$, $C2/c$, $P2_12_12_1$, $P2_1$, Cm, $I422$, $Pmmn$, $Cmcm$, $P2/c$, $Cmca$, $Pbca$, and $C2$ space groups among the low-enthalpy structures. All of those are higher in static enthalpy compared to our $mC24$-high structure. Only the $I422$-10 and $C2/c$-10 are very close in static enthalpy at 2000 GPa with $mC24$-high (Fig. 4.6). Above them, the next 3 lowest static enthalpy structures are the $P2_1/m$-10, $Pbca$-16, and $Cmcm$-6 structures, in ascending order.

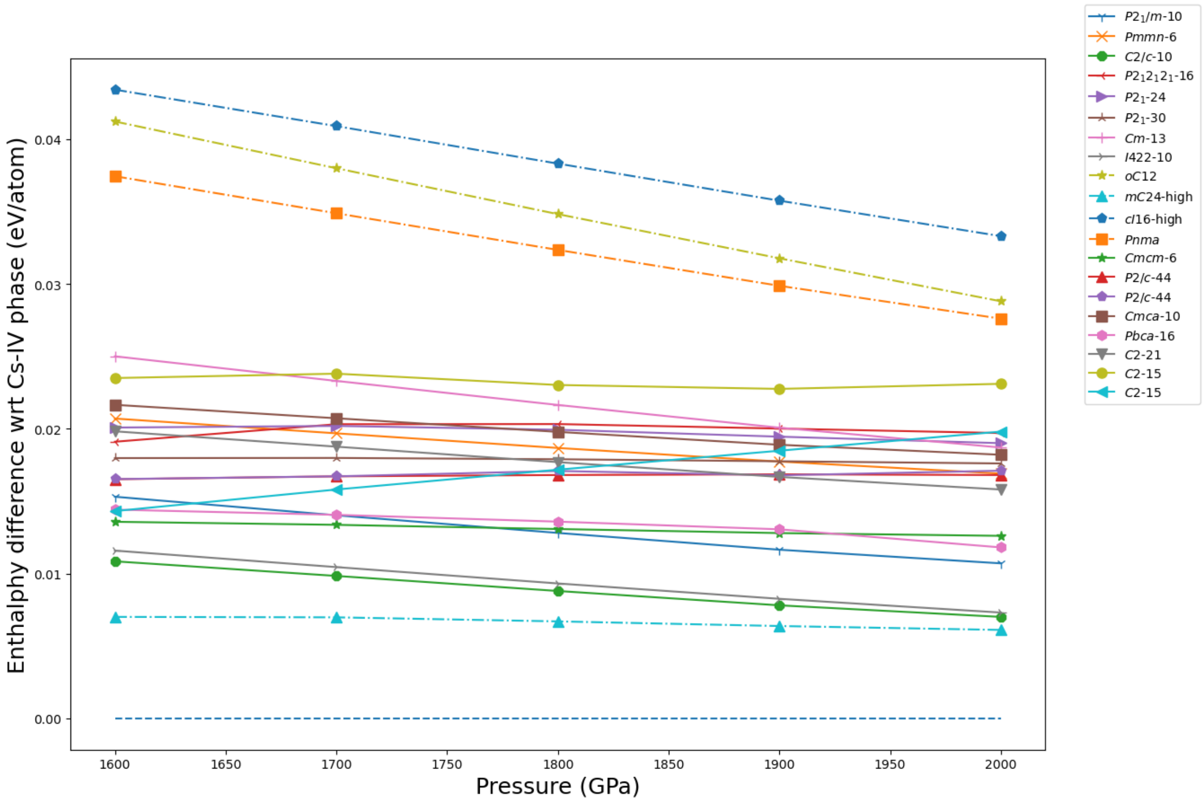

Figure 4.6: The relative static enthalpies per atom of found candidate structures from 1700 GPa structure search with respect to Cs-IV phase. The relative static enthalpy of $mC24$-high, $oC12$, $cI16$, and $Pnma$ are also shown.

4.3 Quantum Monte Carlo Static Enthalpy

4.3.1 For Known Candidate Structures

In Fig. 4.7 below, the DMC static enthalpies of the low-pressure structures are shown with respect to the $C2/c$-24 phase, meanwhile, those for the high-pressure structures are plotted with respect to the Cs-IV phase in Fig. 4.8. In both figures, the error bars are very small, with a maximum value of ±0.26 meV/atom. Hence, they are often obscured by the marker in the plot.

There were some notable differences compared to the DFT static enthalpy results. The phase transition order after $Cmca$-12-low still follows the $Cmca$-12-low → $Cmca$-4-low → $mC24$-low → Cs-IV ordering. However, the DMC calculations predict the $C2/c$-12 phase to be the most stable up to 534 GPa, placing it before the $Cmca$-12 phase. The stability region of $mC24$ also becomes narrower, ranging only from 810 GPa and 886 GPa. This puts the transition to the Cs-IV phase at about 70 GPa lower than the DFT results.

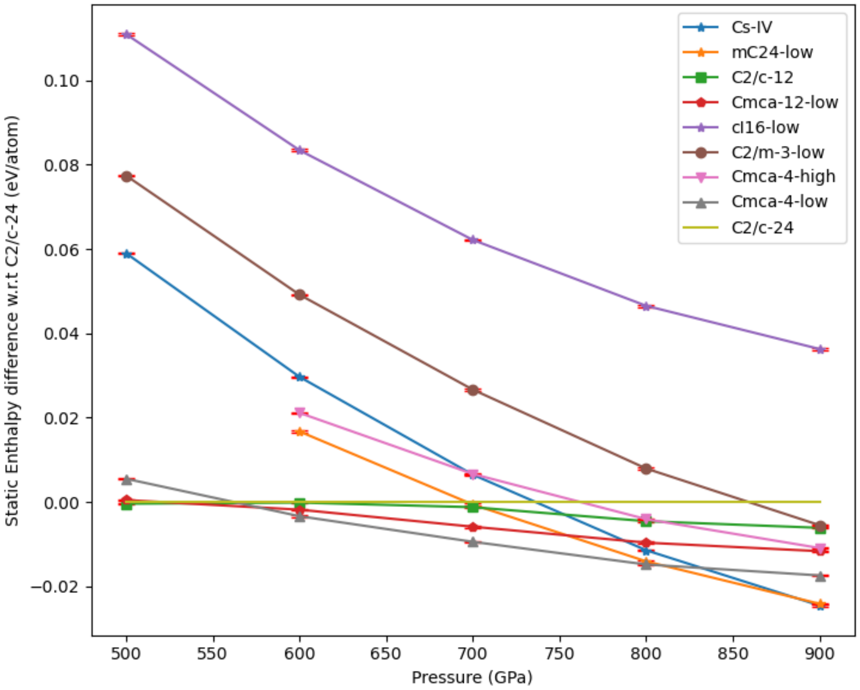

Figure 4.7: The relative DMC static enthalpies of known candidate structures within 500 GPa to 900 GPa, plotted with respect to that of $C2/c$-24 structure.

On the high-pressure structures side, the DMC static enthalpy results also possess some differences when compared to the DFT results. First, the relative static enthalpy of $mC24$-high and $C2/m$-3-low becomes larger and seems to increase with increasing pressure. Second, the gaps between the static enthalpy of known high-pressure structures seem to become larger than the DFT counterparts. At 2000 GPa, both DMC and DFT predicted that among those structures, the $C2/m$-3-high has the lowest relative static enthalpy and $cI16$ has the highest one. However the difference between those 2 structures is 14.5 meV/atom from DMC calculations, but only 6.7 meV/atom from DFT calculations. Furthermore, the next 3 lowest-enthalpy high-pressure known structures from DMC calculations are the $Pnma$, $R3m$, and $R\bar{3}m$ whose enthalpies lie 3.4 meV/atom, 4.5 meV/atom, 5.6 meV/atom higher than $C2/m$-3-high, respectively. Meanwhile, DFT predicted the next 3 most stable structures to be $R3m$, $Pnma$, and $oC12$, which have 0.3 meV/atom, 1 meV/atom, and 2.2 meV/atom higher enthalpy than $C2/m$-3-high. Hence, the DMC calculations made the $oC12$ structure less favorable compared to the DFT prediction.

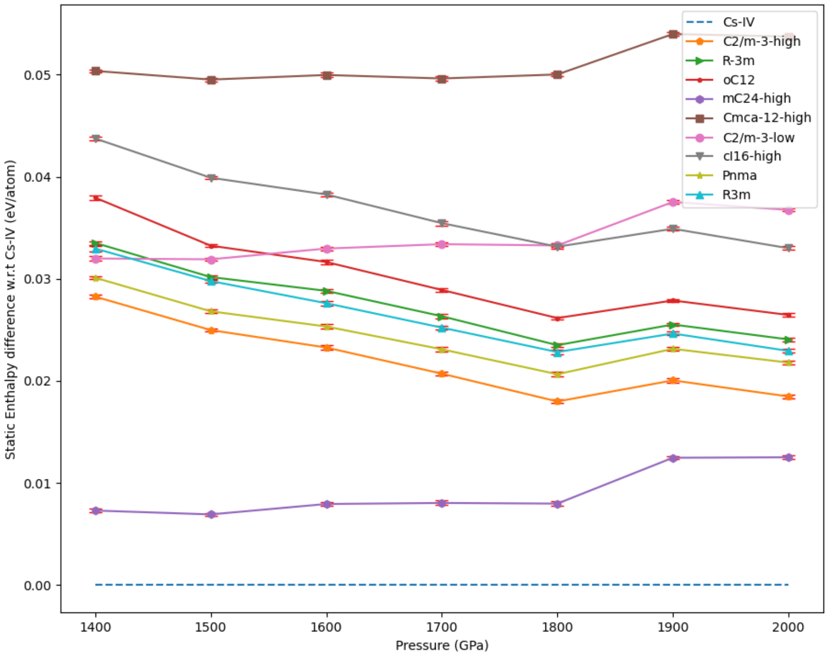

Figure 4.8: The relative DMC static enthalpies of known candidate structures within 1400 GPa to 2000 GPa plotted with respect to that of Cs-IV structure.

For the low-pressure known candidate structures, the DMC static enthalpy calculations in this work resulted in higher transition pressures compared to what was reported in the work of McMinis, et al [40]. Under the absence of $mC24$-low, they predicted the transitions to $Cmca$-12-low and Cs-IV phases to happen at 440 GPa and 684 GPa, respectively, without any $Cmca$-4-low phase between them. The pressures for the aforementioned transitions in our work are more consistent with the results of Monacelli, et al [42]. However, said study also did not predict any $Cmca$-12-low $\rightarrow$ $Cmca$-4-low transition, at least not below 650 GPa. Other than that, this work's DMC relative static enthalpy results for high-pressure known structures (Fig. 4.8) exhibit a "bump" at 1800 GPa to 1900 GPa.

This undesirable feature in the graphs and the discrepancy with previous DMC studies could be attributed to the difference in methods of estimating the pressure. This work used the pressure specified in DFT calculations to calculate the enthalpy. Meanwhile, the common approach is to use the pressure obtained from fitting the DMC energy, or to use the pressure estimated from VMC and DMC calculations.

4.3.2 For Found Candidate Structures

For found candidate structures, the DMC calculations were run with a smaller number of steps compared to the calculations for the known structures in order to save time. This increased the uncertainty of the results, with the error bars reaching up to around ± 1 meV/atom. Generally, the DMC static enthalpy graphs are rougher than their DFT counterparts due to the same reasons explained previously at the end of section 4.3.1.

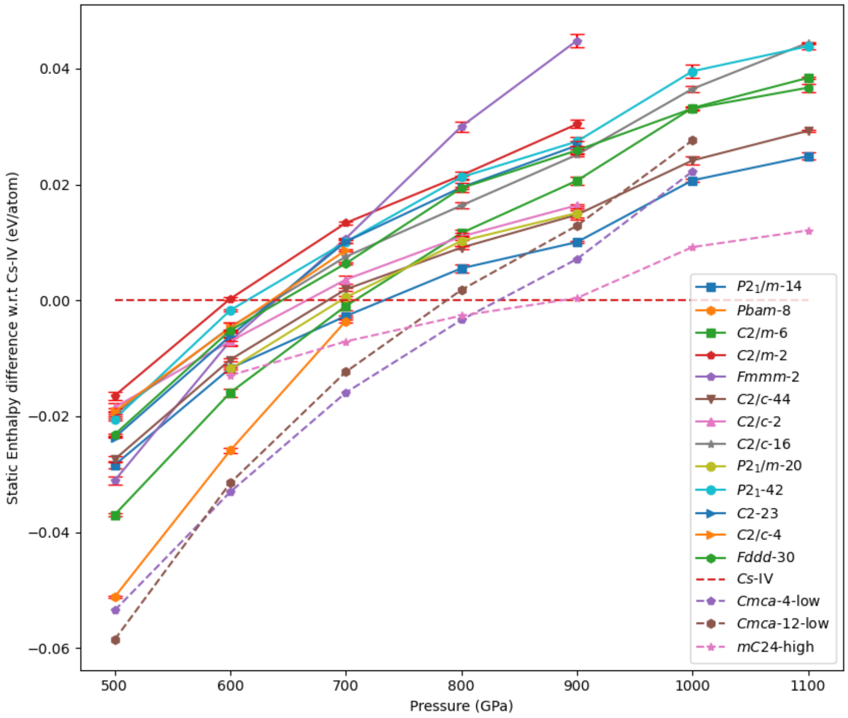

Figure 4.9: The relative DMC static enthalpies of found candidate structures from 500 GPa structure search plotted with respect to that of Cs-IV structure.

The results for 500 GPa and 800 GPa found structure are shown in Fig. 4.9 and Fig. 4.10, respectively. For the 500 GPa found structures, the DMC static enthalpy results did not change the phase diagram of hydrogen within 500 to 1100 GPa. The 5 lowest-enthalpy found structures at 500 GPa are the $Pbam$-8, $C2/m$-6, $Fmmm$-2, $P2_1/m$-14, and $C2/c$-44 structures, whose static enthalpies are within 51 meV/atom to 27 meV/atom lower than that of Cs-IV. These DMC results increased the enthalpy difference between $Pbam$-8 and $C2/m$-6 to 14 meV/atom from 4 meV/atom in the DFT results.

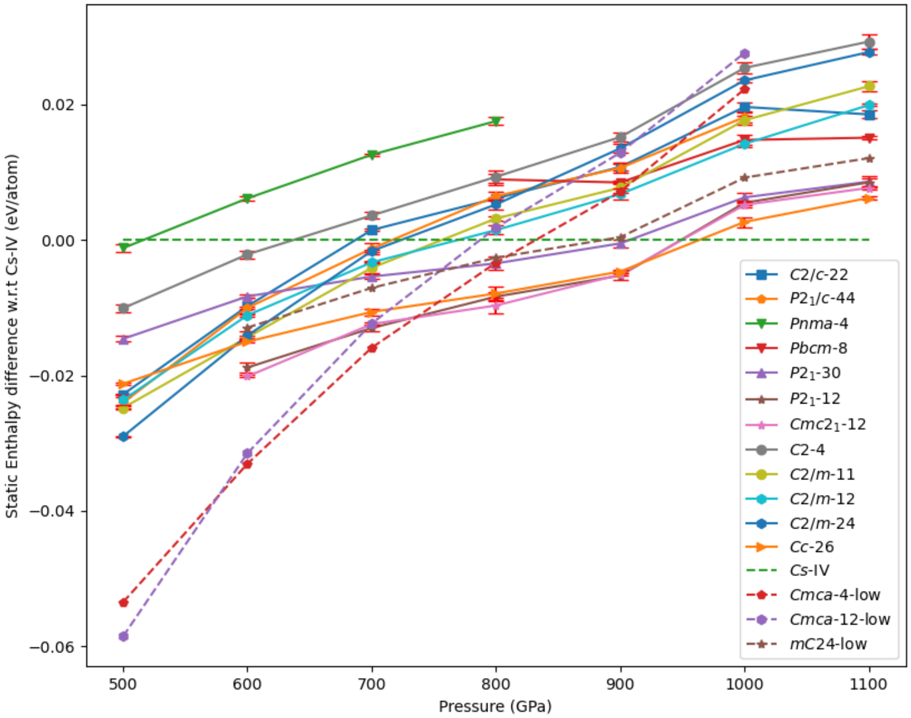

Figure 4.10: The relative DMC static enthalpies of found candidate structures from 800 GPa structure search plotted with respect to that of Cs-IV structure.

For the 800 GPa found structures, the $Cmc2_1$-12 structure remained to have the lowest static enthalpy among the found structures. However, the DMC calculation made its static enthalpy around 7 meV/atom below that of $mC24$-low at 800 GPa. The 5 lowest-enthalpy found structures at 800 GPa are the $Cmc2_1$-12, $P2_1$-12, Cc-26, $P2_1$-30, and $C2/m$-12. One of the found structures, the $C2/m$-24, managed to become as competitive as the lowest-enthalpy found structures from the 500 GPa structure search. At 500 GPa, its static enthalpy is 29 meV/atom below Cs-IV. Note that the result for $C2/c$-6 structure is not shown in Fig 4.10 since it is not stable below 800 GPa, and the pressure range above 800 GPa will be covered by the plot of the next structure searches.

The results for found structures from 1100 GPa, 1400 GPa, and 1700 GPa are shown in Fig 4.11, Fig 4.12, and Fig 4.13, respectively. From the 1100 GPa and 1400 GPa found structures, the only significant change from DFT results is that the DMC calculations made the $C2/c$-6 structure go below $mC24$-low and Cs-IV in static enthalpy at 800 GPa to 900 GPa.

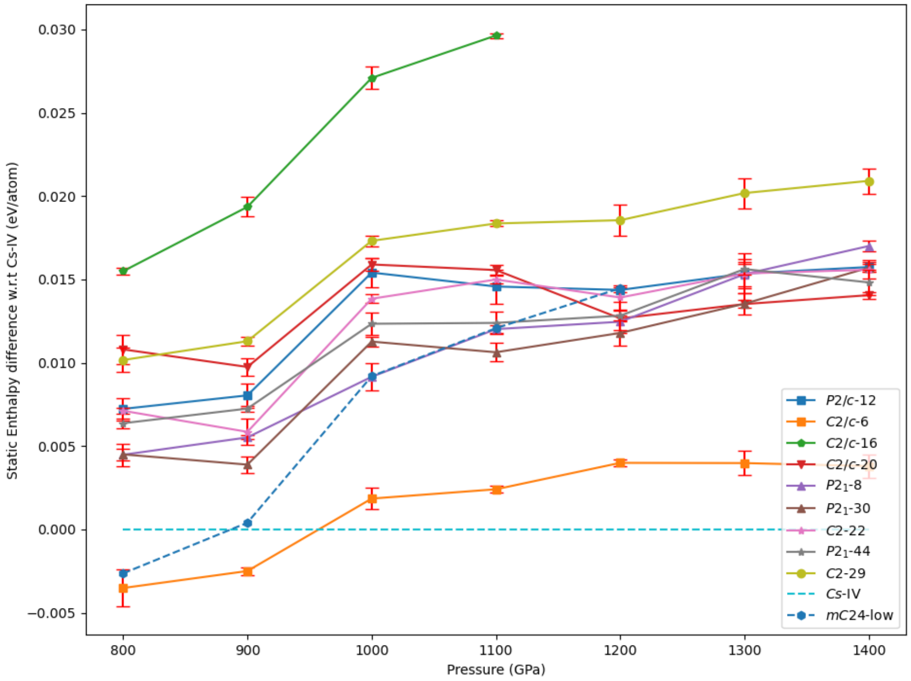

Figure 4.11: The relative DMC static enthalpies of found candidate structures from 1100 GPa structure search plotted with respect to that of Cs-IV structure.

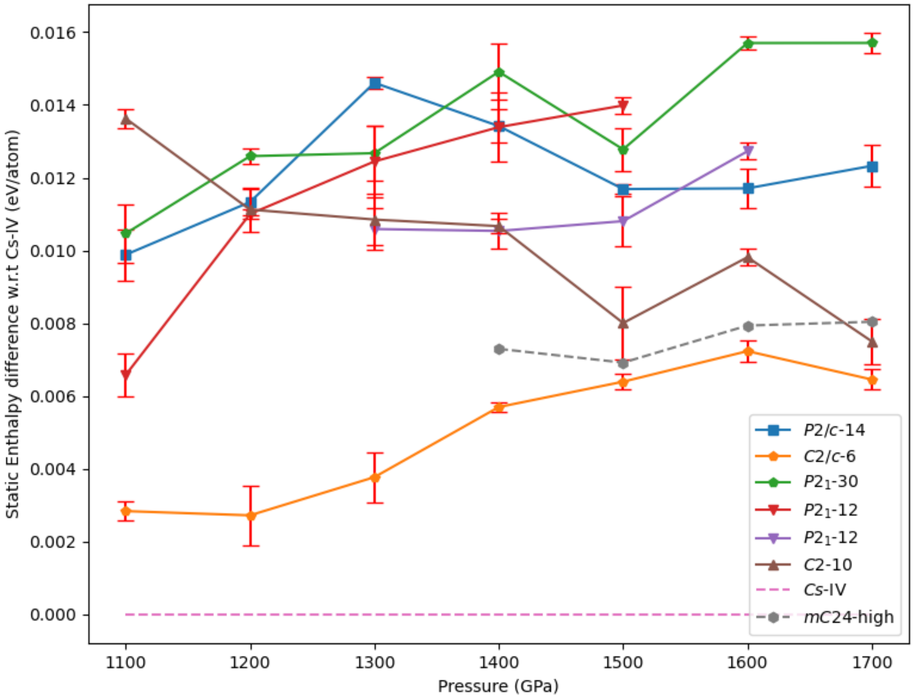

Figure 4.12: The relative DMC static enthalpies of found candidate structures from 1400 GPa structure search plotted with respect to that of Cs-IV structure.

For the 1700 GPa found structures, the DMC calculation made the $C2/c$-10 considerably more stable than $mC24$-high. The static enthalpy of $I422$-10, $P2_1/m$-10, and $Pbca$-16 were also lowered with respect to $mC24$-high and managed to become more stable than $mC24$-high at some point within 1700 GPa to 2000 GPa. These four found structures are also consistently the lowest-enthalpy structure within the pressure range of 1500 GPa to 2000 GPa.

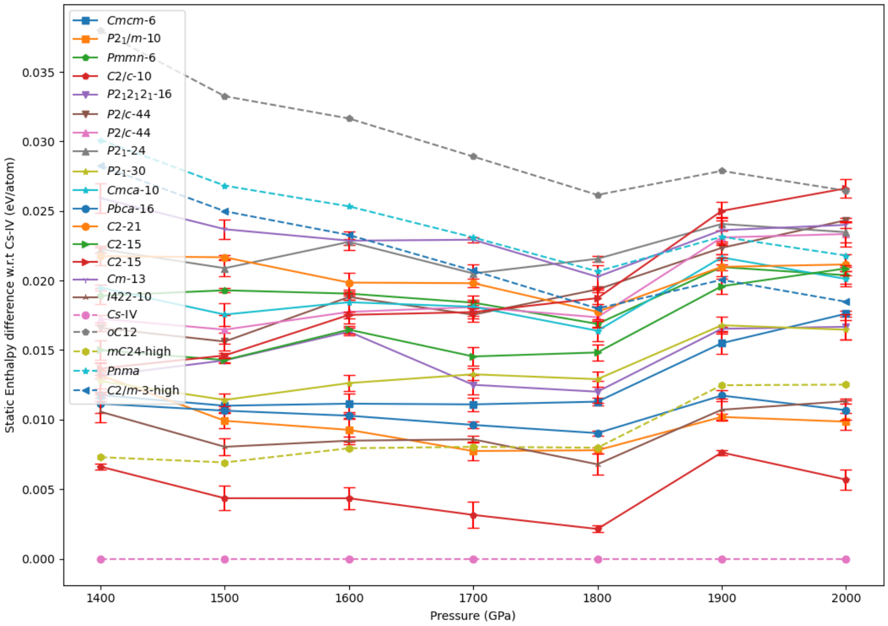

Figure 4.13: The relative DMC static enthalpies of found candidate structures from 1700 GPa structure search plotted with respect to that of Cs-IV structure.

4.4 Phonon Calculation for Known Candidate Structures

The phonon calculations revealed that most of the high-pressures known structures were dynamically unstable. Only two among those structures, the $mC24$-high and $Pnma$, were found to be dynamically stable. The $mC24$-high is stable throughout 1100 GPa to 2000 GPa, meanwhile, $Pnma$ is only stable at 2000 GPa, and starts to exhibit small imaginary phonon modes at 1900 GPa. All of the dynamically unstable high-pressure structures become more stable with increasing pressure, however, there were still some imaginary phonon modes left even at 2000 GPa. Hence, they will be suitable candidate structures in a pressure range beyond 2000 GPa. In Fig. 4.14, the phonon density of state of $R\bar{3}m$ structure is shown, showing how the structure gets more stable with increasing pressure.

From the low-pressure side, the $Cmca$-4-low, $Cmca$-12-low, and $C2/c$-24 were not evaluated by phonon calculations because they are known to be dynamically stable but

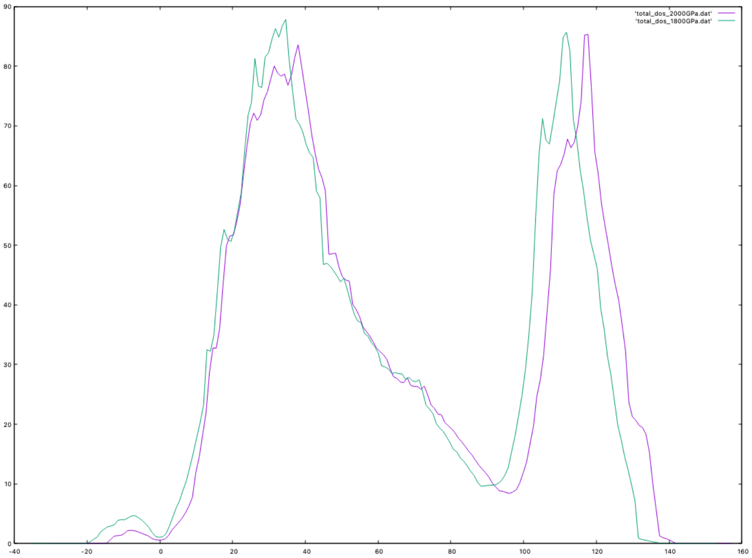

Figure 4.14: The phonon density of states of the optimized $R\bar{3}m$ structure at 1800 GPa (green line) and 2000 GPa (purple line). The graph shows a smaller magnitude of negative frequencies at higher pressure, indicating a more dynamically stable structure.

the inclusion of their ZPE made Cs-IV to be lower in enthalpy at 500 GPa and higher. Among the rest of the low-pressure structures, $cI16$-low, $Cmca$-4-high, and $C2/m$-3-low were found to be dynamically unstable. The former two structures become more stable with decreasing pressure, but the $C2/m$-3-low does not seem to become any less or more stable within 500 GPa to 2000 GPa. This left the $mC24$-low and $C2/c$-12 to be the only dynamically stable low-pressure known structures.

The plot of the relative dynamic enthalpy of $mC24$-low, $C2/c$-12, $mC24$-low, and $Pnma$ is shown in Fig. 4.15. Every single one of them except $Pnma$ becomes much less stable than Cs-IV after the inclusion of the zero-point energy. The $C2/c$-12 structure is very close in enthalpy with Cs-IV at 500 GPa. However, at such pressure, it is still less stable the H_2-PRE phase candidate structure [41], which is around 10 meV/atom lower than Cs-IV dynamic enthalpy. For the $Pnma$ structure, even though its DMC static enthalpy is 22 meV/atom above Cs-IV at 2000 GPa, the ZPE is 33 meV/atom lower which makes it more stable. Hence, there is a new phase transition from Cs-IV to $Pnma$ at somewhere between 1900 GPa and 2000 GPa. Seeing how the other high-pressure known structures are becoming more dynamically stable with increasing pressure, it would be interesting to investigate how their dynamic enthalpy competes with *Pnma* at a pressure regime beyond 2000 GPa. Similar to *Pnma*, the *R*3*m* and *oC*12 structures are also made up of triangular H_3 clusters and could have a ZPE as low as *Pnma* itself.

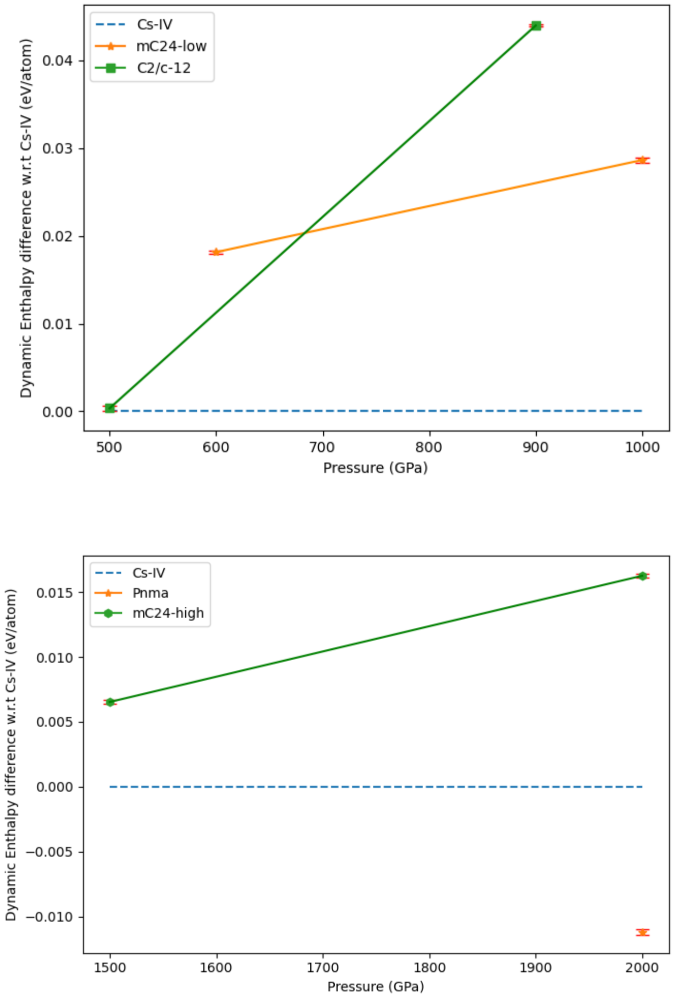

Figure 4.15: The relative dynamic enthalpies of dynamically stable known candidate structures plotted with respect to that of *Cs*-IV structure for the low-pressure side (top), and high-pressure side (bottom).

4.5 Phonon Calculation for Found Structures

For found structures, only a few of them have been evaluated with phonon calculations. The list is as follows:

1. Six lowest-enthalpy structures at 500 GPa, which consist of: $Pbam$-8, $C2/m$-6, $Fmmm$-2, $P2_1/m$-14, and $C2/c$-44 from 500 GPa structure search, plus the $C2/m$-24 structure from 800 GPa structure search.

2. Two lowest-enthalpy structures at 800 GPa which are the $Cmc2_1$-12 and $P2_1$-12 from 800 GPa structure search.

3. The $C2/c$-6 structure from 1100 GPa structure search. As mentioned before, this structure is the only one from 1100 GPa search that managed to become lower in DMC static enthalpy than $mC24$ and Cs-IV, and hence we prioritized its evaluation.

4. The $C2$-10 from 1400 GPa structure search. This structure was prioritized as its DMC static enthalpy graph has a downward trend with increasing pressure and becomes comparable with the lowest-enthalpy structures from 1700 GPa search.

5. Seven lowest-enthalpy structures from 1700 GPa structure search, which are $C2/c$-10, $I422$-10, $P2_1/m$-10, $Pbca$-16, $Cmcm$-6, $P2_12_12_1$-16, and $P2_1$-30.

Phonon calculations for categories (1), (2), and (3) in the list above were done at 2 pressure points within 500 GPa to 1000 GPa. For the chosen structures from the 500 GPa search, only two structures, the $P2_1/m$-14 and $C2/c$-44 were found to be dynamically stable, with the former being only stable at 1000 GPa. All the considered structures from 800 GPa and 1100 GPa search were dynamically stable except for $C2/m$-24 which was only stable at 500 GPa. The dynamical enthalpy for the dynamically stable structures is shown in Fig. 4.16. As can be seen in the figure, the $C2/c$-6 structure is more stable than Cs-IV in the pressure region of 800 GPa to 900 GPa.

The found structures that fall into categories (4) and (5) were evaluated with phonon calculation within 1500 GPa to 2000 GPa. Among those structures, only the $P2_1/m$-10 and the $P2_1$-30 were found to be dynamically unstable. The dynamic enthalpies of dynamically stable structures are shown in Fig. 4.17. The inclusion of zero-point energy made the $C2/c$-10 structure more stable than Cs-IV within 1500 GPa to 2000 GPa. At 1500 GPa, its dynamic enthalpy is 6 meV/atom lower than Cs-IV. To estimate at which

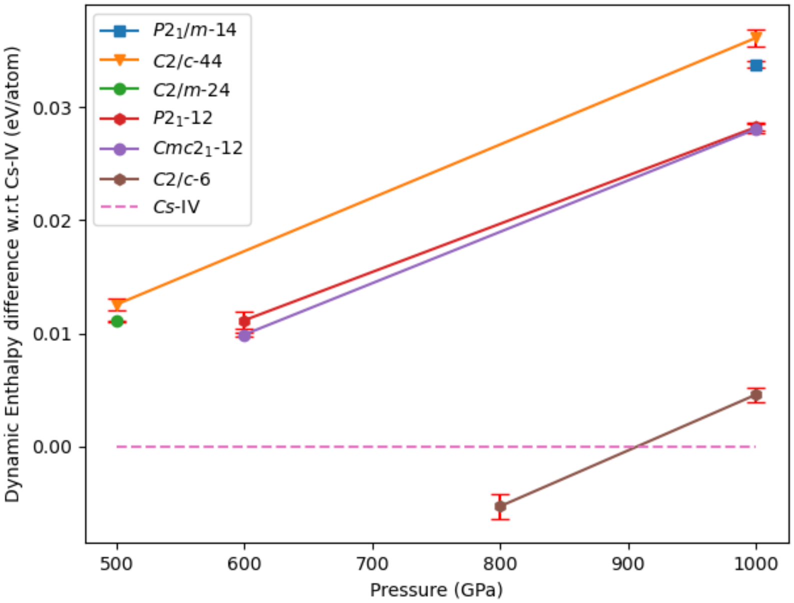

Figure 4.16: The relative DMC dynamic enthalpies of dynamically stable found structures from 500 GPa, 800 GPa, and 1100 GPa structure searches plotted with respect to that of Cs-IV structure.

pressure the Cs-IV transforms into $C2/c$-10, other ZPE evaluations were conducted at 1200 GPa and 1000 GPa. At 1200 GPa, $C2/c$-10 was still 3 meV/atom lower than Cs-IV, however, at 1000 GPa it has already started exhibiting a very small amount of imaginary modes. Hence, we concluded that the transition from Cs-IV $\rightarrow$ $C2/c$-10 happens at a pressure around 1000 GPa.

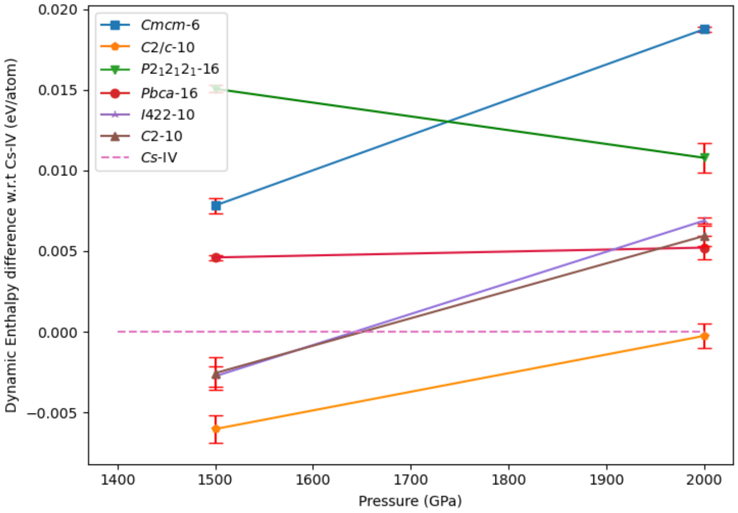

Figure 4.17: The relative DMC dynamic enthalpies of dynamically stable found structures from 1400 GPa and 1700 GPa structure search plotted with respect to that of Cs-IV structure.

Upon combining all the results in this thesis, the final phase diagram of solid hydrogen between 500 GPa to 2000 GPa was modified due to the newly discovered phase transitions. The transition sequence happens in the following order Cs-IV $\rightarrow$ $C2/c$-6 $\rightarrow$ Cs-IV $\rightarrow$ $C2/c$-10 $\rightarrow$ $Pnma$, which happen at 800 GPa, 900 GPa, 1000 GPa, and 2000 GPa. The resulting phase diagram is kind of unusual as after being replaced by $C2/c$-6, the Cs-IV phase becomes the most stable phase again at 900 GPa, only to be taken shortly after by $C2/c$-10 at 1000 GPa. The structure of $C2/c$-6 and $C2/c$-10 can be seen in Fig. 4.18

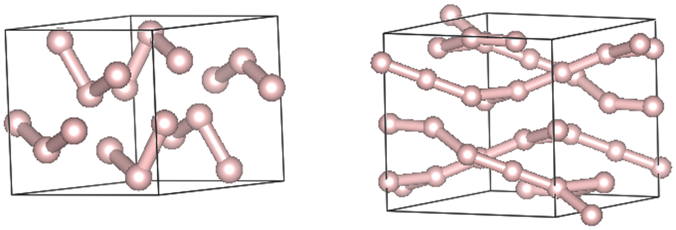

Figure 4.18: The $C2/c$-6 structure (left) and $C2/c$-10 structure (right). Fictitious bonds were drawn for clarity.

Chapter 5

Conclusion and Future Works

5.1 Conclusion

In summary, the phase diagram of solid hydrogen between 500 GPa and 2000 GPa was investigated by using the diffusion quantum Monte Carlo calculation for static enthalpy combined with the zero-point energy evaluated by utilizing harmonic approximation obtained with density functional theory force. An evolutionary crystal structure search was carried out at 500 GPa, 800 GPa, 1100 GPa, 1400 GPa, and 1700 GPa.

Among all the considered known candidate structures, only one of them, namely the $Pnma$ structure, was found to be more stable than Cs-IV at 2000 GPa. From the structure searches, two of the discovered candidate structures were found to be more stable than Cs-IV at some pressure regions within the considered pressure range. The inclusion of this work's results considerably altered the phase diagram of hydrogen. Within 500 GPa to 2000 GPa, hydrogen undergoes 4 phase transitions: Cs-IV $\rightarrow$ $C2/c$-6 $\rightarrow$ Cs-IV $\rightarrow$ $C2/c$-10 $\rightarrow$ $Pnma$, at 800 GPa, 900 GPa, 1000 Gpa, and 2000 GPa, respectively. The Cs-IV phase curiously becomes the most stable phase for the second time between 800 GPa to 900 GPa, after overtaking the short-lived $C2/c$-6 phase.

5.2 Future Works

In this section, we address the possible future study based on the results and challenges encountered in this thesis:

1. First of all, the straightforward extension to this study would be to consider a pressure regime beyond 2000 GPa. Previous DFT studies have considered a pressure as high as 3 to 5 TPa, which is far beyond the range considered in this thesis, but still nevertheless relevant in the context of astrophysics.
2. Investigating the physical properties such as the optical and thermal properties of the most promising candidate structures discovered in this work could also be considered for future studies.
3. The main bottleneck for this thesis comes from the phonon calculations and the DMC calculations. In the future, it is preferable to apply material informatics techniques such as machine learning to accelerate those calculations to significantly increase the number of structures that can be evaluated.
4. The machine learning techniques could also be used to accelerate the structural search. For example, the DFT structural optimization can be replaced by an accurate neural network potential to allow a more thorough exploration of candidate structures. The fact that only a single element needs to be considered should decrease the complexity of designing the neural network.

Bibliography

[1] José A. Flores-Livas, Lilia Boeri, Antonio Sanna, Gianni Profeta, Ryotaro Arita, and Mikhail Eremets. A perspective on conventional high-temperature superconductors at high pressure: Methods and materials. *Phys. Rep.*, 856:1–78, apr 2020.

[2] Ranga P Dias, Ori Noked, and Isaac F Silvera. Quantum phase transition in solid hydrogen at high pressure. *Physical Review B*, 100(18):184112, 2019.

[3] E. Wigner and H. B. Huntington. On the possibility of a metallic modification of hydrogen. *The Journal of Chemical Physics*, 3(12):764–770, dec 1935.

[4] N. W. Ashcroft. Metallic hydrogen: A high-temperature superconductor? *Phys. Rev. Lett.*, 21(26):1748–1749, dec 1968.

[5] Ravit Helled, Guglielmo Mazzola, and Ronald Redmer. Understanding dense hydrogen at planetary conditions. *Nature Reviews Physics*, 2(10):562–574, 2020.

[6] Burkhard Militzer, François Soubiran, Sean M Wahl, and William Hubbard. Understanding jupiter's interior. *Journal of Geophysical Research: Planets*, 121(9):1552–1572, 2016.

[7] Bing Li, Cheng Ji, Wenge Yang, Junyue Wang, Ke Yang, Ruqing Xu, Wenjun Liu, Zhonghou Cai, Jiuhua Chen, and Ho-kwang Mao. Diamond anvil cell behavior up to 4 mbar. *Proceedings of the National Academy of Sciences*, 115(8):1713–1717, 2018.

[8] Paul Loubeyre, Agnes Dewaele, Florent Occelli, Olivier Marie, and Mohamed Mezouar. The toroidal diamond anvil cell for detailed measurements under extreme

static pressures. In *APS March Meeting Abstracts*, volume 2019, pages E17–008, 2019.

[9] P. Loubeyre, R. LeToullec, D. Hausermann, M. Hanfland, R. J. Hemley, H. K. Mao, and L. W. Finger. X-ray diffraction and equation of state of hydrogen at megabar pressures. *Nature*, 383(6602):702–704, oct 1996.

[10] Chris J. Pickard and Richard J. Needs. Structure of phase III of solid hydrogen. *Nat. Phys.*, 3(7):473–476, may 2007.

[11] Paul Loubeyre, Florent Occelli, and Paul Dumas. Synchrotron infrared spectroscopic evidence of the probable transition to metal hydrogen. *Nature*, 577(7792):631–635, 2020.

[12] Ranga P. Dias and Isaac F. Silvera. Observation of the wigner-huntington transition to metallic hydrogen. *Science*, 355(6326):715–718, feb 2017.

[13] Kurt A Johnson and NW Ashcroft. Structure and bandgap closure in dense hydrogen. *Nature*, 403(6770):632–635, 2000.

[14] Jeffrey M McMahon and David M Ceperley. Ground-state structures of atomic metallic hydrogen. *Physical Review Letters*, 106(16):165302, 2011.

[15] Hanyu Liu, Hui Wang, and Yanming Ma. Quasi-molecular and atomic phases of dense solid hydrogen. *The Journal of Physical Chemistry C*, 116(16):9221–9226, 2012.

[16] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136(3B):B864–B871, nov 1964.

[17] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140(4A):A1133–A1138, nov 1965.

[18] John P. Perdew and Yue Wang. Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B*, 45:13244–13249, Jun 1992.

[19] John P. Perdew, Stefan Kurth, Ale š Zupan, and Peter Blaha. Accurate density functional with correct formal properties: A step beyond the generalized gradient approximation. *Phys. Rev. Lett.*, 82:2544–2547, Mar 1999.

[20] Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.

[21] Peter J Reynolds, David M Ceperley, Berni J Alder, and William A Lester Jr. Fixed-node quantum monte carlo for moleculesa) b). *The Journal of Chemical Physics*, 77(11):5593–5603, 1982.

[22] Artem R Oganov and Colin W Glass. Crystal structure prediction using ab initio evolutionary techniques: Principles and applications. *The Journal of chemical physics*, 124(24), 2006.

[23] Andriy O. Lyakhov, Artem R. Oganov, Harold T. Stokes, and Qiang Zhu. New developments in evolutionary structure prediction algorithm USPEX. *Comput. Phys. Commun.*, 184(4):1172–1182, apr 2013.

[24] Artem R Oganov, Andriy O Lyakhov, and Mario Valle. How evolutionary crystal structure prediction works—and why. *Accounts of chemical research*, 44(3):227–237, 2011.

[25] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77(18):3865–3868, oct 1996.

[26] G. Kresse and J. Hafner. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B*, 47(1):558–561, jan 1993.

[27] G. Kresse and J. Hafner. Ab initio molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium. *Phys. Rev. B*, 49(20):14251–14269, may 1994.

[28] G. Kresse and J. Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B*, 54(16):11169–11186, oct 1996.

[29] G. Kresse and J. Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science*, 6(1):15–50, jul 1996.

[30] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist. Van der waals density functional for general geometries. *Phys. Rev. Lett.*, 92:246401, Jun 2004.

[31] Guillermo Román-Pérez and José M. Soler. Efficient implementation of a van der waals density functional: Application to double-wall carbon nanotubes. *Phys. Rev. Lett.*, 103:096102, Aug 2009.

[32] Ji ří Klimeš, David R. Bowler, and Angelos Michaelides. Van der waals density functionals applied to solids. *Phys. Rev. B*, 83:195131, May 2011.

[33] Jeongnim Kim, Andrew D Baczewski, Todd D Beaudet, Anouar Benali, M Chandler Bennett, Mark A Berrill, Nick S Blunt, Edgar Josué Landinez Borda, Michele Casula, David M Ceperley, et al. Qmcpack: an open source ab initio quantum monte carlo package for the electronic structure of atoms, molecules and solids. *Journal of Physics: Condensed Matter*, 30(19):195901, 2018.

[34] Paul RC Kent, Abdulgani Annaberdiyev, Anouar Benali, M Chandler Bennett, Edgar Josué Landinez Borda, Peter Doak, Hongxia Hao, Kenneth D Jordan, Jaron T Krogel, Ilkka Kylänpää, et al. Qmcpack: Advances in the development, efficiency, and application of auxiliary field and real-space variational and diffusion quantum monte carlo. *The Journal of chemical physics*, 152(17), 2020.

[35] Jaron T Krogel. Nexus: A modular workflow management system for quantum simulation codes. *Computer Physics Communications*, 198:154–168, 2016.

[36] Paolo Giannozzi, Stefano Baroni, Nicola Bonini, Matteo Calandra, Roberto Car, Carlo Cavazzoni, Davide Ceresoli, Guido L Chiarotti, Matteo Cococcioni, Ismaila Dabo, Andrea Dal Corso, Stefano de Gironcoli, Stefano Fabris, Guido Fratesi, Ralph Gebauer, Uwe Gerstmann, Christos Gougoussis, Anton Kokalj, Michele Lazzeri, Layla Martin-Samos, Nicola Marzari, Francesco Mauri, Riccardo Mazzarello, Stefano Paolini, Alfredo Pasquarello, Lorenzo Paulatto, Carlo Sbraccia, Sandro Scandolo, Gabriele Sclauzero, Ari P Seitsonen, Alexander Smogunov, Paolo Umari, and Renata M Wentzcovitch. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys.: Condens. Matter*, 21(39):395502, sep 2009.

[37] P Giannozzi, O Andreussi, T Brumme, O Bunau, M Buongiorno Nardelli, M Calandra, R Car, C Cavazzoni, D Ceresoli, M Cococcioni, N Colonna, I Carnimeo, A Dal Corso, S de Gironcoli, P Delugas, R A DiStasio, A Ferretti, A Floris, G Fratesi, G Fugallo, R Gebauer, U Gerstmann, F Giustino, T Gorni, J Jia, M Kawamura, H-Y Ko, A Kokalj, E Küçükbenli, M Lazzeri, M Marsili, N Marzari, F Mauri, N L Nguyen, H-V Nguyen, A Otero de-la Roza, L Paulatto, S Poncé, D Rocca, R Sabatini, B Santra, M Schlipf, A P Seitsonen, A Smogunov, I Timrov, T Thonhauser, P Umari, N Vast, X Wu, and S Baroni. Advanced capabilities for materials modelling with quantum ESPRESSO. *J. Phys.: Condens. Matter*, 29(46):465901, oct 2017.

[38] Atsushi Togo, Laurent Chaput, Terumasa Tadano, and Isao Tanaka. Implementation strategies in phonopy and phono3py. *J. Phys. Condens. Matter*, 35(35):353001, 2023.

[39] Atsushi Togo. First-principles phonon calculations with phonopy and phono3py. *J. Phys. Soc. Jpn.*, 92(1):012001, 2023.

[40] Jeremy McMinis, Raymond C. Clay, Donghwa Lee, and Miguel A. Morales. Molecular to atomic phase transition in hydrogen under high pressure. *Phys. Rev. Lett.*, 114(10):105305, mar 2015.

[41] Tom Ichibha, Yunwei Zhang, Kenta Hongo, Ryo Maezono, and Fernando A Reboredo.

Candidate structure for the h 2-pre phase of solid hydrogen. *Physical Review B*, 104(21):214111, 2021.

[42] Lorenzo Monacelli, Michele Casula, Kousuke Nakano, Sandro Sorella, and Francesco Mauri. Quantum phase diagram of high-pressure hydrogen. *Nature Physics*, 19(6):845–850, 2023.