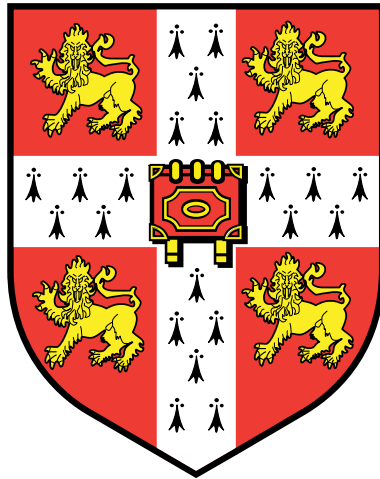# Novel methods to predict solid-state material properties

Michael John Hutcheon

Theory of Condensed Matter

Hughes Hall, University of Cambridge

This thesis is submitted for the degree of

Doctor of Philosophy

September 2021

To my parents,
for looking after so many people, including me

Acknowledgements

I am grateful for being able to work with and around some wonderful people in my time at TCM. I am grateful to my supervisor, Prof. Richard Needs, for his enthusiasm and for the academic freedom that he encouraged. I am also grateful to Prof. Chris Pickard his support and guidance during our work on superconductivity - I learned a great deal about science, not just physics, from Chris. I have also been lucky enough to receive advice from Dr Bartomeu Monserrat, Dr Joseph Nelson and Dr. Pablo López Ríos who collectively provided expertise across a wide variety of topics. I would also like to thank Prof. Mike Payne and Prof. Nigel Cooper for their support of myself and of the TCM group as a whole. I am grateful to Dr. Michael Rutter for writing software that has saved me a great deal of time and for his entertaining tales from the world of IT.

Several PhD students have helped me immensely during my studies. My office mates, Alex Wade and Nick Woods, provided excellent company and a responsive surface from which to bounce ideas. I am grateful to Mark Johnson for sharing his expert knowledge of phonons and for guiding me through the dark forest of FORTRAN. Of all people at TCM, I worked particularly closely with Alice Shipley - working with her has been the highlight of my PhD, not just because of her intelligence and work ethic, but also her friendliness and empathy.

Finally, thanks to my parents and to Helen, for a million things.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Preface. This dissertation contains fewer than 60,000 words including abstract, tables, footnotes and appendices, but excluding table of contents, photographs, diagrams, figure captions, list of figures, list of abbreviations, bibliography and acknowledgements.

Abstract

Solid-state materials find ubiquitous use in modern technology - from semiconductors in electronics to steel in buildings and superconductors in MRI machines. Theoretical understanding of the atomic-scale behavior of these materials can be leveraged to design new materials with desirable properties. In this thesis, we investigate the challenges that arise when this is attempted in practice.

Accurate and inexpensive methods to tackle the atomic-scale problem are a prerequisite for materials discovery. We begin with a description of existing methods. This is followed by the development of a Monte Carlo method to calculate expectation values from the many-body picture without the need for a trial wavefunction, which is both a fundamental, and practical, limitation in existing techniques.

Having explored first-principles methods, we turn to their use in understanding materials, beginning with an investigation of the structure of Lithium. Structure searching calculations result in a mixed-phase model at low temperatures, in good agreement with previous experimental and theoretical results. The quasi-harmonic treatment of finite-temperature thermodynamics is extended to include anharmonic nuclear vibrations, which are shown to not alter the phase diagram despite the small mass of the Li atoms.

Focus then shifts towards leveraging these same methods to discover novel superconductors. This begins with an investigation of the LaH_{10} and YH_{10} compounds, where a new hexagonal phase of LaH_{10} provides an explanation for recent experimental measurements. Machine-learning techniques and novel screening methods are then employed to discover hydrides of Rb and Cs that exhibit superconductivity at significantly lower pressures than LaH_{10}. Optimizations to, and automation of, the workflow then enables the discovery of superconductors on an unprecedented scale, leading to hundreds of new high-temperature superconductors.

Throughout the thesis, the importance of structures that are saddle-points of the energy landscape becomes apparent. The thesis closes with the development of a new algorithm to locate saddle-points that requires no additional information beyond that used by the cheapest existing methods.

This thesis demonstrates that there is progress to be made at every stage of the first-principles materials discovery process and highlights that improving the workflow itself is a non-trivial, but fruitful, pursuit.

Contents

List of Figures

List of Tables

Preface

The work presented in chapter 3 has been published as *"Stochastic nodal surfaces in quantum Monte Carlo calculations"*, Michael Hutcheon, Phys. Rev. E **102**, 042105 (2020) and is my own work.

Large parts of chapter 4 have been published as *"Structural and vibrational properties of lithium under ambient conditions within density functional theory"*, Michael Hutcheon and Richard Needs, Phys. Rev. B **99**, 014111 (2019). I performed all of the calculations and analysis included in this work, with guidance on direction from Prof. Richard Needs.

Chapter 5 has been published as *"Stability and superconductivity of lanthanum and yttrium decahydrides"*, Alice M. Shipley, Michael J. Hutcheon, Mark S. Johnson, Richard J. Needs, and Chris J. Pickard, Phys. Rev. B 101, 224511 (2020). CJP performed structure searching for the La-H system, confirmed by AMS. AMS performed pseudopotential testing and structure searching in the Y-H system and investigated the impact of various approximations. I performed phonon and electron-phonon calculations to derive the Gibbs free energies and critical temperatures of low-energy candidates from the structure searching and investigated the importance of the double-delta smearing parameter on the results. MSJ performed convergence testing.

Chapter 6 has been published as *"Predicting novel superconducting hydrides using machine learning approaches"*, Michael J. Hutcheon, Alice M. Shipley, and Richard J. Needs, Phys. Rev. B **101**, 144505 (2020). AMS and I designed the project, chose model inputs and built the literature dataset. AMS performed structure searching, constructed enthalpy plots and tested two T_c ranking methods based on Gaspari-Gyorffy theory, which she implemented in an existing code (with advice from Dr Po-Hao Chang via email correspondence). I trained and validated the machine learning model, extracted predictions and carried out the final DFPT calculations of T_c.

Most of the work in chapter 7 has been submitted as *"High-throughput discovery of high-temperature conventional superconductors"*, A. M. Shipley, M. J. Hutcheon, R. J. Needs, and C. J. Pickard (with equal contributions from the first two authors). The manuscript is available on the ArXiv (with identifier arXiv:2105.02296). CJP

performed the structure searching. AMS performed the subsequent screening, including calculations of electron-phonon parameters from Gaspari-Gyorffy and constrcuted, trained and validated the T_c regression model. I optimized the existing electron-phonon code, implemented a high-throughput electron-phonon library and calculated critical temperatures for the best candidates and for training the model.

The work in chapter 8 is my own and unpublished.

Chapter 1

Introduction

Theoretical models allow us to explain, and to predict, physical phenomena. Typically, this is also the order in which they are used, whereby models are first used to understand previously unexplained experimental measurements. Once a model is capable of explanation, it may seem a trivial task to extend that model to the task of prediction; surely we can simply apply the model to an otherwise-unstudied system to predict its properties? Unfortunately, this is not the case - an explanatory model is a necessary, but not a sufficient, condition for predictive power.

In this thesis we will focus on the problem of predicting the properties of solid-state materials. We will do this by propagating the quantum-mechanical description of atomic-scale structure to macroscopic-scale properties such as mechanical stability, phase behaviour and superconductivity. We will explore the challenges that arise at several stages of the predictive process, from the computational challenge of solving the microscopic models and the conceptual challenge of understanding those solutions to the wider problem of choosing which of the combinatorially-many systems to study in the first place.

We will employ an atomistic model of materials - describing them as a collection of interacting electrons and nuclei. A vast amount of effort has been put into solving the resulting quantum-mechanical Schrödinger equation, a partial differential equation in dN variables (where N is the number of particles and d the dimension of the space they move in). Because of the complexity of this equation (especially in the many-body $N \gg 1$ limit), numerical, rather than analytic, solutions are often sought. However, even with the exponentially increasing computer power available to academics, the many-body problem is still intractable for all but the smallest systems. Therefore, much of the effort involved in describing materials via quantum mechanics has been focused on the development of approximate solution methods. In this thesis we will investigate and employ two families of solution methods. The first, *quantum Monte Carlo* (QMC) methods, can be traced back to post-war ef-

forts initiated by John Von Neumann to further the understanding of nuclear fission via the stochastic methods pioneered by people such as Nick Metropolis [194] and leveraging the historic ENIAC programmable computer. The method was later developed into a general-purpose solution method for quantum-mechanical problems over the course of many years, culminating in the *Greens function Monte Carlo* of Malvin Kalos [133], adapted for fermionic systems by David Ceperley [42]. The method is closely related with the formulation by Feynman of quantum mechanics as a *path integral* problem; a point of view that we will take in chapter 3 when investigating improvements to the treatment of fermionic systems. QMC methods make the many-body problem tractable by never explicitly storing or operating on the (extremely) large quantum-mechanical state $|\psi\rangle$, but instead statistically sampling only the most important system configurations (or, equivalently, by only evaluating the most important paths in the path-integral).

The second family of solution methods we will concern ourselves with are *density functional theory* (DFT) methods. Where QMC methods reduced the many-body problem by approximating the quantum-mechanical state $|\psi\rangle$, DFT methods instead approximate the quantum-mechanical system itself by approximating the Hamiltonian, H. Influenced by the Thomas-Fermi model of a system of electrons [261, 86], which is based on the electron *density* $n(r)$, Hohenburg and Kohn [114] demonstrated that the energy of a many-electron system was *uniquely* specified by its density. This realisation allowed for the construction of an effective independent-electron theory by Kohn and Sham [142] - a significantly easier quantum-mechanical problem, allowing the explicit construction of the quantum mechanical state. Explicit access to the quantum state has allowed DFT to be extended to the treatment of electronic and quasiparticle (e.g phonon) exitations and even higher-order effects such as the electron-phonon coupling responsible for conventional superconductivity. Access to this extended information allows direct comparison with experiment and, as a result, DFT has become a cornerstone of quantum chemistry. We will explore this interplay between experiment and DFT results at several stages of this thesis. In particular, we will see that the relative ease at which one can perform structural relaxations using DFT will prove complimentary to crystal structure determination via x-ray crystallography - especially when only partial experimental information is available. Indeed we will also see how this process allows for fully *in silico* crystal structure prediction (CSP) without the need for any experimental data. Both of these modes of operation will be apparent in a study of the crystal structure of Lithium presented in chapter 4. Throughout this thesis, we will see that interesting structures correspond not only to minima of potential energy surfaces, but also to saddle points - we investigate algorithms to identify this additional class of structures in chapter 8.

Once one has access to a crystal structure, DFT allows the calculation of many of the properties exhibited by a macroscopic sample of that crystal. Thanks to developments in density functional perturbation theory (DFPT), this routinely includes properties that are mediated by phonons and electron-phonon coupling. The latter of these has become particularly relevant in recent years thanks to renewed interest in conventional superconductivity. This new wave of interest can be traced back to Ashcroft's 1968 prediction of high-temperature superconductivity in metallic hydrogen [17]. Whilst this prediction is widely believed to be accurate, extreme pressures (~ 500 GPa [55]), naturally occurring only deep inside the gas giants, are required to stabilize the metallic phase. Ashcroft went on to suggest that, in order to reduce the required pressure, one could compress the hydrogen via chemical means with the addition of heavy *pre-compressor* elements [15], resulting in a metallic Hydride. Thanks to advances in high-pressure experimental techniques, in particular improved diamond anvil cells, several hydrides have been stabilized in the laboratory. The 2015 measurement of superconductivity at 203 K in SH_3 at 200 GPa [59] (as predicted theoretically [65]) sparked a flurry of interest and quickly lead to even higher temperatures with the measurement of superconductivity at 250 K in LaH_{10} at 170 GPa [61]. We investigate the structural properties of LaH_{10} and the even-higher temperature YH_{10} superconductor in chapter 5 alongside an investigation of numerical sensitivities in the theoretical treatment. We also weigh in on the origin of structural defects in samples of LaH_{10}, which are difficult to characterise with x-ray diffraction because of the small hydrogen cross-section.

Coupled with the large number of possible Hydride compositions, and the ability to determine their structure and superconducting properties *in silico*, the study of high-pressure Hydrides has grown into an extensive field of research [66, 304, 87, 35, 206, 203, 218], culminating in the first measurements of superconductivity above room-temperature in a ternary C-S-H system [248] (albiet at the high pressure of ~ 270 GPa). Looking ahead, the challenge now is to reduce the pressure at which this family of superconductors operate. Because of the large range of compositions and difficulty of the relevant experiments, it is desirable to narrow the search computationally using CSP and DFT. We supplement these techniques with a machine-learning based screening approach in chapter 6, leading to the prediction of high-temperature Rb and Cs polyhydrides at only 50 GPa. We further develop these techniques in chapter 7 to enable a high-throughput search for superconductors on an unprecedented scale. The results elucidate general trends in high-temperature superconductors, where to look for them and the prospects for ambient-conditions superconductivity as well as identifying a large number of new high-T_c superconductors.

Chapter 2

Background theory

2.1 The many-body problem

As one approaches the atomic scale, everything gets blurry. Electrons and nuclei no longer have a well-defined position and are instead described by a joint probability distribution over space, given by $|\psi(r, R)|^2$ where $r = (r_1, r_2, ..., r_N) \in \mathbb{R}^{3N}$ are the positions of the electrons and $R = (R_1, R_2, ..., R_M) \in \mathbb{R}^{3M}$ are the positions of the nuclei. $\psi(r, R)$ is known as the *wavefunction* of the system. To determine ψ, we need only solve one equation, the Schrödinger equation:

$$H |\psi\rangle = E |\psi\rangle , \tag{2.1}$$

where $|\psi\rangle$ is a vector in a Hilbert space $\mathcal{H}$, equipped with an inner product $\langle \cdot | \cdot \rangle$ from which we can obtain the real-space wavefunction as $\psi(r, R) = \langle r, R | \psi \rangle$. Equation 2.1 is simply the eigenvalue equation for the operator $H : \mathcal{H} \rightarrow \mathcal{H}$, known as the *Hamiltonian*, where the eigenvector $|\psi\rangle$ corresponds to a system state with well-defined energy E. The operator H is therefore composed of the individual contributions to the energy of the system. For a system of interacting electrons and nuclei (i.e essentially any material), this operator takes the following form in real space:

$$H = \underbrace{\sum_{r_i} -\frac{1}{2} \nabla^2_{r_i}}_{\text{Electron kinetic}} + \underbrace{\sum_{R_i} -\frac{1}{2m_i} \nabla^2_{R_i}}_{\text{Nuclear kinetic}} + \underbrace{\sum_{r_i, r_j} \frac{1}{|r_i - r_j|}}_{\substack{\text{Electron-electron} \\ \text{repulsion}}} + \underbrace{\sum_{R_i, R_j} \frac{z_i z_j}{|R_i - R_j|}}_{\substack{\text{Nuclear-nuclear} \\ \text{repulsion}}} - \underbrace{\sum_{R_i, r_j} \frac{z_i}{|R_i - r_j|}}_{\substack{\text{Electron-nuclear} \\ \text{attraction}}},$$

$$\tag{2.2}$$

where z_i are the nuclear charges and we work in atomic units. As $\mathcal{H}$ is an infinite-dimensional space, there are an infinite number of solutions to 2.1, which can be indexed in order of increasing energy as $H |\psi_n\rangle = E_n |\psi_n\rangle$ with $E_0 \leq E_1 \leq E_2....$ For a real system at a temperature T, states with $E_n - E_0 \gg k_b T$ will quickly dissipate energy to their surroundings and the system will relax down the ladder towards

lower-energy states. This means that as the temperature decreases, fewer and fewer states become relevant. In particular, we can look at the zero-temperature limit where the system eventually finds itself in the *ground state* $|\psi_0\rangle$. This state can be used as a starting point from which finite-temperatures can be considered in terms of the excitations $|\psi_0\rangle \rightarrow |\psi_n\rangle$ that they permit. Much of this chapter is concerned with methods to find this ground state.

Having set the stage, all we now need do is solve equation 2.1 for the ground state and relevant exited states. Unfortunately, this is easier said than done. The problem lies not in the complexity of the equation itself, or even in the mildly complicated form of the Hamiltonian, but the size of the configuration space in which the problem is posed. Quantum mechanics is formulated not in the real world, but in the *configuration space* of its constituent particles $(r, R) \in \mathbb{R}^{3N} \times \mathbb{R}^{3M}$. Even worse, the central quantity of interest is not an object in this space, but rather a function on it; $\psi(r, R)$. Knowledge of this function involves knowledge of every point in configuration space, which grows rapidly enough with system size to render all but the smallest model systems intractable. This is known as the quantum *many-body* problem. For example, consider the modest case of a single Neon atom ($z = 10$); if we wished to store a real-space representation of the many-electron wavefunction on a grid with a spacing of 0.1Å out to a range of ± 5Å in each direction, we would need to store $(10/0.1)^{3\times10}$ numbers. If we use 64-bit floating point numbers, this means we would need to store 6.4×10^{61} bits of information; around 10,000 times more bits than there are atoms[1] in the sun. Even Moore's law can't help us now.

The remainder of this section starts with an introduction to methods that can tackle the many-body problem by approximating the Hamiltonian H, or the state $|\psi\rangle$. We will then move on to a discussion of how these methods can be applied to the discovery of novel materials. In particular, we will look at methods to determine the crystal structure and superconducting properties of materials from first-principles.

2.2 Born-Oppenheimer approximation

The first approximation we will make involves decoupling the electronic (r) and nuclear (R) degrees of freedom. This is made possible by the considerable mass difference between electrons and nuclei ($M_{\text{proton}}/M_{\text{electron}} \approx 1823$), which means that electrons respond much more quickly than nuclei to changes in their local environment. In particular, the electrons adjust very quickly to any *nuclear* motion (as a school of fish around a whale) and, from the point of view of the electrons, the nuclei effectively appear stationary. This allows us to decouple the electronic

---

[1] Loosely defined

and nuclear degrees of freedom and act on a decoupled electronic state $|\psi\rangle$ (which now depends only on r) with a Hamiltonian that depends only parametrically on the nuclear coordinates R:

$$H_{\text{elec}}(R)\,|\psi\rangle = E_{\text{elec}}(R)\,|\psi\rangle = \left(\underbrace{\sum_{r_i} -\frac{1}{2}\nabla^2_{r_i}}_{\text{Electron kinetic}} + \underbrace{\sum_{r_i,r_j} \frac{1}{|r_i - r_j|}}_{\substack{\text{Electron-electron} \\ \text{repulsion}}} - \underbrace{\sum_{R_i,r_j} \frac{z_i}{|R_i - r_j|}}_{\substack{\text{Electron-nuclear} \\ \text{attraction}}} \right) |\psi\rangle \tag{2.3}$$

Here, the electronic energy eigenvalue $E_{\text{elec}}(R)$ inherits the parametric dependence on the nuclear configuration and forms the energy surface within which the nuclei move. $E_{\text{elec}}(R)$ is known as a *Born-Oppenheimer surface* and is felt as an effective potential by the nuclear state $|\phi\rangle$, appearing in the corresponding nuclear Schrödinger equation:

$$H_{\text{nuc}}\,|\phi\rangle = \left(\underbrace{\sum_{R_i} -\frac{1}{2m_i}\nabla^2_{R_i}}_{\text{Nuclear kinetic}} + \underbrace{\sum_{R_i,R_j} \frac{z_i z_j}{|R_i - R_j|}}_{\substack{\text{Nuclear-nuclear} \\ \text{repulsion}}} + \underbrace{E_{\text{elec}}(R)}_{\substack{\text{Effective electronic} \\ \text{potential}}} \right) |\phi\rangle . \tag{2.4}$$

This separation of electronic and nuclear degrees of freedom is known as the Born-Oppenheimer approximation and is valid whenever the Born-Oppenheimer surfaces from different exited electronic states in Eq. 2.3 are well-seperated ($E^{(0)}_{\text{elec}}(R) \ll E^{(1)}_{\text{elec}}(R) \ll E^{(2)}_{\text{elec}}(R)...$).

When nuclear dynamics are unimportant, one need only solve the electronic Schrödinger equation (Eq. 2.3) once at a particular nuclear configuration of interest (the average nuclear configuration of a crystal, for example). If, on the other hand, nuclear dynamics are important, it is necessary to solve the electronic Schrödinger equation for a variety of different configurations R to build up sufficient information about the Born-Oppenheimer surface $E_{\text{elec}}(R)$ to allow the solution (or approximate solution) of the nuclear Schrödinger equation (Eq. 2.4). For this reason, calculations involving nuclear dynamics are typically much more expensive than those just concerned with electronic properties.

Whilst the Born-Oppenheimer approximation is helpful in reducing the size of the problem, we still need to be able to solve the (now decoupled) electronic and nuclear problems, which are still many-body problems in their own right. In the following sections, we will discuss solution methods for these sub-problems.

2.3 Quantum Monte Carlo

"I consider that I understand an equation when I can predict the properties of its solutions, without actually solving it." - Paul Dirac

An important thing to remember in quantum mechanics is that the wavefunction is an abstract quantity from which we can derive measurable quantities - we don't necessarily care about the form of the wavefunction itself. For an operator O whose eigenvalues correspond to an observable property, the average value of that property over successive measurements for a state $|\psi\rangle$ is given by a quantum-mechanical expectation value:

$$\langle O \rangle = \langle \psi | O | \psi \rangle = \int \psi^*(r) O \psi(r) dr^{3N}. \tag{2.5}$$

This can be rewritten as the expectation of the *local* values of the operator $O_L(r) = \psi(r)^{-1} O \psi(r)$, with respect to the probability density $|\psi(r)|^2$, as

$$\langle O \rangle = \int |\psi(r)|^2 O_L(r) dr^{3N}. \tag{2.6}$$

If we can generate a set of K configurations $\{r^{(i)} : i \in [1, ..., K]\}$ that are distributed according to $|\psi(r)|^2$, then Eq. 2.6 can be approximated as

$$\langle O \rangle \approx \frac{1}{K} \sum_i O_L(r^{(i)}) \tag{2.7}$$

This process is known as *Monte Carlo* integration. The error in evaluating $\langle O \rangle$ via this method can be estimated from the variance of O_L (assuming uncorrelated samples $r^{(i)}$):

$$\sigma_{\langle O \rangle} = \sqrt{Var(\langle O \rangle)} = \sqrt{\frac{Var(O_L)}{K}} = \frac{\sigma_{O_L}}{\sqrt{K}} \tag{2.8}$$

Importantly, this error does not increase with the dimension of the integral, only the variance of the local operator, and can be made arbitrarily small by taking more samples (increasing K). This makes Monte Carlo an attractive choice for extracting observable quantities from the many-body problem. However, we still need to be able to generate the sample configurations $r^{(i)}$ so that they are distributed according to $|\psi(r)|^2$. We will investigate two ways to do this.

2.3.1   Variational Monte Carlo (VMC)

If we have access to a trial wavefunction $\psi_T(r, \alpha)$ that we can evaluate at a particular configuration r and set of variational parameters α, then we can generate a set of samples distributed according to $|\psi_T(r, \alpha)|^2$ via the *metropolis algorithm*:

Algorithm 1 The metropolis algorithm. Given an arbitrary probability density $T(r \to r')$ for a move from r to r', this algorithm will generate a set S of K samples from the distribution $P(r)$.

1: r = random point in configuration space

2: S = empty set

3: **repeat**

4:     draw r' from $T(r \to r')$

5:     $a = T(r' \to r)P(r')/[T(r \to r')P(r)]$

6:     **if** $a >$ random $\in [0,1]$ **then**

7:         r = r'

8:     **end if**

9:     add r to S

10: **until** S contains K samples

A common choice for the transition function is $T(r \to r') \sim \exp(-(r-r')^2/a^2)$, which corresponds to drawing r' from a normal distribution centred at r. By taking $P(r) = |\psi_T(r, \alpha)|^2$, this algorithm will generate a set of samples which can be used to evaluate expectation values via Eq. 2.7. In particular, we can evaluate the expectation value of the energy $E(\alpha) = \langle \psi_T(\alpha)| H |\psi_T(\alpha)\rangle$, which is minimized as $|\psi(\alpha)\rangle \to |\psi_0\rangle$, the quantum mechanical ground state. Therefore, by tweaking our variational parameters α to minimize $E(\alpha)$, we can obtain an improved estimate of the ground state and its energy.

The main drawback of the variational method is that it requires an evaluatable, many-body, trial wavefunction $\psi_T(r, \alpha)$ which, for a suitable value of α, is sufficiently close to the ground state for our purposes; generating such a function is arguably just as difficult as solving the Schrödinger equation in the first place! In the following section, we will explore a method which goes a long way to solving this problem.

2.3.2 Diffusion Monte Carlo (DMC)

In the previous section we saw how VMC provides us with a method to optimize a given, parameterised, wavefunction $\psi_T(r, \alpha)$ so that it approaches the ground state. In principle, by increasing the variational degrees of freedom (i.e increasing the number of parameters) of this trial wavefunction, we can get a closer and closer approximation to the true ground state. In the limit of infinite degrees of freedom, we can treat the entire wavefunction $\psi(r)$ as a variational quantity (one could think of treating the values of the wavefunction at each and every point in configuration space as variational parameters). We can then consider making some small modification to our wavefunction $\psi(r) \to \psi(r) + \delta\psi(r)$, in order to reduce the expectation of our

energy $E = \langle \psi | H | \psi \rangle$. A sensible choice for the step $\delta \psi(r)$ is to move downhill along the negative gradient of E, by some small step size $\delta \tau$:

$$\delta \psi = -\delta \tau \frac{\delta E}{\delta \psi} = -\delta \tau H \psi \tag{2.9}$$

this is analogous to steepest-descent in finite-dimensional optimization problems, but where the gradient is replaced by the functional derivative $\delta E / \delta \psi$. Taking the limit $\delta \tau \to 0$, Eq. 2.9 becomes the differential equation

$$\frac{\partial \psi(r, \tau)}{\partial \tau} = -(H - E_T) \psi(r, \tau) \tag{2.10}$$

where an energy offset E_T has been introduced to maintain normalization of the state as $\tau \to \infty$. This equation is known as the *imaginary time Schrödinger equation* (ITSE)[2], named as such because it can be obtained (to within the offset E_T) by making the substitution $\tau = it$ in the time-dependent Schrödinger equation (TDSE) $i\partial \psi(r,t)/\partial t = H \psi(r,t)$. To gain insight into the behaviour of $\psi(r, \tau)$ as τ varies, we can expand ψ in the basis of eigenstates of H:

$$|\psi(\tau)\rangle = \sum_n C_n(\tau) |\psi_n\rangle \tag{2.11}$$

where $C_n(\tau) = \langle \psi_n | \psi(\tau) \rangle$ and $H |\psi_n\rangle = E_n |\psi_n\rangle$. Substituting this into Eq. 2.10 we obtain

$$\frac{\partial |\psi(\tau)\rangle}{\partial \tau} = -(H - E_T) |\psi(\tau)\rangle = \sum_n \frac{\partial C_n(\tau)}{\partial \tau} |\psi(\tau)\rangle = -\sum_n C_n(\tau)(E_n - E_T) |\psi_n\rangle \tag{2.12}$$

taking the inner product of this equation with $\langle \psi_m |$ we have

$$\frac{\partial C_m(\tau)}{\partial \tau} = -(E_m - E_T) C_m(\tau) \implies C_m(\tau) = C_m(0) e^{-(E_m - E_T)\tau}. \tag{2.13}$$

We see that the exited state components of the wavefunction (with $E_m > E_0$) decay exponentially quickly relative to the ground state component, which we can retain by setting $E_T = E_0$ so that $\psi(r, \tau) \to \psi_0(r)$ as $\tau \to \infty$. Re-inserting the coefficients from Eq. 2.13 into Eq. 2.11 we have

$$|\psi(\tau)\rangle = \sum_n C_m(0) e^{-E_m \tau} |\psi_n\rangle = \exp(-\tau(H - E_T)) |\psi(0)\rangle. \tag{2.14}$$

---

[2]We introduce the ITSE here mainly in the interest of consistency with existing literature. In actual fact, I think that the connection to the TDSE is at best unhelpful and, at worst, misleading. In reality, the ITSE is simply a tool used to attack the variational problem (in the limit $\tau \to \infty$) and should not be elevated to the level of a fundamental equation which we are aiming to solve (after all, it does not correspond to a physical evolution of the system). Later in this thesis we will treat the variational problem as fundamental (as it should be) and we will see how the offset E_T appears simply as a Lagrange multiplier corresponding to the normalization constraint.
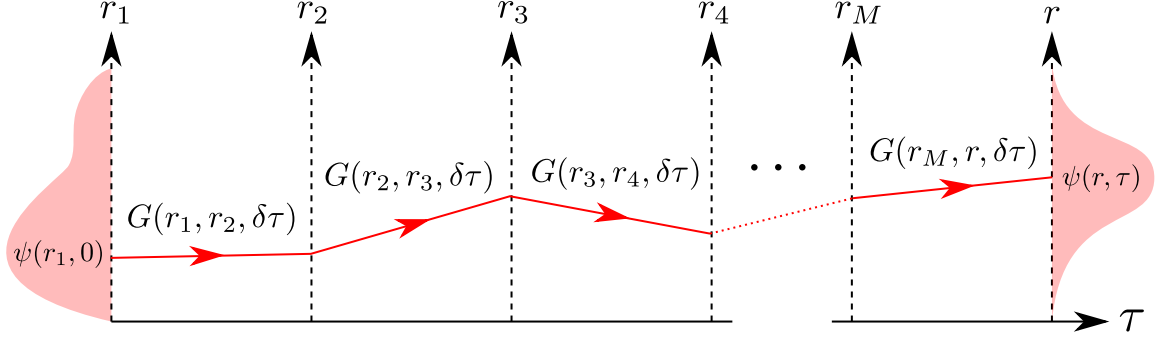
Figure 2.1: The discretised path integral (Eq. 2.17), showing a particular path in red along with the Green's function factors accumulated along it.

We see that the state at τ can be obtained by applying the imaginary-time propagation operator $\exp(-\tau(H - E_T))$. We can split up the action of this operator into $M = \tau/\delta\tau$ individual timesteps

$$\exp(-\tau(H - E_T)) = \exp(-M\delta\tau(H - E_T)) = \prod_{i=1}^{M} \exp(-\delta\tau(H - E_T)) \qquad (2.15)$$

Inserting this into Eq. 2.14, and moving to the position representation, we have

$$\psi(r, \tau) = \langle r | \psi(\tau) \rangle = \left\langle r \left| \prod_{i=1}^{M} \exp(-\delta\tau(H - E_T)) \right| \psi(0) \right\rangle \qquad (2.16)$$

Inserting a resolution of the identity $I = \int |r\rangle \langle r| \, d^{3N}r$ at each timestep, we have

$$\psi(r, \tau) = \int \langle r | \exp\left(-\delta\tau(H - E_T)\right) | r_M \rangle \langle r_M | \exp\left(-\delta\tau(H - E_T)\right) | r_{M-1} \rangle \ldots$$
$$\ldots \langle r_3 | \exp\left(-\delta\tau(H - E_T)\right) | r_2 \rangle \langle r_2 | \exp\left(-\delta\tau(H - E_T)\right) | r_1 \rangle \qquad (2.17)$$
$$\langle r_1 | \psi(0) \rangle d^{3N}r_1 d^{3N}r_2 \ldots d^{3N}r_M$$

The integrand of this equation can be thought of as a path $r_1, r_2, r_3, \ldots, r_M, r$ through configuration space where each segment contributes a factor of the *Green's function* $G(r_i, r_{i+1}, \delta\tau) = \langle r_{i+1} | \exp(-\delta\tau(H - E_T)) | r_i \rangle$ (see figure 2.1). The integration over $\{r_i\}$ can therefore be thought of as a *path integral* summing the contributions to $\psi(r, \tau)$ from all possible paths in configuration space linking back to $\psi(r_1, 0)$[3]. Each of these paths can be thought of as the trajectory of a point-like object wandering through configuration space. We call these objects *walkers*. Because $\psi(r, \tau) \to \psi_0(r)$ as $\tau \to \infty$, the walkers eventually end up distributed according to the ground state wavefunction - exactly what we need to evaluate expectation values via Eq. 2.7! We can represent the ith walker by a configuration path $x_i(\tau)$

---

[3]If we had done this with the TDSE instead of the ITSE, we would have derived path-integral quantum mechanics.

and a weight $w_i(\tau)$ given by the Green's function contributions accumulated along that path:

$$w_i(\tau_k) = \prod_{j=0}^{k-1} G(x_i(\tau_j), x_i(\tau_{j+1}), \delta\tau) \tag{2.18}$$

where we have introduced the notation $\tau_j = j\delta\tau$ for the timesteps. Our wavefunction (Eq. 2.17) can then be written as

$$\psi(x, \tau) = \lim_{W \to \infty} \sum_{i=0}^{W} w_i(\tau)\delta(x - x_i(\tau)) \tag{2.19}$$

where the limit of infinite walkers $\lim_{W \to \infty}$ corresponds to the infinite number of possible paths $x_i(\tau)$ in the path integral. Given an initial configuration $\{x_i(\tau = 0), w_i(\tau = 0)\}$, a simple way to generate a walker path is to iterate this configuration forward at each timestep according to some transition probability $T(x_i(\tau_j) \to x_i(\tau_{j+1}))$. According to Eq. 2.18, this transition modifies the weight by a factor of $G(x_i(\tau_j), x_i(\tau_{j+1}), \delta\tau)$, leading to the following propagation scheme:

$$\begin{aligned} x_i(\tau_{j+1}) \ \ \text{sampled from} \ \ &T(x_i(\tau_j) \to x_i(\tau_{j+1})) \\ w_i(\tau_{j+1}) = w_i(\tau_j) \times \ &G(x_i(\tau_j), x_i(\tau_{j+1}), \delta\tau)/T(x_i(\tau_j) \to x_i(\tau_{j+1})) \end{aligned} \tag{2.20}$$

where we have divided the updated weight by a factor of $T(x_i(\tau_j) \to x_i(\tau_{j+1}))$, to remove the bias towards particular paths introduced by T.

In practice, we can't apply the above propagation scheme to an infinite number of walkers. By approximating the wavefunction (Eq. 2.19) using a finite walker population W, we arrive at the wavefunction of *diffusion Monte Carlo* (DMC)

$$\psi_{\text{DMC}}(x, \tau) = \sum_{i=0}^{W} w_i(\tau)\delta(x - x_i(\tau)). \tag{2.21}$$

In principle, when making this approximation, we can choose to retain whichever configuration paths $x_i(\tau)$ we like, so long as the weights are properly accounted for; this is reflected in our freedom to choose $T(x_i(\tau_j) \to x_i(\tau_{j+1}))$ in Eq. 2.20. However, it makes sense to attempt to retain the paths which contribute the most to Eq. 2.17, corresponding to the walkers with the largest weights w_i. In order to do this we treat the walkers with a *birth-death* algorithm whereby, each iteration, a walker with weight w_i is replaced by $\lfloor w_i + u \rfloor$ walkers, each with weight 1 at the same configuration. Here u is a uniform random number $\in [0, 1]$ and $\lfloor \cdot \rfloor$ is the floor function. In this way, walkers that have insignificant weight will have an insignificant probability of surviving to the next iteration. In contrast, walkers that have accumulated too much weight will divide into several children, to ensure that sufficiently many different paths are sampled. Note that this procedure leaves the total weight $W = \sum_i w_i$ unchanged on average.

The fluctuations in walker population arising from the birth-death algorithm introduced in the previous section lead to corresponding fluctuations in expectation values calculated via Eq. 2.7. This leads to an increased variance (and therefore error) in our Monte Carlo estimate of the expectation value. By considering Eq. 2.20, we can see that fluctuations in the weights w_i can be removed entirely by choosing $T(x \to x') = G(x, x', \delta\tau)$. In this way, we see that the Green's function $G(x, x', \delta\tau)$ could have been interpreted as a transition probability from x to x'. Indeed, by suitable choices for T in Eq. 2.20, we can choose to partition the Green's function factor in whatever way we wish between contributing to the transition probability $T(x_i(\tau_j) \to x_i(\tau_{j+1}))$ or the weight update $w_i(\tau_j) \to w_i(\tau_{j+1})$. To inform this decision, we need to consider the form of G;

$$G(x, x', \delta\tau) = \langle x' | \exp(-\delta\tau(H - E_T)) | x \rangle = \langle x' | \exp(-\delta\tau(H - E_T)) | x \rangle \quad (2.22)$$

We can partition H into kinetic and potential terms

$$H = T + V = -\frac{1}{2}(M^{-1}\nabla)^2 + V(x) \quad (2.23)$$

where the many-body potential $V(x)$ includes all inter-particle interactions exactly and $M = \mathrm{diag}(m_1, m_1, m_1, m_2, m_2, m_2...)$ is the mass tensor (with m_i being the mass of the ith particle). With this partitioning of H, we can use the Suzuki-Trotter [264] expansion of the operator $\exp(-\delta\tau H)$ for small $\delta\tau$ (see Appendix A.1) leading to

$$G(x, x', \delta\tau) = G_V(x, x', \delta\tau)G_D(x, x', \delta\tau)\exp(\delta\tau E_T) \quad (2.24)$$

with the potential-dependent part given by

$$G_V(x, x', \delta\tau) = \exp(-\delta\tau[V(x') + V(x)]/2) \quad (2.25)$$

and the kinetic-dependant part given by a $3N$-dimensional Gaussian, with variance $\delta\tau/M_{ii}$ in the ith direction in configuration space:

$$G_D(x, x', \delta\tau) = \frac{|M|}{(2\pi\delta\tau)^{3N/2}} \exp\left(\frac{-|M(x - x')|^2}{2\delta\tau}\right). \quad (2.26)$$

We could have obtained $G = G_D$ by setting $V = E_T = 0$ in the ITSE, leading to a diffusion equation:

$$\frac{1}{2}(M^{-1}\nabla)^2\psi(r, \tau) = \frac{\partial\psi(r, \tau)}{\partial\tau}. \quad (2.27)$$

We therefore refer to G_D as the *diffusive* part of the Green's function (hence the subscript D). Indeed, sampling moves according to G_D alone would lead to a diffusive random walk across configuration space.

Now that we have an explicit form for G, we can return to the question of how it is divided between the transition probability $T(x_i(\tau_j) \to x_i(\tau_{j+1}))$ and the

weight update $w_i(\tau_j) \rightarrow w_i(\tau_{j+1})$. As mentioned previously, we should absorb as much of G as possible into the transition probability, to reduce fluctuations in walker weights. However, it is, in general, difficult to sample moves from an arbitrary transition probability, prohibiting the naive choice of $T(x_i(\tau_j) \rightarrow x_i(\tau_{j+1})) = G(x_i(\tau_j), x_i(\tau_{j+1}), \delta\tau)$. Instead we note that it is comparatively straightforward to sample x' from $G_D(x, x', \delta\tau)$ thanks to the fact it has the same form as a normal distribution (which we can generate samples from using a Box-Muller transformation [36]). This prompts the choice $T(x_i(\tau_j) \rightarrow x_i(\tau_{j+1})) = G_D(x_i(\tau_j), x_i(\tau_{j+1}), \delta\tau)$, leading to the following realisation of Eq. 2.20:

$$
\begin{aligned}
&x_i(\tau_{j+1}) \text{ sampled from } G_D(x_i(\tau_j), x_i(\tau_{j+1}), \delta\tau) \\
&w_i(\tau_{j+1}) = w_i(\tau_j) \times G_V(x_i(\tau_j), x_i(\tau_{j+1}), \delta\tau) \exp(\delta\tau E_T)
\end{aligned}
\tag{2.28}
$$

This corresponds to moving the walkers around configuration space according to the diffusion equation 2.27, but updating their weights according to G_V, which depends only on the local potential $V(x)$, and the renormalization factor $\exp(\delta\tau E_T)$. This is the propagation scheme employed in a *diffusion Monte Carlo* (DMC) calculation.

2.3.3 The fermion sign problem

The DMC scheme derived in the previous section relies on the fact that application of the imaginary-time propagator $\exp(-\tau(H - E_T))$ will project out the lowest-energy component of our initial wavefunction:

$$
\lim_{\tau \rightarrow \infty} \exp(-\tau(H - E_T)) |\psi(\tau = 0)\rangle = |n_{\min}\rangle
\tag{2.29}
$$

where $|n_{\min}\rangle$ is the lowest-energy eigenstate of H that has non-zero overlap with the initial state $|\psi(\tau = 0)\rangle$. If this state is to describe our system correctly, it must also have the correct exchange symmetry under the particle exchange operators P_i;

$$
P_i |n_{\min}\rangle = \begin{cases} |n_{\min}\rangle \text{ if } P_i \text{ exchanges identical bosons} \\ -|n_{\min}\rangle \text{ if } P_i \text{ exchanges identical fermions} \end{cases}
\tag{2.30}
$$

Because the Hamiltonian commutes with particle exchange $[P_i, H] = 0$, if we start in a state with well-defined exchange symmetry, we should retain this symmetry under the action of Eq. 2.29. However, in the DMC scheme derived in the previous section, we simulate the action of the propagator $\exp(-\tau(H - E_T))$ *probabilistically*. Whilst, on average, this reproduces the symmetry-preserving behaviour of the propagator, symmetry is **not** preserved on a per-timestep basis. This means that the DMC method will eventually find its way to the absolute ground state of H, irrespective of its exchange symmetry. This is a particular problem for electronic systems, where the fermionic ground state is typically not the absolute ground state of H and so
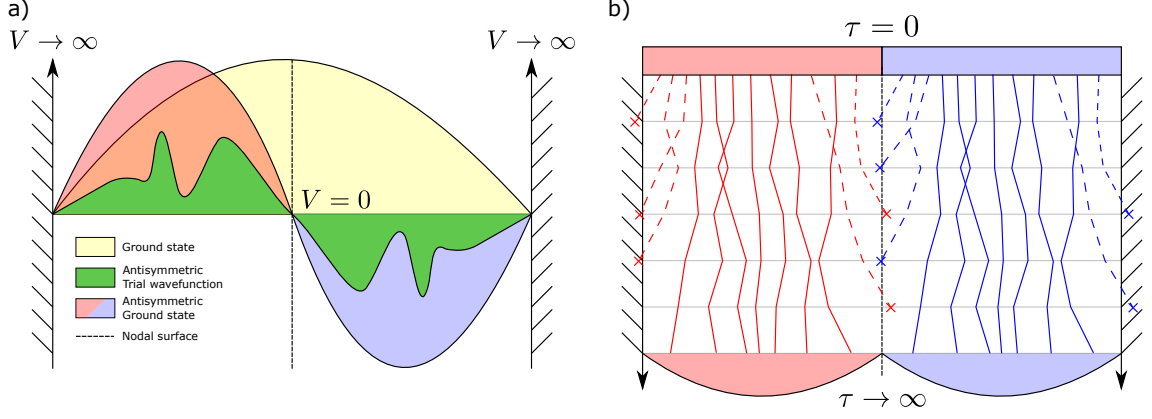
Figure 2.2: An illustration of the fixed node approximation for an infinite square well potential. On the left, we see the absolute ground state (yellow), the trial wavefunction (green) and the antisymmetric ground state (first exited state, positive in red, negative in blue). On the right we see how the antisymmetric ground state can be obtained by propagating a collection of +ve (red lines) and -ve (blue lines) walkers, subject to removal (indicated by crosses) of walkers that cross the nodal surface (dotted line at the centre), or that stray into the infinite-potential region (edges). The removal of walkers creates a bias in the walker density away from the edges/centre as $\tau \to \infty$, resulting in the correct sine-like wavefunction. This can be thought of as two independent DMC simulations (red and blue) confined to the *nodal pockets* (regions of continuous sign) of the trial wavefunction.

is projected away along with all of the exited states, exponentially quickly. This is known as the *fermion sign problem* and, because of the exponential decay of the states we are interested in, is a very serious problem.

The fixed node approximation

In order to overcome the sign problem, we need to enforce exchange symmetry for the ensemble of walkers at every timestep. One way to achieve this is by reference to a known wavefunction of the correct symmetry, known as the *trial wavefunction*, $\psi_T(x)$. We can enforce the symmetry of ψ_T upon the walkers by requiring that the sign of each walker is equal to the sign of ψ_T at every timestep (i.e sign$[w_i(\tau)] \overset{!}{=}$ sign$[\psi_T(x_i(\tau))] \; \forall \; \tau$). Typically, this is enforced by removing from the simulation any walker who's weight is not of the correct sign. The resulting scheme is known as *fixed-node diffusion Monte Carlo* (FNDMC), because the zero-crossings $\psi(x) = 0$ (nodes) of our DMC wavefunction are fixed to that of the trial wavefunction. This is illustrated for the case of finding the first exited state of a square-well potential in Fig. 2.2.

The fixed node approximation is one of the most widely-employed quantum

Monte Carlo schemes and forms the basis of some of the most important quantum mechanical calculations ever performed [40]. However, the reliance on a trial wavefunction is a major drawback for several reasons. Firstly, one must be able to generate ψ_T *a priori*, in practice by using alternative quantum mechanical methods (e.g by building a Slater determinant using the non-interacting states from Hartree-Fock theory, or density functional theory [4]). Because of this reliance on alternative methods, FNDMC is not a self-contained solution method for the many-body problem in its own right. Secondly, by enforcing the nodal surface of some approximate trial wavefunction, we introduce a *fixed-node error* into expectation values which is expensive to systematically reduce and is often the limiting factor in obtaining higher accuracy results [105, 227]; Because the trial wavefunction is often generated using methods that do not contain explicit electron-electron interactions, the resulting nodal surface does not exhibit the full influence of electron-electron correlations. In order to remedy this, one can introduce additional correlated degrees of freedom into the wavefunction by including a Jastrow factor [130] or backflow corrections [154]. Whilst this can often recover more than 90% of the correlation energy [90], it introduces additional parameters that must be optimized, either beforehand using VMC (see Sec. 2.3.1), or (preferably) in parallel with DMC propagation. The careful generation and optimization of trial wavefunctions for use with FNDMC is an involved process and is an active field study in its own right. In chapter 3, we explore an alternative method to FNDMC to enforce exchange symmetry.

2.4   Density functional theory (DFT)

The Quantum Monte Carlo methods explored in section 2.3 work directly with the full many-body Hamiltonian H. Whilst this means that all interactions are included explicitly and without approximation, it also means that we must work in the full $3N$-dimensional configuration space. In order to make this manageable, we had to approximate the state $|\psi\rangle$. In this section we will instead shift our focus to the Hamiltonian. Working within the Born-Oppenheimer approximation, we can write our electronic Hamiltonian as $H_{\text{elec}} = T + V_{ee} + V_{ext}$ where

$$T = \sum_{r_i} -\frac{1}{2}\nabla^2_{r_i}, \quad V_{ee} = \sum_{r_i,r_j} \frac{1}{|r_i - r_j|}, \quad V_{ext} = \sum_{r_i} V_{ext}(r_i) \qquad (2.31)$$

are the electron kinetic energy, electron-electron repulsion and external potential respectively. We have slightly generalized the form considered in Eq. 2.3 to an

---

[4]Using non-interacting single-particle states to construct the interacting many-body state in this way does not lead to the exact many-body state, or many-body nodal surface; this is an important source of errors in practical calculations.

arbitrary potential $V_{ext}(r)$ that acts the same way on each electron (this includes the specific case of electron-nuclear interaction). For an electronic state $|\psi\rangle$, the energy can be written as

$$E = \langle\psi| H_{\text{elec}} |\psi\rangle = \langle\psi| T + V_{ee} + V_{ext} |\psi\rangle \tag{2.32}$$

Focusing just on the contribution from the external potential, we have

$$\langle\psi| V_{ext} |\psi\rangle = \int |\psi(r_1, ..., r_N)|^2 \sum_i V_{ext}(r_i) d^3r_1...d^3r_N \tag{2.33}$$

$$= \sum_i \int \left(V_{\text{ext}}(r_i) \int |\psi(r_1, ..., r_N)|^2 d^3r_1...d^3r_N \setminus d^3r_i \right) d^3r_i \tag{2.34}$$

where $\setminus d^3r_i$ indicates that the inner integral is not performed over r_i. Because of exchange symmetry we can permute any of the arguments of $|\psi|^2$ and thus rewrite this integral as

$$\int |\psi(r_1, ..., r_N)|^2 d^3r_1...d^3r_N \setminus d^3r_i = \int |\psi(r_i, ..., r_N)|^2 d^3r_2...d^3r_N = \frac{n(r_i)}{N} \tag{2.35}$$

where $n(r)$ is the electron density (normalized to N). Substituting this into Eq. 2.34, we have

$$\langle\psi| V_{ext} |\psi\rangle = \sum_i \int V_{\text{ext}}(r_i) \frac{n(r_i)}{N} d^3r_i = \int V_{\text{ext}}(r) n(r) d^3r \tag{2.36}$$

Hence, the energy due to an external potential depends only on the electron *density* rather than the full many-body wavefunction. The former is a function on $\mathbb{R}^3$ rather than $\mathbb{R}^{3N}$ and thus is exponentially easier to store, and perform operations on, in a computer. If we can also write the other contributions to the energy in terms of just the electron density, we will be able to do away with the wavefunction altogether and instead treat $n(r)$ as the fundamental quantity. This is the goal of *density functional theory* (DFT), named because the most general form of a density-dependant energy is a *functional* of the density. Writing the contribution from the external potential explicitly, this functional can be written as

$$E[n(r)] = F[n(r)] + \int V_{ext}(r) n(r) d^3r. \tag{2.37}$$

We note that V_{ext} contains all of the system-dependent terms in the Hamiltonian. This means that the remaining contribution $F[n(r)]$ is a *universal* functional which is independent of the system in question. The proof of the existence of such a functional is due to Hohenberg and Kohn [114]. We demonstrate this proof by employing an explicit construction of the functional F due to Levy [155]:

$$F[n(r)] = \min\{\langle\psi|T + V_{ee}|\psi\rangle : |\psi(r)|^2 = n(r)\} \equiv \langle\psi_{\min}[n]| T + V_{ee} |\psi_{\min}[n]\rangle. \tag{2.38}$$

Substituting this into Eq. 2.37 we have

$$E[n(r)] = \langle\psi_{\min}[n]| \, T + V_{ee} \, |\psi_{\min}[n]\rangle + \int V_{ext}(r)n(r)d^3r$$

$$= \langle\psi_{\min}[n]| \underbrace{T + V_{ee} + V_{ext}}_{H_{\text{elec}}} |\psi_{\min}[n]\rangle \geq E_{\text{GS}} \qquad (2.39)$$

where the second line follows from Eq. 2.36 and the inequality follows from the quantum-mechanical variational principle applied to H_{elec}. Equation 2.39 amounts to a variational principle of the energy functional with respect to the density. From the definition of F, it can be clearly seen that equality is obeyed if $n(r) = n_{\text{GS}}(r)$, the ground state electron density;

$$E[n_{\text{GS}}(r)] = \min\{\langle\psi|T + V_{ee}\,|\psi\rangle : |\psi(r)|^2 = n_{\text{GS}}(r)\} + \int V_{ext}(r)n(r)d^3r$$

$$= \min\{\langle\psi|H\,|\psi\rangle : |\psi(r)|^2 = n_{\text{GS}}(r)\} \qquad (2.40)$$

$$\leq \langle\psi_{\text{GS}}|H\,|\psi_{\text{GS}}\rangle = E_{\text{GS}}$$

where the inequality is a consequence of the minimization. Combining equations 2.39 and 2.40, we must have

$$E[n_{\text{GS}}(r)] = E_{\text{GS}} \qquad (2.41)$$

Equations 2.39 and 2.41 are known as the *Hohenberg-Kohn theorems* and are central to the development of DFT.

Kohn and Sham unpacked the form of F so as to include known contributions explicitly, leading to:

$$E[n(r)] = T_0[n(r)] + \int \left(V_{ext}(r) + \frac{1}{2}\phi(r)\right) d^3r + E_{xc}[n(r)] \qquad (2.42)$$

where $\phi(r)$ is the Hartree potential and $T_0[n(r)]$ is the kinetic energy of a system of non-interacting electrons [5]. Exchange and correlation effects are included in the so-called *exchange correlation functional*, $E_{xc}[n(r)]$. Thanks to the Hohenberg-Kohn theorems, we know that we are aiming to minimize the energy with respect to the density. We need to do this with a fixed number of electrons $N = \int n(r)d^3r$, so we introduce the Lagrange multiplier μ and extremise

$$S[n(r)] = E[n(r)] + \mu\left[N - \int n(r)d^3r\right] \qquad (2.43)$$

---

[5]The non-interacting kinetic energy functional $T_0[n(r)]$ is actually an unkown quantity and known approximations are not sufficiently accurate for many applications. However, as we will see later, the Kohn-Sham formalism involves the construction of an explicit wavefunction (from the single-particle Kohn-Sham orbitals) from which the kinetic energy can be calculated. The lack of an accurate kinetic energy functional hinders the development of so called *orbital-free* DFT methods, that work exclusively with the density.

With respect to $n(r)$ and μ. Extremisation with resepct to $n(r)$ leads to:

$$\frac{\delta S}{\delta n(r)} \stackrel{!}{=} 0 \implies \frac{\delta E}{\delta n(r)} = \frac{\delta T_0}{\delta n(r)} + V_{ext}(r) + \phi(r) + \frac{\delta E_{xc}}{\delta n(r)} = \mu \qquad (2.44)$$

If we compare this to the same equation for a system *without* electron-electron interactions, but subject to a potential $V_{KS}(r)$:

$$\frac{\delta E}{\delta n(r)} = \frac{\delta T_0}{\delta n(r)} + V_{KS}(r) = \mu \qquad (2.45)$$

We see that the system with electron-electron interactions is formally equivalent to a system without electron-electron interactions, but with an effective potential:

$$V_{KS}(r) = V_{ext}(r) + \phi(r) + \frac{\delta E_{xc}}{\delta n(r)} \qquad (2.46)$$

This leads to the (single-particle) Kohn-Sham reformulation of the Schrödinger equation:

$$\left(-\frac{1}{2}\nabla^2 + V_{KS}(r)\right)\phi_i(r) \equiv H_{KS}\phi_i(r) = \epsilon_i\phi_i(r) \qquad (2.47)$$

which gives us the electron density via

$$n(r) = \sum_{i=1}^{N} |\phi_i(r)|^2 \qquad (2.48)$$

from which, in principle, the exact energy may be extracted via $E[n(r)]$. However, the form of E_{xc} is not known exactly and must be approximated. Two different approximations are employed in this thesis. The first of these is the *local density approximation* (LDA), where the potential depends only on the *local* electron density

$$V_{xc} \approx V_{xc}(n(r)) \qquad (2.49)$$

The form of $V_{xc}(n(r))$ is obtained from quantum Monte Carlo calculations of the homogeneous electron gas [40]. The second is a *generalized gradient approximation* (of Perdew-Burke-Ernzerhof form [215]), in which the gradient of the electron density is also considered;

$$V_{xc} \approx V_{xc}(n(r), \nabla n(r)) \qquad (2.50)$$

Once a form for V_{xc} is chosen, the Kohn-Sham equations (2.47) can then be solved, self-consistently, in a basis of our choice. Since we are investigating crystalline systems with periodic boundary conditions a convenient choice of basis is the plane waves $\{e^{ik_n r}\}$ where $\{k_n\}$ are our reciprocal lattice vectors (hence preserving periodic boundary conditions). Obviously this infinite set of lattice vectors has to be truncated at some cutoff $|k| < k_{\text{cut}}$ (known as the plane-wave cutoff, typically referred to by the corresponding kinetic energy $E_{\text{cut}} = k_{\text{cut}}^2/2m_e$), but our results can be readily converged against this. The CASTEP [48] and Quantum Espresso [101] implementations of plane-wave DFT are used in this thesis.

2.4.1   K-points and finite-size errors

The crystalline systems that we are interested in are effectively infinite, so we might expect that an infinite amount of time would be needed to apply DFT. However, we can go a long way to remedying this issue by noting that these systems are periodic; the potential $V_{KS}(r)$ obeys

$$V_{KS}(r + L) = V_{KS}(r) \ \forall \ L \in \{n_1 a_1 + n_2 a_2 + n_3 a_3 : n_1, n_2, n_3 \in \mathbb{Z}\} \qquad (2.51)$$

where L is a lattice vector formed from the primitive lattice vectors a_1, a_2 and a_3. We can define the lattice translation operators, T_i, via their action on an arbitrary state:

$$T_i \psi(r) = \psi(r + a_i) \qquad (2.52)$$

Then, from Eq. 2.51, it is clear that $[T_i, H_{KS}] = 0$. As a result, the Hamiltonian and the translation operators share the same eigenbasis. The action of a translation operator (n times) on an energy eigenstate $\psi(r)$ is then

$$T_i^n \psi(r) = t_i^n \psi(r) = \psi(r + n a_i) \qquad (2.53)$$

Because this state must be normalized we have $|t_i^n| = 1$ and are free to make the association

$$t_i = \exp(2\pi i \theta_i) \qquad (2.54)$$

for some $\{\theta_i\}$; we can then define $k = \theta_1 b_1 + \theta_2 b_2 + \theta_3 b_3$, where $\{b_i\}$ are the primitive reciprocal lattice vectors (satisfying $b_i \cdot a_j = 2\pi \delta_{ij}$), in order to write the action of an arbitrary lattice translation L as

$$T_L \psi(r) = \psi(r + L) = \exp(ik \cdot L)\psi(r) \qquad (2.55)$$

Note that displacing k by a reciprocal lattice vector makes no difference to the resulting state (as $\exp(ib_i \cdot L) = 1$), so we can take $\{\theta_i\} \in [0, 1]$ which implies that the values of k in the 1$^{\text{st}}$ Brillouin zone are sufficient to enumerate all distinct states. Finally, this allows us to define a function $u(r) = \exp(-ik \cdot r)\psi(r)$, which inherits the periodicity of the lattice:

$$
\begin{aligned}
u(r + a_i) &= \exp(-ik \cdot (r + a_i))\psi(r + a_i) \\
&= \exp(-ik \cdot (r + a_i))T_i \psi(r) \\
&= \exp(-ik \cdot (r + a_i))\exp(ik \cdot a_i)\psi(r) \\
&= u(r)
\end{aligned}
\qquad (2.56)
$$

This means that all eigenstates of H_{KS} can be written as a product of a plane wave and a function which inherits the periodicity of the lattice:

$$\psi(r) = \exp(ik \cdot r)u(r) \qquad (2.57)$$

This result is known as *Bloch's theorem* and is central to the study of crystalline systems. In particular, in DFT this allows us to choose a particular k and to then solve the resulting Kohn-Sham equations for the possible states, or *bands*. The resulting single-particle Kohn-Sham sates may be indexed by a k vector and a band index n as $\psi_{k,n}(r) = \exp(ik \cdot r)u_{nk}(r) \equiv \langle r | k, n \rangle$ where $H_{KS} | k, n \rangle = \epsilon_{nk} | k, n \rangle$. Important quantities (density of states, charge density, matrix elements etc...) can then be obtained as integrals over the 1st Brillouin zone;

$$I = \frac{1}{\Omega_{BZ}} \int_{BZ} f(k) d^3 k \tag{2.58}$$

In order to evaluate these, we sample the Brillouin zone using a weighted sum over discrete k-points;

$$\frac{1}{\Omega_{BZ}} \int_{BZ} f(k) d^3 k \rightarrow \sum_i \omega_i f(k_i) \tag{2.59}$$

Often, an equally spaced (*Monkhorst-Pack* [197]) grid over the Brillouin zone is used to choose these sampling points. The final result is that we have reduced the problem of solving the quantum mechanics of an infinite, periodic system to the problem of a finite sampling from an infinite set of k-points in the first Brillouin zone. If the difference between the sum and integral in Eq. 2.59 is significant then it means that the set of k-points sampled was insufficient. We refer to such errors as *finite size* errors as a result of the correspondence between the infinite set of k-points and the infinite lattice. Similarly to increasing the plane-wave cutoff introduced earlier, this k-point grid can be made systematically denser in order to reduce these finite size errors.

2.5 Phonons in DFT

2.5.1 The harmonic approximation

Section 2.4 described a formalism to calculate the (approximate) electronic energy $E_{\text{elec}}(R)$ for a given set of nuclear coordinates R. However, this is only the first half of the Born-Oppenheimer problem described in section 2.2. We must now derive a treatment for the nuclear system which possesses the Hamiltonian

$$\begin{aligned} H_{\text{nuc}} | \phi \rangle &= \left(\sum_{R_i} -\frac{1}{2m_i} \nabla_{R_i}^2 + \sum_{R_i, R_j} \frac{z_i z_j}{|R_i - R_j|} + E_{\text{elec}}(R) \right) | \phi \rangle \\ &\equiv \left(\sum_{R_i} -\frac{1}{2m_i} \nabla_{R_i}^2 + V(R) \right) | \phi \rangle . \end{aligned} \tag{2.60}$$

Through which we have defined the effective potential $V(R)$ felt by the nuclei (composed of the nuclear-nuclear repulsion and Born-Oppenheimer energy $E_{\text{elec}}(R)$). We

can expand this potential in terms of the perturbations to nuclear coordinates δ_i (of which there are $3 \times$ the number of nuclei) as

$$V(R + \delta) = V(R) + \sum_i \delta_i \partial_i V(R) + \frac{1}{2} \sum_{ij} \delta_i \delta_j \partial_i \partial_j V(R) + \mathcal{O}(\delta^3) \qquad (2.61)$$

where $\partial_i V(R) \equiv \partial V(R)/\partial \delta_i$. For an equilibrium crystal structure $R^{(0)}$, $\partial_i V(R)|_{R^{(0)}} = 0 \; \forall \; i$. As a result, the lowest order in which we can consider vibrational effects is the *harmonic approximation* in which we neglect terms $\mathcal{O}(\delta^3)$ and above.

At this stage it is useful to consider the classical system of nuclei within the harmonic approximation, where the equations of motion for the atoms can be written as

$$m_i \ddot{\delta}_i = -\partial_i V(R) = -\sum_j \delta_j \partial_i \partial_j V(R)|_{R^{(0)}} \qquad (2.62)$$

where m_i is the mass of the atom to which the ith coordinate belongs and the last equality follows from differentiating equation 2.61. If we take a plane-wave ansatz for our atomic displacements (with wavevector q, frequency $\omega(q, \sigma)$ and amplitude $P_i(q, \sigma)/\sqrt{m_i}$):

$$\delta_i = \frac{1}{\sqrt{m_i}} P_i(q, \sigma) \exp\left[i(\omega(q, \sigma)t - q \cdot R_i^{(0)})\right] \qquad (2.63)$$

where $R_i^{(0)}$ is the equilibrium position of the atom to which the ith coordinate belongs, then Eq. 2.62 becomes

$$\begin{aligned}
\omega^2(q, \sigma) P_i(q, \sigma) &= \sum_j \frac{1}{\sqrt{m_i m_j}} \exp\left[iq \cdot (R_i^{(0)} - R_j^{(0)})\right] P_j(q, \sigma) \partial_i \partial_j V(R)|_{R^{(0)}} \\
&= \sum_j D_{ij}(q) P_j(q, \sigma)
\end{aligned} \qquad (2.64)$$

where

$$D_{ij}(q) = \frac{1}{\sqrt{m_i m_j}} \exp\left[iq \cdot (R_i^{(0)} - R_j^{(0)})\right] \partial_i \partial_j V(R)|_{R^{(0)}} \qquad (2.65)$$

is known as the *dynamical matrix*, and is Hermitian (due to our inclusion of the $1/\sqrt{m_i}$ factor in our ansatz). We see that Eq. 2.64 is simply the eigenvalue problem for this dynamical matrix. The resulting eigenvalues $\omega^2(q, \sigma)$ are the squared frequencies of the phonon modes (indexed by σ), at a particular phonon wavevector q. The eigenvector component P_i gives the amplitude of the modal oscillations in the coordinate δ_i.

Returning to a fully quantum treatment, we define *phonon coordinates* $p_{q,\sigma}$, in terms of our eigenvectors $P_i(q, \sigma)$ via

$$p_{q,\sigma} = \sum_i \sqrt{m_i} \exp(-iq \cdot R_i^{(0)}) P_i(q, \sigma) \delta_i \qquad (2.66)$$

then the Hamiltonian in the harmonic approximation:

$$H_{\text{nuc}}^{(2)} = \sum_{R_i} -\frac{1}{2m_i}\nabla_{R_i}^2 + \frac{1}{2}\sum_{i,j}\delta_i\delta_j\partial_i\partial_j V(R)|_{R^{(0)}} \qquad (2.67)$$

can be rewritten in terms of phonon coordinates as (see appendix A.2)

$$H^{(2)} = \sum_{q,\sigma} -\frac{1}{2}\frac{\partial^2}{\partial p_{q,\sigma}^2} + \frac{1}{2}\omega_{q,\sigma}^2 p_{q,\sigma}^2. \qquad (2.68)$$

This Hamiltonian corresponds to a sum of Hamiltonians for a series of non-interacting simple harmonic oscillators with frequencies $\omega_{q,\sigma}$.

2.5.2 Anharmonic corrections

The derivation of the Hamiltonian in the harmonic approximation (Eq. 2.68) neglects terms in the potential of third and higher order in the nuclear displacements (or, equivalently, phonon coordinates). This is a reasonable approximation only if the amplitude of the nuclear motion is small. In some situations however, this is not the case. Indeed, as the temperature of the system increases the expected displacement of the nuclei also increases. For example, the average RMS displacement of a harmonic oscillator (with $m = \hbar = \omega = 1$) is given by the square root of

$$\begin{aligned}
\langle x^2 \rangle &= \int x^2 \sum_n \frac{1}{e^{\beta(E_n-\mu)} - 1}|\psi_n(x)|^2 dx \\
&= \int x^2 \sum_n \frac{1}{e^{\beta(E_n-\mu)} - 1}|N_n H_n(x)\exp(-x^2/2)|^2 dx
\end{aligned} \qquad (2.69)$$

where $\beta = 1/k_b T$, $H_n(x)$ are the Hermite polynomials and N_n is the normalization coefficient of the nth exited state. We choose the Lagrange multiplier μ so that the total state occupation is one. The resulting RMS displacement is plotted in figure 2.3 and increases with temperature which, for real systems, would push us into the anharmonic regime. The temperature at which this effect becomes important is proportional to the frequency of oscillation, so systems with low phonon frequencies are more likely to display anharmonicity.

In order to calculate anharmonic corrections, we must generalize our vibrational Hamiltonian. Replacing the harmonic potential in Eq. 2.68 with the Born-Oppenheimer energy surface $V(R)$ gives us the fully anharmonic Hamiltonian

$$H_A = \sum_{q,\sigma} -\frac{1}{2}\frac{\partial^2}{\partial p_{q,\sigma}^2} + V(R(\{p_{q,\sigma}\})) \qquad (2.70)$$

Neglecting coupling between different phonon modes, we approximate this Hamiltonian as

$$H_A \approx \sum_{q,\sigma}\left(-\frac{1}{2}\frac{\partial^2}{\partial p_{q,\sigma}^2} + V_{q,\sigma}(p_{q,\sigma})\right) = \sum_{q,\sigma} H_{q,\sigma} \qquad (2.71)$$

36
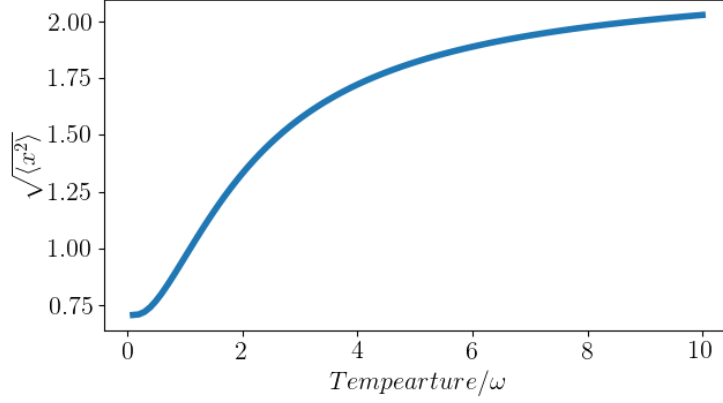
Figure 2.3: The RMS displacement for a harmonic oscillator as a function of temperature.

where

$$V_{q,\sigma}(p_{q,\sigma}) = V(R(0,0,\ldots,p_{q,\sigma},\ldots,0,0)) - V(R^{(0)}) \qquad (2.72)$$

is the so called *independent phonon* contribution to the Born-Oppenheimer energy [199]. The resulting independent phonon Hamiltonians, $H_{q,\sigma}$, can then be individually diagonalized in the basis of harmonic single phonon states $|n\rangle_{q,\sigma}$ defined via

$$\left(-\frac{1}{2}\frac{\partial^2}{\partial p_{q,\sigma}^2} + \frac{1}{2}\omega_{q,\sigma}^2 p_{q,\sigma}^2\right)|n\rangle_{q,\sigma} = \left(n + \frac{1}{2}\right)\omega_{q,\sigma}|n\rangle_{q,\sigma} \qquad (2.73)$$

This basis corresponds to the 1D harmonic states $\langle p|n\rangle = \psi_n(p)$ that appeared earlier in Eq. 2.69 (here p is a phonon coordinate, not to be confused with a momentum). The independent phonon matrix elements in this basis are calculated by explicitly mapping the independent phonon potential $V_{q,\sigma}(p_{q,\sigma})$ and carrying out the integration

$$H_{q,\sigma;n,m} = \langle n| H_{q,\sigma} |m\rangle = \int \psi_n^*(p)\left(-\frac{1}{2}\frac{\partial^2}{\partial p^2} + V_{q,\sigma}(p)\right)\psi_m^(p)dp \qquad (2.74)$$

The resulting eigenvalues of $H_{q,\sigma}$ are the anharmonic vibrational energies of that mode, which can be used to calculate anharmonic corrections to thermodynamic quantities.

2.6 Electron-phonon coupling in DFT

We determined how to calculate the properties of Kohn-Sham electrons and phonons independently in the previous two sections. However, in doing so we assumed the Born-Oppenheimer approximation. In order to calculate the effects of coupling between Kohn-Sham electrons and phonons, we must consider leading-order corrections to the Born-Oppenheimer approximation in nuclear displacements. Expanding

our Kohn-Sham potential in terms of these displacements leads to

$$V_{KS}(R + \delta R) = V_{KS}(R) + \sum_{\kappa,p} \frac{\partial V_{KS}}{\partial R_{\kappa,p}} \cdot \delta R_{\kappa,p} + O(\delta R^2). \tag{2.75}$$

Because phonons typically break the translational symmetry of the crystal (i.e they can not be contained within a single unit cell of the unperturbed crystal), we must consider a larger supercell (that *is* periodic) containing N_p unperturbed unit cells. In the above expression, $R_{\kappa,p}$ is then the position of atom κ in the pth unit cell. Displacements of this atom can then be written in terms of phonon creation and annihilation operators [102] as

$$\delta R_{\kappa,p} = \frac{1}{\sqrt{N_p M_\kappa}} \sum_{q\nu} e^{iq \cdot R_p} \frac{1}{\sqrt{2\omega_{q\nu}}} \left(a_{q\nu} + a_{-q\nu}^\dagger \right) e_{\kappa\nu}(q) \tag{2.76}$$

where $e_{k\nu}(q)$ and $\omega_{q,\nu}$ are, respectively, the eigenvector and frequency of the phonon mode with creation operator $a_{q\nu}^\dagger$, R_p is the origin of the pth unit cell and M_κ is the mass of atom κ. Substituting this into Eq. 2.75 we obtain

$$V_{KS}(R + \delta R) = V_{KS}(R) + \frac{1}{\sqrt{N_p}} \sum_{q\nu} G_{q\nu}(a_{q\nu} + a_{-q\nu}^\dagger) \tag{2.77}$$

where

$$G_{q\nu} = \frac{1}{\sqrt{2\omega_{q\nu}}} \sum_\kappa \frac{e_{\kappa\nu}(q)}{\sqrt{M_\kappa}} \cdot \sum_p e^{iq \cdot R_p} \frac{\partial V_{KS}}{\partial R_{\kappa,p}} \tag{2.78}$$

This allows us to write down the resulting electron-phonon coupling Hamiltonian in second-quantized form as

$$\begin{aligned}
H_{ep}(\delta R) &= \sum_{nkn'k'} \langle n, k| V_{KS}(R + \delta R) - V_{KS}(R) |n', k'\rangle c_{nk}^\dagger c_{n'k'} \\
&= \frac{1}{\sqrt{N_p}} \sum_{q\nu} \left[\sum_{nkn'k'} \langle n, k| G_{q\nu} |n'k'\rangle c_{nk}^\dagger c_{n',k'} \right] (a_{q\nu} + a_{-q\nu}^\dagger)
\end{aligned} \tag{2.79}$$

Where $c_{nk}^\dagger$ creates a Kohn-Sham electron in orbital n, wavevector k (i.e occupies the Bloch state $u_{nk}(x) \exp(ik \cdot x)/\sqrt{N_p}$). Substituting our definition of $G_{q\nu}$ we have

$$\langle n, k| G_{q\nu} |n', k'\rangle = \frac{1}{\sqrt{2\omega_{q\nu}}} \sum_\kappa \frac{e_{\kappa\nu}(q)}{\sqrt{M_\kappa}} \cdot \sum_p e^{iq \cdot R_p} \langle n, k| \frac{\partial V_{KS}}{\partial R_{\kappa,p}} |n', k'\rangle \tag{2.80}$$

Now

$$\begin{aligned}
\langle n, k| \frac{\partial V_{KS}}{\partial R_{\kappa,p}} |n', k'\rangle &= \int N_p^{-1/2} u_{nk}^*(x) e^{-ik \cdot x} \frac{\partial V_{KS}}{\partial R_{\kappa,p}}(x) N_p^{-1/2} u_{n'k'}(x) e^{ik' \cdot x} \, dx \\
&= \int N_p^{-1/2} u_{nk}^*(x - R_p) e^{-ik \cdot (x - R_p)} \frac{\partial V_{KS}}{\partial R_{\kappa,p}}(x - R_p) \\
&\quad \times N_p^{-1/2} u_{n'k'}(x - R_p) e^{ik' \cdot (x - R_p)} \, dx \\
&= e^{iR_p \cdot (k - k')} \int_{1^{\text{st}} \text{ unit-cell}} u_{nk}^*(x) e^{-ik \cdot x} \frac{\partial V_{KS}}{\partial R_{\kappa,0}}(x) u_{n'k'}(x) e^{ik' \cdot x} \, dx
\end{aligned} \tag{2.81}$$

where in the last line we have used Bloch's theorem and the fact that

$$\frac{\partial V_{KS}}{\partial R_{\kappa,p}}(x - R_p) = \frac{\partial V_{KS}}{\partial R_{\kappa,0}}(x) \tag{2.82}$$

where $R_{\kappa,0}$ is the position of atom κ in the first unit cell. we may now write Eq. 2.80 as

$$\langle n, k | G_{q\nu} | n'k' \rangle = \frac{1}{\sqrt{2\omega_{q\nu}}} \sum_{\kappa} \frac{e_{\kappa\nu}(q)}{\sqrt{M_\kappa}} \cdot \langle n, k | \frac{\partial V_{KS}}{\partial R_{\kappa,0}} | n', k' \rangle_{\text{uc}} \underbrace{\sum_{p} e^{i(q+(k-k'))\cdot R_p}}_{N_p \delta_{q,k-k'}} \tag{2.83}$$

where the subscript "uc" on the ket here means integration only over the first unit cell. Finally we obtain the DFT electron-phonon coupling Hamiltonian

$$H_{ep} = \frac{1}{\sqrt{N_p}} \sum_{q\nu knm} \langle m, k+q | G_{q\nu,\text{uc}} | n, k \rangle_{\text{uc}} c^\dagger_{m,k+q} c_{n,k}(a_{q\nu} + a^\dagger_{-q\nu}) \tag{2.84}$$

where we have defined

$$G_{q\nu,\text{uc}} = \frac{1}{\sqrt{2\omega_{q\nu}}} \sum_{\kappa} \frac{e_{\kappa\nu}(q)}{\sqrt{M_\kappa}} \cdot \frac{\partial V_{KS}}{\partial R_{\kappa,0}} \tag{2.85}$$

This allows us to write down the Hamiltonian for an interacting Kohn-Sham-electron-phonon system, correct to first order in the electron-phonon coupling constants $g_{mn\nu}(k, q) = \langle m, k+q | G_{q\nu,\text{uc}} | n, k \rangle_{\text{uc}}$:

$$H = \underbrace{\sum_{kn} \epsilon_{nk} c^\dagger_{nk} c_{nk}}_{\text{Electronic dispersion}} + \underbrace{\sum_{q\nu} \omega_{q\nu} \left(a^\dagger_{q\nu} a_{q\nu} + \frac{1}{2} \right)}_{\text{phonon dispersion}} + \\ \underbrace{\frac{1}{\sqrt{N_p}} \sum_{kqmn\nu} g_{mn\nu}(k, q) c^\dagger_{m,k+q} c_{nk} \left(a_{q\nu} + a^\dagger_{-q\nu} \right)}_{\text{electron-phonon coupling}} . \tag{2.86}$$

2.6.1 Superconductivity

The Electron-Phonon coupling term in Eq. 2.86 is formed of two contributions, corresponding to the Feynman diagrams shown in Figs. 2.4(a) and 2.4(b)

$$H_{\text{el-ph}} = \frac{1}{\sqrt{N_p}} \sum_{kqmn\nu} \underbrace{g_{mn\nu}(k, q) c^\dagger_{m,k+q} c_{nk} a_{q\nu}}_{\text{Phonon absorption}} + \underbrace{g_{mn\nu}(k, q) c^\dagger_{m,k+q} c_{nk} a^\dagger_{-q\nu}}_{\text{Phonon emission}} . \tag{2.87}$$

The resulting (lowest order) effective electron-electron interaction is then due to the virtual phonon exchange shown in figure 2.4(c). Applying the momentum-space Feynman rules, we obtain the phonon-mediated electron-electron interaction

$$H_{\text{el-ph-el}} = \sum_{\substack{k_1 nn' \\ k_2 mm' \\ q\nu}} V_{\text{el-ph-el}}(k_1, k_2, q)_{nn'mm'} c^\dagger_{k_1+q,n'} c_{k_1,n} c^\dagger_{k_2-q,m'} c_{k_2,m} \tag{2.88}$$

(a) Phonon absorption

(b) Phonon emission
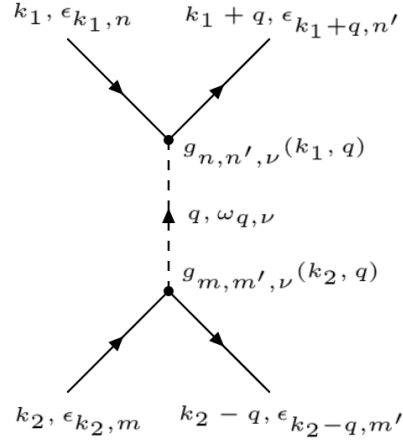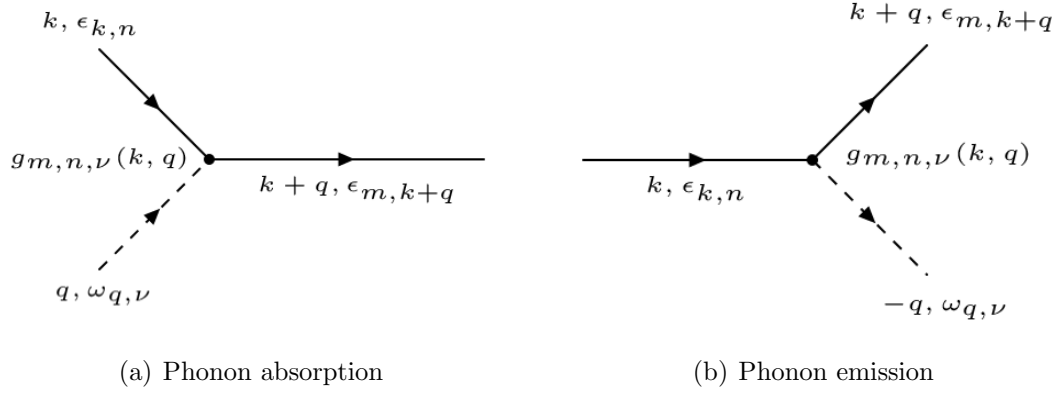
(c) The effective phonon-mediated electron-electron interaction.

Figure 2.4: Feynman diagrams from $H_{\text{el-ph}}$.

where the effective electron-electron interaction $V_{\text{el-ph-el}}(k_1, k_2, q)_{nn'mm'}$ is given by

$$V_{\text{el-ph-el}}(k_1, k_2, q)_{nn'mm'} = g_{nn'\nu}(k_1, q)g_{mm'\nu}(k_2, q)\frac{\omega_{q,\nu}}{(\epsilon_{k_1,n} - \epsilon_{k_1+q,n'})^2 - \omega_{q,\nu}^2} \quad (2.89)$$

At temperature T the accessible region of unoccupied final states will lie within $\sim k_b T$ of the Fermi surface. This means that, in the expression above, we will have $(\epsilon_{k_1,n} - \epsilon_{k_1+q,n'})^2 \sim (k_b T)^2$. At temperatures of interest for superconductors $\langle \omega_{q,\nu} \rangle >> k_b T$ and so $V_{\text{el-ph-el}}(k_1, k_2, q)_{nn'mm'} < 0$ is an attractive interaction between the states at k_1 and k_2.

2.6.2 BCS theory

The Bardeen-Cooper-Schrieffer (BCS) theory of superconductivity [24] was the first widely-successful theory explaining the origin of superconductivity from the attractive interaction given in Eq. 2.89. Whilst this is not the theory of superconductivity that we use later in this thesis to derive critical temperatures, it provides significant insight and demonstrates several important approximations often employed in the field. We start by considering the Hamiltonian for a metal where only one band (indexed by k) lies within $k_b T$ of the Fermi level and is subject to the phonon-mediated electron-electron interaction arising from a single phonon band (indexed by q);

$$H_{\text{BCS}} = \sum_k \epsilon(k)c_k^\dagger c_k + \sum_{k_1,k_2,q} V(k_1, k_2, q)c_{k_1+q}^\dagger c_{k_1} c_{k_2-q}^\dagger c_{k_2} \quad (2.90)$$

where $c_k^\dagger(c_k)$ creates (annihilates) an electron in the band at k with energy $\epsilon(k)$ and $V(k_1, k_2, q)$ is our phonon-mediated electron-electron interaction:

$$V(k_1, k_2, q) = g(k_1, q)g(k_2, q)\frac{\omega_q}{(\epsilon(k_1) - \epsilon(k_1+q))^2 - \omega_q^2} \quad (2.91)$$

As noted previously, at the temperatures of interest for superconductivity $\langle \omega_q \rangle \gg |\epsilon(k_1) - \epsilon(k_1+q)| \sim k_b T$ so this will be an attractive interaction for typical virtual phonons. BCS[6] used this observation to motivate the following approximate form for V:

$$V_{\text{BCS}} = \begin{cases} -|g_{\text{eff}}|^2 & \text{if } k_1, k_2, k_1 + q, k_2 - q \in S \\ 0 & \text{Otherwise} \end{cases} \quad (2.92)$$

Where $S = \{k : \epsilon_F \leq \epsilon(k) \leq \epsilon_F + \langle \omega_q \rangle\}$ is the thin shell of states just above the Fermi level that are energetically accessible via a kick from the average phonon. The momentum-dependence of the matrix elements has also been neglected, motivated by

---

[6]There is some confusion in the literature about the form of V_{BCS}, with some using the constraint $|\epsilon(k) - \epsilon(k+q)| \leq \langle \omega_q \rangle$, motivated by Eq. 2.91 to ensure $V_{\text{BCS}} < 0$. In this thesis, I instead use the definition employed by Cooper ($k \in S$) [51] so that I do not need to re-normalize the pair binding energy with a post-hoc energy cutoff.
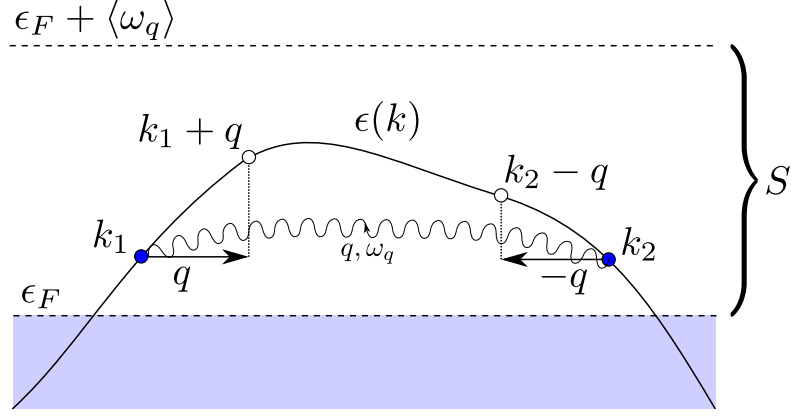
Figure 2.5: The Cooper problem. Two electronic states (k_1 and k_2) lying above a filled Fermi sea (blue rectangle) are coupled by a virtual phonon (q, ω_q). All states involved lie within $\langle \omega_q \rangle$ of the Fermi level (i.e in S), resulting in an attractive interaction.

the fact that only a small region of the relevant electronic band will lie within $k_b T$ of the Fermi level, within which the matrix elements should not vary strongly. We then have $g_{\text{eff}} \sim \langle g^2(k_F, q)/\omega_q \rangle$ where k_F corresponds to a representative electronic state near the Fermi level and we average over our phonon band. With this approximation, our Hamiltonian Eq. 2.90 can be written as

$$H_{\text{BCS}}^{(\text{eff})} = \sum_k \epsilon(k) c_k^\dagger c_k - |g_{\text{eff}}|^2 \sum_{k_1, k_2, q} c_{k_1+q}^\dagger c_{k_1} c_{k_2-q}^\dagger c_{k_2} \tag{2.93}$$

Where the momenta in the second sum are restricted such that $k_1, k_2, k_1+q, k_2-q \in S$.

The Cooper problem

Cooper [51] considered the important case of two electrons outside a filled Fermi sea, evolving according to $H_{\text{BCS}}^{(\text{eff})}$ (as shown in Fig. 2.5). The wavefunction of these two electrons can be expanded in the basis of Slater determinants of Kohn-Sham states:

$$|\psi\rangle = \sum_{k_1, k_2 \in I} C_{k_1, k_2} |k_1, k_2\rangle \tag{2.94}$$

where $I = \{k : \epsilon(k) > \epsilon_F\}$ are the states above the Fermi sea and

$$\langle r_1, r_2 | k_1, k_2 \rangle = \begin{vmatrix} \psi_{k_1}(r_1) & \psi_{k_1}(r_2) \\ \psi_{k_2}(r_1) & \psi_{k_2}(r_2) \end{vmatrix}. \tag{2.95}$$

are our Slater determinants of single-particle Kohn-Sham states $\psi_{k_i}(r_j)$. It will turn out that the lowest energy state $|\psi\rangle$ will have no centre-of-mass motion and so only

states with total momentum $k_1 + k_2 = 0$ will contribute, allowing us to index the Slater determinants with a single momentum $k = k_1 = -k_2$:

$$|\psi\rangle = \sum_{k \in I} C_k |k, -k\rangle. \tag{2.96}$$

The Schrödinger equation for this state then reads

$$H_{\text{BCS}}^{(\text{eff})} |\psi\rangle = E |\psi\rangle = \sum_{k \in I} C_k 2\epsilon(k) |k, -k\rangle - |g_{\text{eff}}|^2 \sum_{\substack{k \in S \\ q : k+q \in S}} C_k |k+q, -k-q\rangle, \tag{2.97}$$

where the first term is the contribution arising from the bare Kohn-Sham states (using the fact that $\epsilon(k) = \epsilon(-k)$). The second term arises from the scattering of $|k, -k\rangle$ into $|k+q, -k-q\rangle$ via a virtual phonon with momentum q, subject to the constraint that the initial and final states are in S (see Fig. 2.5). Because the sum over q runs over the entire first B.Z, the second sum in Eq. 2.97 can instead be taken over final states $k' = k + q \in S$, leading to

$$E |\psi\rangle = \sum_{k \in I} C_k 2\epsilon(k) |k, -k\rangle - |g_{\text{eff}}|^2 \sum_{k, k' \in S} C_k |k', -k'\rangle. \tag{2.98}$$

Taking the inner product with $\langle k, -k|$ (for $k \in S$)[7], we obtain

$$E C_k = 2\epsilon(k) C_k - |g_{\text{eff}}|^2 \sum_{k \in S} C_k. \tag{2.99}$$

Defining $A = \sum_{k \in S} C_k$, we can re-arrange this for C_k:

$$C_k = \frac{-|g_{\text{eff}}|^2 A}{E - 2\epsilon(k)} \tag{2.100}$$

For self-consistency we must then have (from the definition of A):

$$A = \sum_{k \in S} \frac{-|g_{\text{eff}}|^2 A}{E - 2\epsilon(k)} \implies 1 = -|g_{\text{eff}}|^2 \sum_{k \in S} \frac{1}{E - 2\epsilon(k)}. \tag{2.101}$$

Replacing the sum over k with an integral over the density of states $g(\epsilon)$ (which we assume to be constant for $k \in S$) gives

$$1 = -|g_{\text{eff}}|^2 g(\epsilon_F) \int_{\epsilon_F}^{\epsilon_F + \langle \omega_q \rangle} \frac{1}{E - 2\epsilon} d\epsilon. \tag{2.102}$$

Where the prefactor $\lambda = |g_{\text{eff}}|^2 g(\epsilon_F)$ is the so-called *electron-phonon coupling parameter*, and is assumed to be small within BCS theory. Carrying out the integral and rearranging gives

$$E = 2\epsilon_F - \frac{2\langle \omega_q \rangle}{\exp(2/\lambda) - 1} \underset{\lambda \ll 1}{\approx} 2\epsilon_F - 2\langle \omega_q \rangle \exp(-2/\lambda). \tag{2.103}$$

---

[7]If we had instead taken the inner product with a state outside of S we would have seen that the energy was just the unaltered KS energy. This is because BCS Hamiltonian neglects the renormalization of the KS energies due to the electron-phonon interaction.
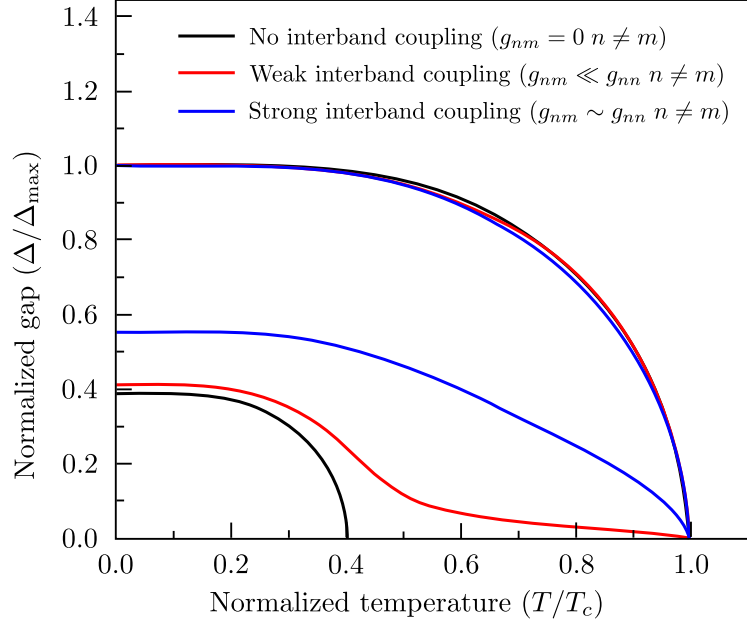
Figure 2.6: Temperature dependence of the superconducting gaps in a multi-band superconductor (intraband coupling constants are those for MgB_2 from [269]).

This energy is lower than the energy of two electrons at the Fermi level by an amount given by

$$\Delta = 2\langle\omega_q\rangle\exp(-2/\lambda). \tag{2.104}$$

BCS realised that this means the entire Fermi surface will be unstable to the formation of bound pairs of electrons with binding energy Δ. This bound state is called a *Cooper pair* and an estimate for the critical temperature of a superconductor is given by the thermal energy needed to break such pairs

$$k_b T_c = \Delta \implies T_c = \frac{2\langle\omega_q\rangle}{k_b}\exp(-2/\lambda) \tag{2.105}$$

Which gives reasonable agreement with weakly coupled ($\lambda \ll 1$) conventional superconductors. The agreement can be improved by noting that when the entire Fermi surface becomes unstable to the formation of Cooper pairs, we are in actual fact dealing with a many-electron problem, rather than a two-electron problem. Solving the many-body BCS problem at finite temperature leads to

$$T_c = \frac{1.13\langle\omega_q\rangle}{k_b}\exp(-1/\lambda) \tag{2.106}$$

Which is remarkably similar to Cooper's result (especially in the limit $\lambda \to 0$).

44

2.6.3 Approximations in BCS theory

Extension to the multi-band case

In our discussion of BCS theory, we only considered coupling within a single electronic band mediated by phonons from a single mode. Including additional phonon modes is relatively straightforward, amounting to summing contributions to $|g_{\text{eff}}|^2$ from each mode. However, the generalisation to multiple electronic bands is more involved. Each band that crosses the Fermi level will in general have different electronic character and so will couple to the phonons differently. This will result both in different intra-band interactions $g_{mm}(k_1, k_2, q)$, but also in *inter*-band couplings $g_{mn}(k_1, k_2, q)$ with $n \neq m$. Neglecting inter-band coupling would result in each band acting as an independent sub-system with its own T_c - the most significant being that with the highest T_c. Systems where inter-band coupling cannot be neglected are better described by a single T_c, with a non-trivial dependence on the mixing of coherent states from each band (this is the case in the multi-band superconductor MgB_2, for example, as shown in Fig. 2.6).

Renormalization and retardation

The approximation of $V(k_1, k_2, q)$ by the BCS potential V_{BCS} neglects electron-phonon coupling for all but a very thin shell of states near the Fermi surface. The result is that the single-particle electronic energies $\epsilon(k)$ remain unaltered outside of this shell. In reality, electron-phonon coupling will renormalize the electronic energies and, consequently, the Fermi surface itself. This argument goes the other way too; electron-phonon coupling will also renormalize the phonon frequencies. As well as neglecting this renormalization of the single-particle states, BCS theory only crudely accounts for retardation effects due to the 'sluggishness' of the phonon response (by cutting off the electron-phonon interaction above $\langle \omega_q \rangle$). The retardation of the phonon response allows two electrons to exchange a phonon whilst not being at the same place at the same time. As a result of this separation, the two electrons feel a reduced Coulomb repulsion from one another. This will (reasonably directly) increase the pair binding energy and thus the critical temperature, so including the effects of retardation correctly is key to understanding the strong-binding/high-T_c case.

2.6.4 Eliashberg theory

As motivated in the previous section, the correct treatment of retardation in the phonon response is a crucial step towards understanding strongly-coupled superconductors. Extending the BCS theory to include these retardation effects is the

goal of Eliashberg theory [72], which derives the temperature-dependent imaginary-time Green's function for electrons evolving under a Hamiltonian with first-order electron-phonon coupling (Eq. 2.86). The full derivation [188] is reasonably involved and does not provide a large amount of additional insight beyond BCS theory (at least for the purposes of this thesis), so we will simply summarise the main results. Roughly speaking, one uses a Wick decomposition to re-write the electron Green's function in terms of a normal-state Green's function G, corresponding to the time-ordered thermodynamic expectation $G(k, \tau) = -\langle T_\tau c_k(\tau) c_k^\dagger(0) \rangle$ and an anomalous (superconducting-pair) Green's function $F(k, \tau) = -\langle T_\tau c_k(\tau) c_k(0) \rangle$. In practice, it is convenient to move from the imaginary-time (τ) domain to the imaginary-frequency ($i\omega$) domain. Considering just the contribution from first-order Feynman diagrams, and confining all quantities to lie on the Fermi surface, the normal and anomalous Green's functions can be determined from the following pair of coupled equations, known as the *Eliashberg equations*:

$$[1 - Z(i\omega_n)] i\omega_n = -\frac{\pi}{\beta} \sum_m \frac{Z(i\omega_m) i\omega_m}{\Xi(i\omega_m)} \int \frac{2\omega \alpha^2 F(\omega)}{(\omega_n - \omega_m)^2 + \omega^2} d\omega$$

$$\phi(i\omega_n) = -\mu^* \frac{\pi}{\beta} \sum_m \frac{\phi(i\omega_m)}{\Xi(i\omega_m)} \theta(\omega_c - |\omega_m|) + \frac{\pi}{\beta} \sum_m \frac{\phi(i\omega_m)}{\Xi(i\omega_m)} \int \frac{2\omega \alpha^2 F(\omega)}{(\omega_n - \omega_m)^2 - \omega^2} d\omega$$

$$(2.107)$$

where

$$\Xi(i\omega_n) = \sqrt{[Z(i\omega_n)\omega_n]^2 + \phi^2(i\omega_n)} \tag{2.108}$$

In these equations, Z (proportional to the normal Green's function) and ϕ (proportional to the anomalous Green's function) are the only unknowns and are uniquely specified via Eqs. 2.107 from the parameters in our Hamiltonian. These parameters enter Eqs. 2.107 via the so-called *Eliashberg function* $\alpha^2 F(\omega)$:

$$\alpha^2 F(\omega) = \frac{1}{2\Omega_{\text{BZ}}} \sum_\nu \int_{\text{BZ}} \omega_{q\nu} \lambda_{q\nu} \delta(\omega - \omega_{q\nu}) dq \tag{2.109}$$

where

$$\lambda_{q,\nu} = \frac{1}{N(\epsilon_F)\omega_{q\nu}\Omega_{\text{BZ}}} \sum_{nm} \int_{\text{BZ}} |g_{mn\nu}(k,q)|^2 \delta(\epsilon_{n,k} - \epsilon_F) \delta(\epsilon_{m,k+q} - \epsilon_F) dk \tag{2.110}$$

are the electron-phonon coupling strengths associated with each phonon mode. In practice, given $\alpha^2 F(\omega)$, it is relatively straightforward to numerically solve Eqs. 2.107 self-consistently to obtain Z and ϕ, from which T_c is obtained as the temperature where the pair-density $\phi \to 0$ (to do this I use the ELK code [1], who's notation [232] we employ in Eqs. 2.107). However, it is common in the field to instead use the approximation due to McMillan, Allen and Dynes [192, 11]:

$$T_c = \frac{\omega_{\log}}{1.2} \exp\left(\frac{-1.04(1 + \lambda)}{\lambda(1 - 0.62\mu^*) - \mu^*}\right) \tag{2.111}$$

where

$$\lambda = \sum_{q\nu} \lambda_{q\nu} = 2 \int \alpha^2 F(\omega) \frac{d\omega}{\omega}, \tag{2.112}$$

is the electron-phonon coupling parameter and

$$\omega_{\log} = \exp\left(\frac{2}{\lambda} \int \alpha^2 F(\omega) \log(\omega) \frac{d\omega}{\omega}\right) \tag{2.113}$$

is a particular choice for the typical phonon frequency $\langle \omega_q \rangle$. μ^* is the Morel-Anderson pseudopotential [200], which is typically treated as an empirical parameter with values between 0.1 and 0.2.

2.7   *Ab initio* random structure searching (AIRSS)

2.7.1   The global searching problem

Two factors determine if a particular arrangement of atoms in a crystalline structure is observed in nature. The first is *dynamic stability* and is a necessary condition (although dynamical instabilities can sometimes disappear when moving to a higher level of theory, see chapter 8). Dynamic stability means that there is no infinitesimal perturbation to the crystal structure that will result in a reduction in the energy $\implies$ the structure is a local minima of the energy. The second factor is thermodynamic preference towards lower energy structures. In particular, at zero temperature and neglecting dynamical effects (zero-point energy), the most thermodynamically stable configuration is the global minimum of the energy. Working within the Born-Oppenheimer approximation, we would like to be able to find such low energy minima of the Born-Oppenheimer surface $E_{\text{elec}}(R)$.

When calculating the total DFT energy, $E_{\text{elec}}(R)$ of a particular crystal structure R, we also compute the DFT ground state $|\psi\rangle$ (typically in a self-consistent manner). The Hellmann-Feynman theorem gives us a way to go on to cheaply calculate the force $F = \partial\langle H \rangle / \partial R$ on the atoms:

$$
\begin{aligned}
\frac{\partial \langle H \rangle}{\partial R} &= \langle \partial\psi/\partial R | \, H \, |\psi\rangle + \langle\psi| \, \partial H/\partial R \, |\psi\rangle + \langle\psi| \, H \, |\partial\psi/\partial R\rangle \\
&= E(\langle \partial\psi/\partial R \, |\psi\rangle + \langle\psi| \, \partial\psi/\partial R\rangle) + \langle\psi| \, \partial H/\partial R \, |\psi\rangle \\
&= E\frac{\partial}{\partial R} \underbrace{\langle \psi \, |\psi\rangle}_{1} + \langle\psi| \, \partial H/\partial R \, |\psi\rangle \\
&= \langle\psi| \, \partial H/\partial R \, |\psi\rangle = F
\end{aligned}
\tag{2.114}
$$

We simply take the expectation value of the derivative of the Hamiltonian with respect to the state we've already worked out. Because we can cheaply evaluate the force, we can efficiently minimize the energy using gradient descent. This gives us a

way of finding *local* minima in the Born-Oppenheimer energy surface, which correspond to dynamically stable crystal structures. We call this *geometry optimization.*

The search for minima in the Born-Oppenheimer surface has an intrinsic problem: the number of local minima scales exponentially with system size. This can be seen from a simple argument given in [220] and reproduced here; consider a system of N atoms divided into M subsystems, each of N/M atoms. If the subsystems are large enough they will have independent stable configurations. The total number of stable configurations $n_s(N)$ therefore satisfies

$$n_s(N) = n_s^M(N/M) \tag{2.115}$$

The solution to which is

$$n_s(N) \propto e^{\alpha N} \tag{2.116}$$

This exponential scaling means that in order to efficiently study large systems we must reduce our search space *a priori*, by somehow biasing the search towards lower energy regions of the Born-Oppenheimer surface. This can be done in several ways, amongst which are Genetic algorithms [262, 207, 144, 180], Simulated annealing [58, 138], Basin hopping [272], Particle swarms [276], Data mining [113] and Random sampling [220]. In the next section, we define the method that is used in this thesis, based on random sampling.

2.7.2 AIRSS

Leveraging large multiprocessor computers, we can carry out many geometry optimizations simultaneously to enumerate the minima of the Born-Oppenheimer surface. It turns out that if we begin each of these geometry optimizations at a random sensible location in configuration space, we can typically recover many of the lowest-energy minima (those which are likely to correspond to thermodynamically-stable structures). When we say *sensible* random configurations, we mean configurations chosen to satisfy physically motivated constraints. Some examples of these constraints are

- Physical atom-atom separations (atoms not overlapping).

- Atoms arranged as structural units (i.e crystalline water arranged using H_2O molecules).

- Preference towards higher crystalline symmetry.

- Incorporation of experimental information.

These choices are based on the following general properties of the Born-Oppenheimer surface [220]:

- The portion of the Born-Oppenheimer surface where atoms are very close together contains almost no minima.

- This high-energy part of the Born-Oppenheimer surface forms a large fraction of the total surface.

- Minima with lower energies tend to have larger basins of attraction (in general the basin size follows a power law of the energy).

- Minima tend to correspond to more symmetrical structures.

We see that, simply by requiring that atoms are sensibly far apart, we can reduce the fraction of the surface that must be searched substantially. Combined with the fact that the remaining minima will tend to have large basins of attraction, and will tend to be symmetric, we can efficiently find the lowest energy minima. The method of generating random, sensible structures and relaxing to a nearby local minimum using DFT is known as *ab initio* random structure searching (AIRSS) [220].

Chapter 3

Stochastic nodal surfaces in quantum Monte Carlo calculations

3.1 Motivation

Quantum Monte Carlo (QMC) methods were responsible for the data underpinning all DFT calculations [40] and remain amongst the most accurate methods available for calculating ground state properties of quantum systems [186]. However, as we saw in Sec. 2.3.3, for certain systems QMC suffers from the infamous *fermion sign problem*, the general solution to which has been shown to be NP-hard [266]. In this thesis, we focus on the specific case of the *diffusion Monte Carlo* (DMC) methods examined in Sec. 2.3.2 [90, 21, 263], which converge on the many-body ground state by iteratively projecting out exited state components from the wavefunction. Here, the sign problem arises due to exchange symmetry dividing the wavefunctions into regions of different sign, known as *nodal pockets*, separated by a *nodal surface*. This increases the fermionic ground state energy relative to that of the bosonic ground state and, as a result, the former is projected out, typically exponentially decaying away as the iterative procedure progresses [20].

Despite this exponential decay of the fermionic component, methods such as *release-node* DMC can extract information about the fermionic ground state from the transient behaviour of the wavefunction [41]. However, this transient behaviour leads to a statistical error that grows with system size, requiring a formidable computational effort to mitigate [268]. The most popular approach to obtain a stable fermionic ground state in DMC is the *fixed-node approximation* that we discussed in Sec. 2.3.3, whereby the nodal surface is fixed to that of some trial wavefunction, which must be known *a priori*. We focus on electronic systems, where it is conjectured that the presence of many-body correlation leads to the minimal case of only two nodal pockets [196], which may make the electronic problem more tractable

than the general NP-hard case.

The work presented in this chapter was, in part, motivated by my own frustrations with carrying out DMC calculations of solid Lithium - in particular the fact that the majority of my time was spent generating, optimizing and transferring trial wavefunctions from preliminary DFT calculations to aid the main DMC calculation. I wanted to better understand the sign problem as it applies to the electronic systems that would typically be treated using the fixed node approximation. In particular, to understand more deeply what goes wrong if we do not stabilise the simulation with a trial wavefunction and to investigate the possibility of alternative methods of stabilization that do not require *a priori* knowledge of the solution.

3.2   Overview

In the rest of this chapter, we develop a formalism of fermionic DMC that makes no reference to a trial wavefunction. In section 3.3 we show that including exchange symmetry as a constraint in the energy minimization problem leads to a modified DMC scheme, resulting in a new propagation channel in the Green's function that couples populations of signed walkers. In section 3.4 we go on to show how this propagation results in the formation of a *stochastic* nodal surface, constructed from the entire population of walkers, that is free to vary and minimize the energy. We propose a diffusion scheme to maximise its stability.

Compared to previous methods employing signed walkers (such as *the fermion Monte Carlo* method of Kalos and Pederiva [134], the early work of Anderson and Traynor [14] or the second-quantized approach of Umrigar [270]) the work in this chapter represents an alternative way to manage the propagation, and cancellations between, signed walkers, informed by the underlying optimization problem. In section 3.5 we discuss the relationship to these previous schemes. Finally, we provide an open-source implementation of the method [2] and demonstrate that it obtains a stable fermionic ground state for the harmonic and atomic systems considered.

3.3   Formalism

We start by formulating the fermionic problem for N particles in d dimensions as the following *constrained* optimization problem:

Find $|\psi\rangle$ to minimize $\langle\psi|\,H\,|\psi\rangle$ such that

$$\langle\psi\,|\psi\rangle = 1 \qquad \text{(Normalization)} \qquad (3.1)$$

$$\langle x\,|\psi\rangle = -\langle P_i x\,|\psi\rangle \qquad \text{(Antisymmetry)} \qquad (3.2)$$

$$\forall \ P_i \in \mathcal{E}, \ x \in \mathbb{R}^{dN}$$

where $\mathcal{E}$ is the set of pairwise identical-fermion exchanges. If the system contains M identical fermions, there are $M(M-1)/2$ such exchanges which can be combined to generate the set $\mathcal{P}$ of the $M!$ permutations of identical fermions (we discuss the alternative choice of using the permutations $\mathcal{P}$ to construct the antisymmetry constraint in Sec. 3.5.1). Introducing the Lagrange multipliers E_T and $\mu_i(x)$ the optimization problem is equivalent to extremizing the Lagrangian

$$\mathcal{L} = \underbrace{\langle\psi|\,H\,|\psi\rangle}_{\text{Objective function}} + E_T \underbrace{[1 - \langle\psi\,|\psi\rangle]}_{0 \text{ iff } \psi \text{ is normalized}} + \sum_i \int \psi^*(x)\mu_i(x) \underbrace{(P_i + 1)\psi(x)}_{0 \text{ iff } \psi \text{ is antisymmetric}} \ dx \quad (3.3)$$

with respect to ψ, ψ^*, E_T and the $\mu_i(x)$'s. We note that $\mathcal{L}$ can be written as

$$\mathcal{L} = E_T + \langle\psi|\,\overbrace{H - E_T + \sum_i \mu_i(x)(P_i + 1)}^{H_X}\,|\psi\rangle \quad (3.4)$$

allowing us to define an effective Hamiltonian H_X. Because $(P_i + 1)|S\rangle = 2|S\rangle$ for symmetric states $|S\rangle$ and $(P_i + 1)|A\rangle = 0$ for antisymmetric states $|A\rangle$, the term involving the Lagrange multipliers μ_i can be interpreted as a cost function that penalises the appearance of a symmetric component in the wavefunction (so long as we take $\mu_i(x) > 0$). Extremization of $\mathcal{L}$ with respect to ψ and ψ^* (see appendix A.5) leads to

$$H_X\psi = 0 = H\psi - E_T\psi + \left[\sum_i \mu_i(x)(P_i + 1)\right]\psi. \quad (3.5)$$

Because of the antisymmetry constraint, the term in square brackets vanishes at the extremum of $\mathcal{L}$, leading to the Schrödinger equation $H\psi = E_T\psi$. This allows us to identify E_T as the fermionic ground state energy.

To actually perform the extremization we follow the DMC method outlined in Sec. 2.3.2 and propagate the imaginary time ($\tau = it$) Schrödinger equation, but for H_X instead of H:

$$\frac{\partial\,|\psi(\tau)\rangle}{\partial\tau} = -H_X\,|\psi(\tau)\rangle \quad (3.6)$$

which can be written in integral form as

$$\underbrace{\langle x\,|\psi(\tau + \delta\tau)\rangle}_{\substack{\text{Propagated wavefunction} \\ \psi(x,\tau+\delta\tau)}} = \int \underbrace{\langle x|\exp(-\delta\tau H_X)|x'\rangle}_{\substack{\text{Green's function} \\ G(x,x',\delta\tau)}} \underbrace{\langle x'\,|\psi(\tau)\rangle}_{\substack{\text{Old wavefunction} \\ \psi(x',\tau)}} dx'. \quad (3.7)$$

Continuing to follow traditional DMC, we sample our wavefunction with a discrete set of walkers, each representing a particular point in configuration space x_i and carrying a corresponding weight w_i, leading to Eq. 2.21, repeated here:

$$\psi_{\text{DMC}}(x, \tau) = \sum_i w_i(\tau)\delta(x - x_i(\tau)). \quad (3.8)$$

Eq. 3.7 can then be interpreted as an evolution equation for the walkers, where the Green's function $G(x, x', \delta\tau)$ determines the propagation of the weights w_i and configurations x_i. In Sec. 2.3.2, we saw that we can choose how to partition the influence of the Green's function between the propagation of the walker configuration $x_i(\tau) \to x_i(\tau + \delta\tau)$ and the update procedure of the walker weights $w_i(\tau) \to w_i(\tau + \delta\tau)$. As was the case in Sec. 2.3.2, in order to determine how to distribute this influence, we need the explicit form for the Green's function. Writing $H = T + V$, where T is the kinetic energy operator and V is the (local) many-body potential, allows us to define the well-known potential and diffusive parts of the Green's function, derived in Sec. 2.3.2;

$$G_V(x, x', \delta\tau) \equiv \exp\left(-\delta\tau[V(x) + V(x')]/2\right),$$

$$G_D(x, x', \delta\tau) \equiv \langle x| \exp(-\delta\tau T) |x'\rangle \propto \exp\left(-\frac{|x - x'|^2}{2\delta\tau}\right).$$

For sufficiently small timesteps $\delta\tau \ll 1$, our full Green's function can then be written (see appendix A.6) as

$$G(x, x', \delta\tau) = \overbrace{\left[\mathcal{N}(x') - \sum_i \mathcal{X}_i(x') P_i\right]}^{G_X \Longleftrightarrow \text{Exchange moves}}$$

$$\times \underbrace{G_V(x, x', \delta\tau)}_{\text{Potential weighting}} \times \underbrace{G_D(x, x', \delta\tau)}_{\text{Diffusion}} \times \underbrace{\exp(\delta\tau E_T)}_{\text{Population control}}, \tag{3.9}$$

with

$$\mathcal{X}_i(x') = \delta\tau \mu_i(P_i x'),$$
$$\mathcal{N}(x') = 1 - \sum_i \delta\tau \mu_i(x'). \tag{3.10}$$

We note that if we were to neglect the fermionic constraint (i.e take $\mu_i \to 0 \ \forall \ i$), we would recover the Green's function of traditional DMC (see section 2.3.2). The part of the Green's function arising from the fermionic constraint is labelled G_X and can be applied to a walker at x' with weight w by carrying out the fermionic exchange $\{x' \to P_i x', w \to -w\}$ with probability $\mathcal{X}_i(x')$. These non-local *exchange moves* enforce the antisymmetry of the wavefunction by allowing walkers sampling one nodal pocket to stochastically switch to sampling any symmetry-related nodal pocket (see Fig. 3.1). The tiling theorem [39] then implies that any walker can access and contribute weight to all nodal pockets. As a result, rather than each walker simply contributing to the wavefunction at a particular point in configuration space, it can now contribute to all symmetry-related points.

For simplicity, in our implementation we choose the probabilities $\mathcal{X}_i(x')$ and $\mathcal{N}(x')$ so that each of the exchange moves (including no exchange) are equiprobable (i.e we take $\mathcal{X}_i(x') = \mathcal{N}(x') \ \forall \ i$). This corresponds to a particular positive,
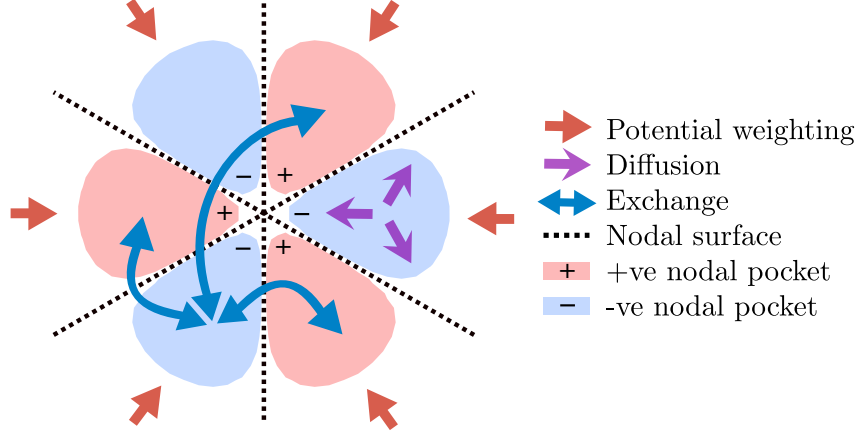
Figure 3.1: Schematic of wavefunction formation arising from competing walker propagation channels (shown for three fermions in a harmonic well as in Fig. 3.4).

and constant, choice of the Lagrange multipliers $\mu_i(x)$, which can be interpreted as the rate at which the exchange P_i is applied to walkers at x. Whilst this does result in a Hamiltonian that penalises symmetric components, strictly speaking $\mu_i(x)$ should be varied with τ to extremise $\mathcal{L}$ (in the same way that we will vary E_T to maintain a constant population in Sec. 3.4.3). This would correspond to biasing the simulation to carry out particular exchanges for walkers at particular configurations. One can see how this might work - regions in configuration space that have a large symmetric component could be specifically targeted for exchange in the corresponding coordinates. However, in practice it is very difficult to identify such regions and, in any case, any random exchange of particles will quickly wipe out symmetric components[1]. Without being able to identify which combinations of configurations and exchanges should be targeted, there is no reason to bias towards particular exchanges - hence our choice of equiprobable exchanges. As we will see later, this does not adversely affect the stability of the resulting fermionic state.

3.4 Implementation

3.4.1 Stochastic nodal surface

Now that we have identified an additional propagation channel arising from the antisymmetry constraint, we need to think about how it works alongside the propagation channels we already had, in particular the diffusive channel. The diffusive part of

---

[1]In actual fact random exchange will, on average, wipe out *all* components leading to $\langle \psi \rangle = 0$ (which we refer to later as bosonic collapse), but we will see later how this can be remedied with a correlated diffusion process.
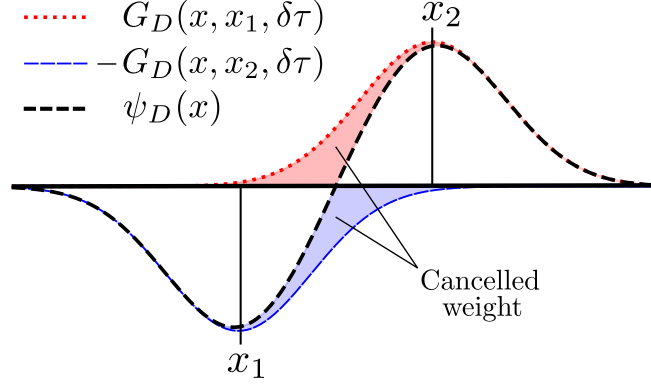
Figure 3.2: The diffusive propagation of two nearby walkers of opposite sign located at x_1 and $x_2 \implies \psi_D(x) = G_D(x, x_2, \delta\tau) - G_D(x, x_1, \delta\tau) \equiv G_2 - G_1$ (black dashed line). The red (blue) shaded region show the portion of G_2 (G_1) that can be cancelled in the propagation.

the Green's function applied to a set of walkers leads to a *diffused* wavefunction

$$\psi_D(x) = \sum_i w_i G_D(x, x_i, \delta\tau) \tag{3.11}$$

as shown in Fig. 3.2 for two opposite-sign walkers. If we represent this new wavefunction as a combination of walkers with weights ± 1 with configurations sampled from the distributions $P_\pm(x)$ respectively, we must have

$$P_+(x) - P_-(x) = \psi_D(x). \tag{3.12}$$

In traditional DMC each walker diffuses independently by an amount sampled from G_D, resulting in

$$
\begin{aligned}
P_+(x) = \psi_+(x) &\equiv \sum_{w_i > 0} w_i(\tau) G_D(x, x_i, \delta\tau) \\
P_-(x) = \psi_-(x) &\equiv \sum_{w_i < 0} |w_i(\tau)| G_D(x, x_i, \delta\tau).
\end{aligned}
\tag{3.13}
$$

A drawback of this scheme when applied to signed walkers is that it allows +ve walkers to move into a region where ψ_D is -ve, and vice versa, as can be seen from the overlap of $P_+(x)$ and $P_-(x)$ in Fig. 3.3(a). This prohibits the emergence of well-separated regions of +ve and -ve walkers, corresponding to nodal pockets. Without stable nodal pockets, the walkers end up sampling the bosonic ground state with a randomly fluctuating sign. This is known as *bosonic collapse* and arises in a similar fashion to the exponentially decaying signal-to-noise ratio in so-called *release-node* DMC [20]. An example is shown in Fig. 3.4(a) for a system of three non-interacting fermions in a harmonic well.
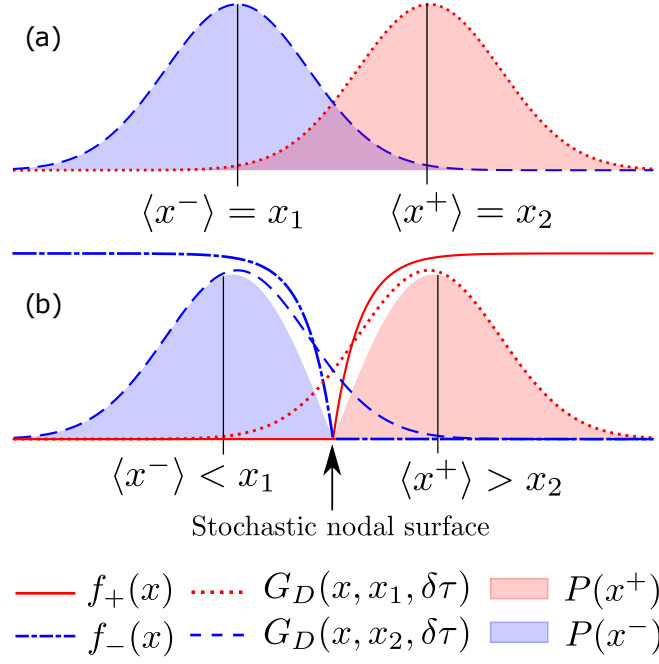
Figure 3.3: Propagation schemes satisfying Eq. 3.12, applied to the walkers in Fig. 3.2. (a) traditional DMC propagation (Eq. 3.13). (b) our propagation scheme (Eq. 3.15). Note that in (a) there is overlap of the +ve and -ve walker distributions. The same is not true for (b).

To avoid bosonic collapse, one particular sign of walker should dominate at each point in configuration space. Typically this sign is chosen according to the fixed-node approximation as being equal to that of the trial wavefunction (see Fig. 2.2), fixing the nodal surface of the DMC wavefunction to that of the trial wavefunction. We instead exploit the freedom to choose P_+ and P_- in Eq. 3.12 to design a propagation scheme that encourages the formation of a *stochastic* nodal surface (that is free to vary in order to minimize the energy). In order to do this, we seek the form of $P_\pm(x)$ satisfying Eq. 3.12 that maximizes the expected separation of +ve and -ve walkers, given by

$$\langle |x_+ - x_-| \rangle = \int P_+(x_+) P_-(x_-) |x_+ - x_-| dx_+ dx_- \qquad (3.14)$$

This leads, independently of the form of $\psi_D(x)$ (see appendix A.7), to

$$P_\pm(x) = \begin{cases} |\psi_D(x)| & \text{if } \text{sign}(\psi_D(x)) = \pm 1, \\ 0 & \text{otherwise.} \end{cases} \qquad (3.15)$$

These distributions have no overlap, as can be seen in Fig. 3.3(b). However, because $P_\pm(x)$ are no longer simple sums of Gaussian terms (as $\psi_\pm$ were in Eq. 3.13), they are difficult to sample moves from directly. This can be remedied by exploiting our freedom to absorb part of the propagation into the weight update by factorising
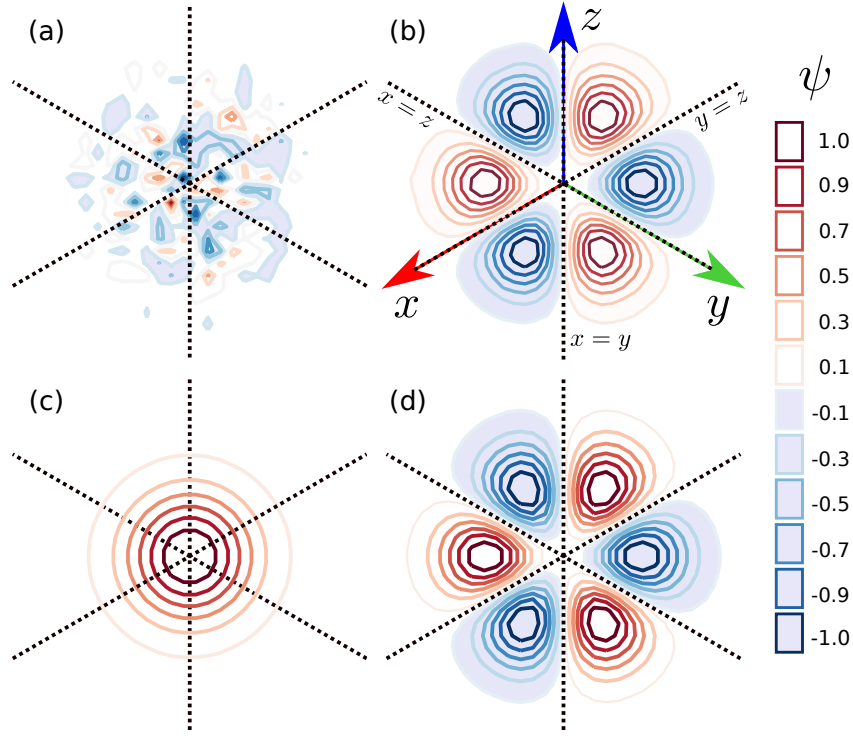
Figure 3.4: The wavefunction of three non-interacting fermions with coordinates x, y and z in a one-dimensional harmonic well, integrated and viewed along the $(1,1,1)$ projection. The analytic nodal surface is shown as a dotted black line. From this projection, the nodal pockets can be clearly seen. (a) Bosonic collapse from DMC with exchange moves but without a stochastic nodal surface. (b) From DMC with exchange moves and a stochastic nodal surface. (c) Analytic bosonic ground state. (d) Analytic fermionic ground state.

$P_\pm(x)$ into (see appendix A.8)

$$P_\pm(x) = \psi_\pm(x) f_\pm(x) \tag{3.16}$$

where

$$f_\pm(x) = \max\left(1 - \psi_\mp(x)/\psi_\pm(x),\ 0\right) \in [0, 1] \tag{3.17}$$

can be interpreted as reweighting functions, as shown in Fig. 3.3(b). We can then interpret Eq. 3.16 as a diffusion according to $\psi_\pm(x)$ (corresponding to diffusive moves sampled from $G_D(x, x', \delta\tau)$) followed by a corrective reweighting $w \to f_\pm(x)w$, due to cancellation of +ve and -ve walkers.

Applying this scheme to the same system of three non-interacting fermions in a harmonic well results in the wavefunction shown in Fig. 3.4(b). Comparing to Fig. 3.4(d) we see that the analytic nodal surface is reproduced.

3.4.2 Initialization

We initialize the walkers in such a way as to speed up their equilibration into an antisymmetric state. This is achieved by defining a unique ordering of the walker configurations, whereby the particles are ordered by their increasing coordinates. For example, in 2 spatial dimensions, the particles are ordered first by increasing x coordinate, then by increasing y coordinate. Starting with walker configurations distributed according to a normal distribution $x_i \sim \mathcal{N}(\mu = 0, \sigma = 1\text{a.u})$, we apply exchange moves to the walkers $x = (\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)$ until their constituent particles are increasing according to this order (i.e $\mathbf{r}_1 \le \mathbf{r}_2 \le \mathbf{r}_3...$), and set their weights to +1. This is the same as the ordering used in Ref. [150], except here we only use this procedure for initialization of the walkers. As the simulation proceeds, walkers will be propagated into antisymmetric images of this initial positive-definite group, quickly setting up a large antisymmetric component.

3.4.3 Energy estimation

The Lagrange multiplier associated with normalization, E_T, corresponds to an energy offset which appears in our effective Hamiltonian H_X. As the algorithm progresses, the value of E_T is updated to keep the total weight of walkers, $W(\tau) = \sum_i |w_i(\tau)|$, roughly constant. The expected total weight after propagation from τ to $\tau + \delta\tau$ is given by

$$\langle W(\tau + \delta\tau) \rangle = \sum_i |\langle w_i(\tau + \delta\tau) \rangle| = \sum_i |w_i(\tau) G(x_i(\tau + \delta\tau), x_i(\tau), \delta\tau)|. \tag{3.18}$$

Separating this into contributions from different parts of the Green's function (see Eq. 3.9) we have

$$\langle W(\tau + \delta\tau)\rangle = \sum_i |w_i(\tau) G_X^{(i)} G_V^{(i)} G_D^{(i)} \exp(\delta\tau E_T)| \qquad (3.19)$$

where we have used the shorthand notation $G_I^{(i)} \equiv G_I(x_i(\tau + \delta\tau), x_i(\tau), \delta\tau)$ for $I \in [X, V, D]$. We keep the total weight roughly constant by requiring

$$\langle W(\tau + \delta\tau)\rangle \overset{!}{=} W(\tau) = \sum_i |w_i(\tau)| \implies$$

$$E_T(\tau) = \frac{1}{\delta\tau} \ln\left(\frac{\sum_i |w_i(\tau)|}{\sum_i |w_i(\tau) G_X^{(i)} G_V^{(i)} G_D^{(i)}|}\right). \qquad (3.20)$$

This is known as the *growth estimator* of the energy and, in order to keep the population stable, will converge to the lowest eigenvalue of H_X, which we can therefore estimate by averaging the value of $E_T(\tau)$ (after equilibration) over many iterations. However, because each iteration is correlated with the previous iteration, one must be careful in estimating the uncertainty of such averages. In this chapter we use the widely-employed reblocking method [89] to estimate statistical uncertainties.

If one has access to a suitable trial wavefuncion $\psi_T(x)$, that has non-zero overlap with the exact fermionic ground state, the fermionic energy can be estimated directly using the so-called projection estimator:

$$E_{\mathrm{proj}}(\tau) = \frac{\sum_i w_i(\tau) H \psi_T(x_i)}{\sum_i w_i(\tau) \psi_T(x_i)} \qquad (3.21)$$

Note that it is H, not H_X, that appears in Eq. 3.21. It is well-known [150] that the statistical uncertainties in $E_{\mathrm{proj}}(\tau)$ are typically smaller than that of $E_T(\tau)$, due to the reduced dependence on the fluctuating population. Cancellations between signed walkers contribute to these fluctuations, resulting in larger statistical errors for fermionic systems when using the growth estimator. This can be clearly seen by comparing DMC calculations of the ground state and first exited state of the Helium atom in Fig. 3.5. However, in order to use Eq. 3.21, we require both access to a trial wavefunction ψ_T and for the denominator to remain finite for sufficiently many DMC timesteps to build up accurate statistical averages. In transient methods, such as *release-node* DMC, the exponential decay of the fermionic component leads to an exponential decay of the denominator of Eq. 3.21 and a correspondingly small set of usable iterations from which to build up such averages. In contrast, we find that the fermionic state obtained from propagating the Green's function of H_X leads to stable (at least on the timescales we have probed in obtaining the results for this chapter) non-zero denominator of Eq. 3.21, as can be seen in Fig. 3.6, allowing straightforward use of projection-based estimators. However, both for simplicity

59
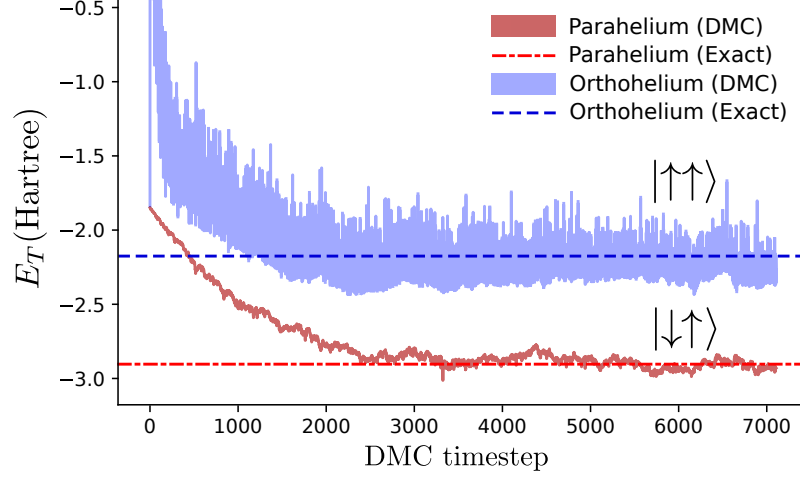
Figure 3.5: DMC calculations of the Helium atom ground state, with electrons having opposite spin (parahelium), and the exited (triplet) state with electrons having parallel spin (orthohelium), calculated using a stochastic nodal surface. A timestep of $\delta\tau = 10^{-3}$ atomic units was used with $\delta\tau_{\text{eff}} = 0.5$ atomic units (see section 3.4.4). It is clear to see that the exited (fermionic) state shows larger fluctuations than the (bosonic) ground state. This is due to cancellations between oppositely signed walkers contributing to fluctuations in the growth estimator of the energy. The reference energies are from VMC optimization of many-parameter trial wavefunctions [10], accurate to within a few μHa.
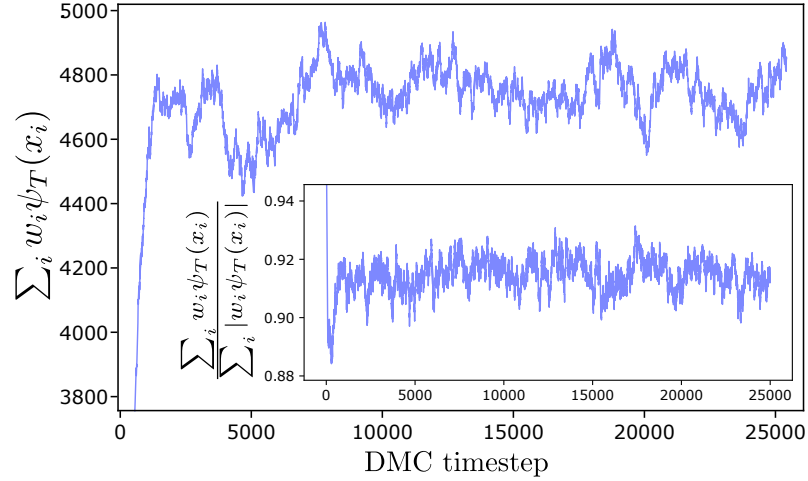
Figure 3.6: The denominator of Eq. 3.21 vs. DMC timestep for the simulation used to produce Fig. 3.4(b) (10^4 walkers, $\delta\tau = \delta\tau_{\text{eff}} = 10^{-3}$ atomic units). Note the y-axis scale. For the purposes of this plot, the trial wavefunction was set to the analytic fermionic ground state (shown in Fig. 3.4(d)). We can see that the denominator remains large and roughly constant. Inset: the denominator as a fraction of its maximum possible value (obtained if the walkers are all of the same sign as the analytic wavefunction).

and as a proof-of-concept, we restrict ourselves to considering implementations that require no trial wavefunction, and so are limited to using the growth estimator. Any symmetric component that remains (despite the exchange-moves and cancellations) will influence this estimator towards the symmetric ground state energy (as we will see later), rather than being removed as would be the case with a projection-type estimator. The extension of the method to include a trial wavefunction (when one is easily available) to allow both the use of the projection estimator and importance sampling (see Ref. [90]) would be a straightforward way to improve the efficiency of the method in future.

As is typical in DMC, after modifying the weights according to each part of the Green's function, we treat them with the *birth-death* algorithm described in Sec. 2.3.2. In atomic systems, timestep error can lead to a walker diffusing too close to a configuration where an electron overlaps with a nucleus and obtaining a correspondingly divergent (+ve) weight. The birth-death algorithm will then attempt to convert this single walker into a divergent number of child walkers. This is known as a *population explosion*. We mitigate this outcome by defining a maximum walker weight w_{max} and rejecting any DMC iteration where $\max(|w_i|) > w_{\text{max}}$ (we take $w_{\text{max}} \geq 4$ resulting in only 1 in every ~ 5000 iterations being rejecting). We also use a softened version of the coulomb interaction of the form

$$V_{c,\text{soft}}(r, r_s) = \frac{1}{r + r_s} \tag{3.22}$$

For the calculations performed for this chapter, $r_s \leq 10^{-5}$ which introduces a bias that is much smaller than the timestep error. We note that schemes to reduce the error due to coulomb singularities exist [175, 143], but are not employed here.

3.4.4 Effective nodal surface timestep

For systems existing in one spatial dimension the nodal surface is entirely specified by the antisymmetry constraint [39]. As a result, fermionic methods must be tested on higher-dimensional systems, which present a significantly increased challenge. For a fixed number of walkers, the average walker-walker separation $\langle |x_i - x_j| \rangle_{i \neq j}$ increases exponentially with the dimensionality of configuration space; a manifestation of the sign problem (see section 2.3.3). This allows the +ve and -ve walkers more space to slip past one another and induce the bosonic collapse of the wavefunction. To mitigate this outcome we introduce an effective timestep $\delta\tau_{\text{eff}} \geq \delta\tau$ and enforce the nodal surface of the corresponding diffused wavefunction

$$\psi_{D,\text{eff}}(x) = \sum_i w_i G_D(x, x_i, \delta\tau_{\text{eff}}). \tag{3.23}$$
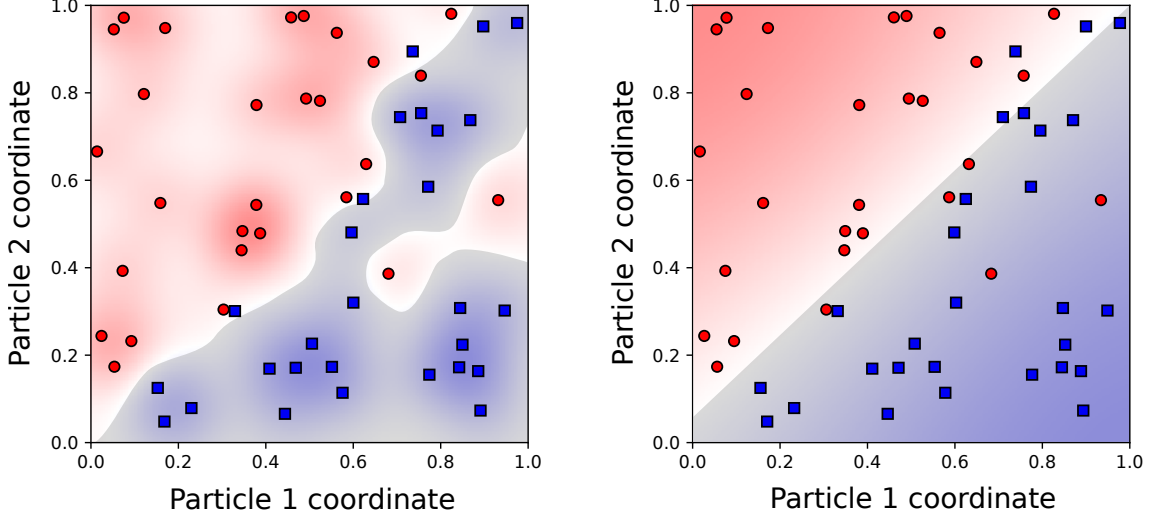
Figure 3.7: The effect of $\delta\tau_{\mathrm{eff}}$ on the nodal surface of a two-fermion system in 1D. Left: $\delta\tau_{\mathrm{eff}} = 0.1$ Right: $\delta\tau_{\mathrm{eff}} = 1.0$. Red circles (blue squares) represent the location of positive (negative) walkers. The background is shaded according to $\psi_{D,\mathrm{eff}}(x)$, where the nodal surface can be seen as a bright line separating the positive (red) and negative (dark blue) nodal pockets. It can be clearly seen that increasing $\delta\tau_{\mathrm{eff}}$ leads to a smoother nodal surface that is closer to the analytic nodal surface at $x = y$.

By increasing $\delta\tau_{\mathrm{eff}}$ we obtain a long-range ansatz for the nodal surface which, as before, is still free to vary in order to minimize the energy. Taking $\delta\tau_{\mathrm{eff}} > \delta\tau$ can be justified on physical grounds, as the kinetic energy contribution penalises wavefunctions that fluctuate over small length scales. By increasing $\delta\tau_{\mathrm{eff}}$, we are effectively smoothing out such fluctuations, as can be seen in Fig. 3.7. However, this is still an approximation and, as such, large values of $\delta\tau_{\mathrm{eff}}$ introduce a bias into the DMC energy which grows larger as more features of the nodal surface become unresolvable on the scale of $\delta\tau_{\mathrm{eff}}$. In order to keep this bias as small as possible, the long range nodal surface is applied post-hoc; $\delta\tau_{\mathrm{eff}}$ does not enter into the diffusive step or into evaluation of the functions $f_{\pm}(x)$ (see section 3.4.5).

We can see how increasing $\delta\tau_{\mathrm{eff}}$ takes us from the bosonic ground state to the fermionic ground state of a lithium atom in Fig. 3.8. The DMC energy plateaus at the fermionic energy as $\delta\tau_{\mathrm{eff}}$ increases above ~ 0.6 atomic units. We note that the resulting fermionic state is stable for long times, in contrast with transient methods such as *release-node* DMC. On increasing $\delta\tau_{\mathrm{eff}}$ beyond ~ 1.5, we enter the regime where $\delta\tau_{\mathrm{eff}}$ is too large to resolve the analytic nodal surface and a positive bias is introduced to the energy (see Fig. 3.8, inset). This is similar to the situation in fixed-node DMC where the energy is bounded from below by the true ground state energy and variational with respect to antisymmetric trial wavefunctions.
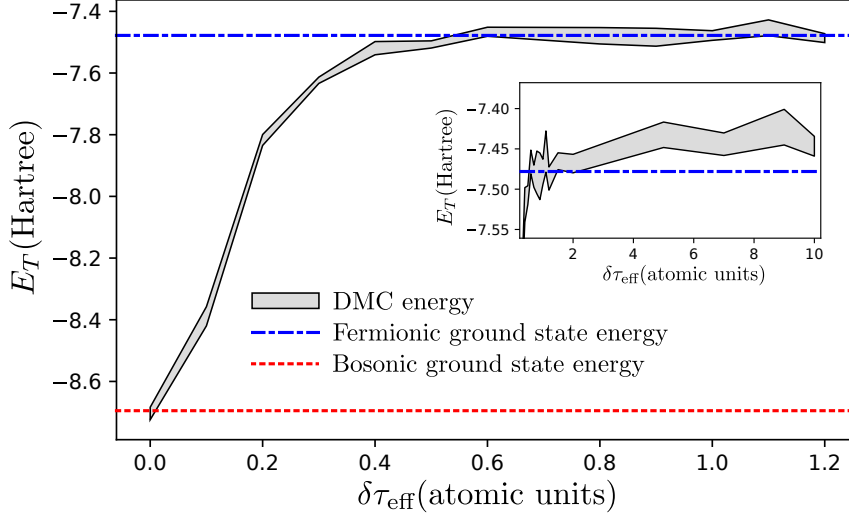
Figure 3.8: The DMC energy of a Lithium atom as a function of the effective timestep $\delta\tau_{\text{eff}}$ used to define the stochastic nodal surface. For each value of $\delta\tau_{\text{eff}}$, the energy was obtained from a simulation of 10^4 walkers for 10^5 iterations with a timestep of 10^{-3} atomic units. The DMC energy is shaded to $\pm$ the reblocked error. The blue dotted line is at the non-relativistic fermionic energy obtained from a Hylleraas-type expansion, accurate to within a basis set error of $< 10^{-9}$ Ha [225]. The inset shows the effect of increasing $\delta\tau_{\text{eff}}$ beyond sensible values.

Clearly, it would be useful to be able to identify a sensible value for $\delta\tau_{\text{eff}}$ without having to construct plots such as Fig. 3.8. From the form of Eq. 3.23, $\delta\tau_{\text{eff}}$ can be interpreted as the range of influence of a walker on the nodal surface (see also Fig. 3.7). A natural choice for its value is the expected midpoint distance between a +ve walker and its nearest -ve neighbour:

$$\delta\tau_{\text{eff}} = \left\langle \min_{x^-} \frac{|x^+ - x^-|}{2} \right\rangle_{x^+} \tag{3.24}$$

Where the minimization is over the positions x^- of all of the negative walkers, and the average is over the positions x^+ of all of the positive walkers. In a preliminary calculation of the beryllium atom, the value given by Eq. 3.24 fluctuates around $\delta\tau_{\text{eff}} = 1.35$. Carrying out an extended DMC calculation of the Beryllium ground state energy using this value for $\delta\tau_{\text{eff}}$ produces Fig. 3.9, from which the ground state energy is estimated as -14.665 ± 0.07 Ha, well within the error bars of the exact value of $-14.66654 \pm 2 \times 10^{-4}$ Ha obtained via Hylleraas-type expansions [244].

3.4.5   Summary of method

Combining the propagation stages explored in the preceding sections, we arrive at Algorithm. 2, which we refer to as *Exchange-Diffusion Monte Carlo* (XDMC). We note that steps 3-6 commute and can be applied in any order.
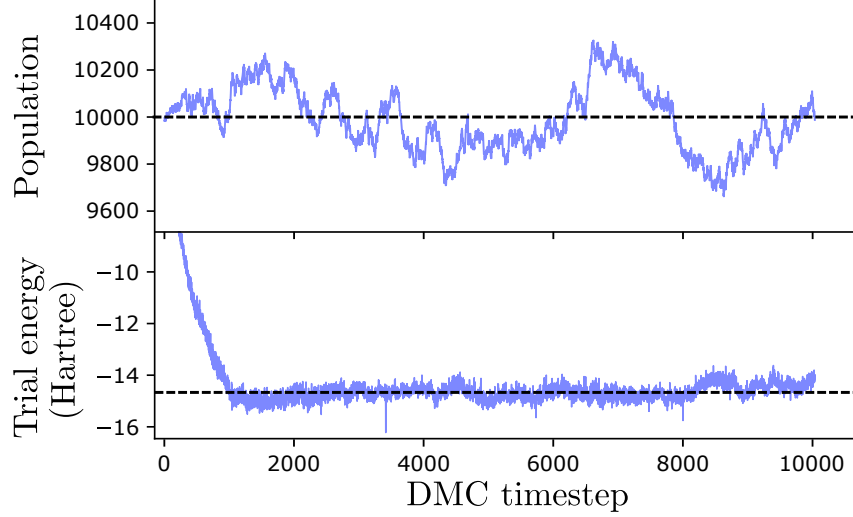
Figure 3.9: The evolution of a DMC calculation of a Beryllium atom. The calculation was carried out using 10^4 walkers for 10^4 timesteps, each of $\delta\tau = 10^{-3}$ atomic units. An effective timestep of $\delta\tau_{\text{eff}} = 1.35$ atomic units, derived in a preliminary calculation from Eq. 3.24, was used to describe the stochastic nodal surface. In the upper panel, the dashed line is at the target population. In the lower panel, the dashed line is at -14.66654 Ha, the energy obtained from a Hylleraas-type expansion, accurate to within 2×10^{-4} Ha of the exact value [244]. The DMC estimate of the energy is -14.665 ± 0.07 Ha.

Algorithm 2 Exchange-diffusion Monte Carlo (XDMC).

1: **Initialization** Initialise the walkers according to the procedure discussed in Sec. 3.4.2.

2: **repeat**

3:     **Exchange moves** To each walker, apply one of the exchange moves $x \to P_i x$, $w \to -w$ (or no exchange $x \to x$, $w \to w$), each with equal probability.

4:     **Diffusion** Diffuse each walker from $x \to x'$ with probability $G_D(x, x', \delta\tau)$.

5:     **Potential reweighting** For each walker, apply the potential reweighting $w \to w G_V(x, x', \delta\tau)$.

6:

    Cancellation Reweight positive walkers according to $w \to f_+(x')w$ and negative walkers according to $w \to f_-(x')w$ and, if $\delta\tau_{\text{eff}} > \delta\tau$, enforce the extended-range nodal surface of Eq. 3.23.

7:     **Birth-death** Replace each walker with $\lfloor |w| + u \rfloor$ walkers each of weight sign(w) where u is a uniformly-distributed random number in $[0, 1]$.

8: **until** Expectation values have converged to the required tolerance

---

64

3.5 Discussion

3.5.1 Relation to previous schemes

Step 6: cancellation

For the purposes of this thesis we opted not to explicitly pair walkers for cancellation and instead enforce a stochastic nodal surface defined by the entire population (see Fig. 3.10). However, the two methods are closely related. If we consider the limiting case of cancellation between two walkers with weights $w_1 > 0$ and $w_2 < 0$ at x_1 and x_2 respectively, then Eq. 3.13 reads

$$
\begin{aligned}
\psi_+(x) &= w_1 G_D(x, x_1, \delta\tau), \\
\psi_-(x) &= |w_2| G_D(x, x_2, \delta\tau).
\end{aligned}
\tag{3.25}
$$

From which we can construct the cancellation function $f_\pm(x)$ according to Eq. 3.17. The reweighting given by $w \to f_\pm(x)w$ now takes the form

$$
\begin{aligned}
w_1 &\to \max(w_1 - w_2 G_D(x, x_2, \delta\tau)/G_D(x, x_1, \delta\tau), 0), \\
w_2 &\to \max(w_2 - w_1 G_D(x, x_1, \delta\tau)/G_D(x, x_2, \delta\tau), 0).
\end{aligned}
\tag{3.26}
$$

This pairwise cancellation is the same as that proposed in Refs. [14] and [134]. Ref. [14] goes on to show that it is possible to extend this scheme to facilitate cancellations within a collection of more than two walkers, from which Eq. 3.17 can be recovered in the entire-population limit. In this thesis, Eq. 3.17 was instead obtained directly by requiring maximal separation of the walkers into nodal pockets (see appendix A.7). The schemes given in Refs. [14] and [134] can therefore be thought of as limiting cases of the maximal-separation scheme when only subsets of the population are considered for cancellation. This is a sensible approximation to make if each subset consists of walkers that are near to one another, due to the limited range of the diffusive Green's function. Indeed, one could approximate $\psi_D(x)$ by only considering the k nearest-neighbouring walkers to x, leading to

$$
\psi_D(x) \approx \psi_D^{(k)}(x) = \sum_{i=1}^{k} w_i G_D(x, x_i, \delta\tau)
\tag{3.27}
$$

where x_i are understood to be in order of increasing distance from x. Taking the $k = 1$ case corresponds to a Voronoi tiling of configuration space, where the sign of the diffused wavefunction at x is given by the sign of the nearest walker to x (as shown in Fig. 3.11). The same form of nodal surface is obtained on heuristic grounds in Ref. [195], where it is shown that it produces sensible results for low-dimensional ($D < 20$) configuration spaces. It is also known that the nodes of the free fermion density matrix approach that of the Voronoi wavefunction in the high temperature limit [39].
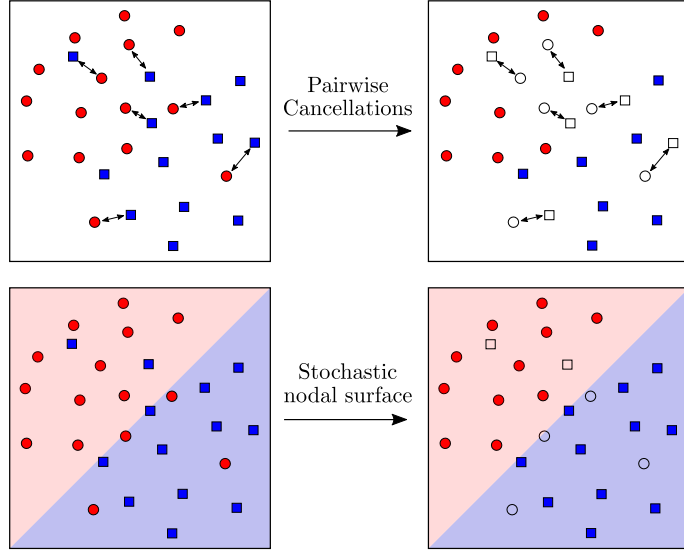
Figure 3.10: Schematic of cancellations via explicit pairing (upper two panels) and a stochastic nodal surface (lower two panels). Red circles represent positive walkers, blue squares represent negative walkers and empty shapes represent cancelled walkers. When using explicit pairing, walkers are first paired according to some criterion and then cancelled. This cancellation is often only partial and may take place over several iterations [270, 14, 134]. When using a stochastic nodal surface, the diffused wavefunction is evaluated for the configuration of each walker, and any walker with the wrong sign is immediately removed from the simulation.
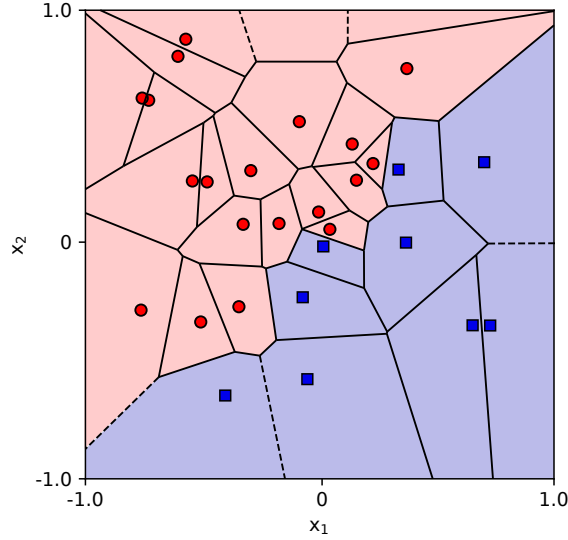
Figure 3.11: A Voronoi wavefunction for two non-interacting fermions in a 1D harmonic oscillator. Red circles (blue squares) represent positive (negative) walkers. The wavefunction is positive (negative) in red (dark blue) shaded regions. The emerging stochastic nodal surface at $x_1 = x_2$ can be clearly seen. Any walker crossing this surface in the next iteration will be removed from the simulation.

66

Step 3: Exchange moves

It is clear that the strictly *local* influence of a DMC walker is the limiting factor in describing antisymmetric wavefunctions. In this work, non-local information is introduced via the exchange moves. However, it is possible to incorporate this information implicitly in the form of the DMC walker itself. The simplest way to do this is to modify each walker to represent a set of symmetry-related points in configuration space, rather than just a single configuration:

$$\delta(x - x_i) \rightarrow \sum_{P \in \mathcal{P}} \text{sign}(P)\delta(x - Px_i) \tag{3.28}$$

where $\mathcal{P}$ is the set of all fermionic permutations of the system. We can obtain this representation by imposing antisymmetry constraints for the wavefunction under any of the permutations $\mathcal{P}$, rather than the pairwise exchanges $\mathcal{E}$ that we have used thus far. Whilst these are equivalent problems (the exchanges $\mathcal{E}$ generate the permutations $\mathcal{P}$), they lead to different propagation schemes. If we were to use the permutations $\mathcal{P}$, then the constraints that must be satisfied are

$$\psi(x) = \text{sign}(P)\psi(Px) \ \forall \ P \in \mathcal{P}. \tag{3.29}$$

Similarly to the exchange case, we can also derive an effective Hamiltonian for the permutation case:

$$H_P = H - E_T + \sum_{P \in \mathcal{P}} \mu_P(x)(1 - \text{sign}(P)P). \tag{3.30}$$

The Green's function for this Hamiltonian is given by

$$G_P(x, x', \delta\tau) = \langle x| \exp\left(-\delta\tau H_P\right)|x'\rangle. \tag{3.31}$$

Following the derivation for H_X, for small timesteps $\delta\tau$ we have

$$G_P(x, x', \delta\tau) \approx G_V(x, x', \delta\tau) \times \langle x| \exp(-\delta\tau T) \left[1 - \sum_{P \in \mathcal{P}} \delta\tau\mu_P(x)(1 - \text{sign}(P)P)\right]|x'\rangle$$

$$= G_V(x, x', \delta\tau) \times$$

$$\left[\underbrace{\left(1 - \sum_{P \in \mathcal{P}} \delta\tau\mu_P(x')\right)}_{\mathcal{N}_P(x')} \underbrace{G_D(x, x', \delta\tau)}_{\text{Diffusion from } x' \rightarrow x} + \sum_{P \in \mathcal{P}} \underbrace{\delta\tau\mu_P(Px')}_{\mathcal{X}_p(x')} \text{sign}(P) \underbrace{G_D(x, Px', \delta\tau)}_{\text{Diffusion from } Px' \rightarrow x}\right]. \tag{3.32}$$

Choosing the $\mu_P(x)$s such that $\mathcal{N}_P = 0$ and $\mathcal{X}_P(x')$ is constant (similarly to what we do in the H_X case), and noting that for a matrix A with entries $A_{i,j}$

$$\det(A) = |A| = \sum_{P \in \mathcal{P}} \text{sign}(P) \prod_i A_{i,P_i}, \tag{3.33}$$

we can obtain the form of the Green's function given in Eq. 13 of Ref. [270] for walkers of the form given in Eq. 3.28:

$$G_P(x, x', \delta\tau) =$$

$$G_V(x, x', \delta\tau) \begin{vmatrix} g(x_1, x'_1) & g(x_1, x'_2) & \dots & g(x_1, x'_n) \\ g(x_2, x'_1) & g(x_2, x'_2) & \dots & g(x_2, x'_n) \\ \vdots & & & \vdots \\ g(x_N, x'_1) & g(x_N, x'_2) & \dots & g(x_N, x'_N) \end{vmatrix} \qquad (3.34)$$

where

$$g(x_i, x'_j) = \frac{1}{\sqrt{2\pi\delta\tau}} \exp\left(-\frac{(x_i - x'_j)^2}{2\delta\tau}\right) \qquad (3.35)$$

is a diffusive Green's function taking single-particle coordinates $x_1, x_2...$ from the primed and unprimed configurations as arguments.

The propagation of walkers according to Eq. 3.34 is an alternative method to the propagation using exchange moves. It can be interpreted as the propagation of a collection of *second-quantized* walkers, each consisting of $N!$ symmetry-related delta-function walkers (as in Eq. 3.28). Due to its determinental form Eq. 3.34 can be evaluated in $O(N^3)$ time, rather than $O(N!)$ time [270]. Whilst this is still more expensive than the exchange-moves scheme (where the Green's function can be evaluated in $O(N)$ time) the additional walker images introduced by the permutations can be used to increase the cancellation rate between +ve and -ve walkers [270]. In order to determine the additional scope for cancellations, we should consider how much closer we can make walkers of opposite signs thanks to the additional permutations. In Fig. 3.12(a), we plot the distribution of the minimum separation obtainable by applying permutations to two walkers, each distributed according to a Gaussian with $\sigma = 1$ atomic unit. Mathematically speaking, we are plotting the distribution of the distance D given by

$$D = \min_{P \in S} |x - Py| \text{ where } x, y \sim \mathcal{N}(\mu = 0, \sigma = 1) \qquad (3.36)$$

where S is either the set of exchanges $\mathcal{E}$ or the set of permutations $\mathcal{P}$. Ideally, this distance would be as small as possible to facilitate cancellations between oppositely-signed walkers. In Fig. 3.12(b), we plot the average value of D against the number of fermions in the system. This distance increases linearly as the fermion count increases, resulting in an exponentially decreasing possibility for cancellation; a manifestation of the sign problem. This increase is slower when using the full set of permutations $\mathcal{P}$, because it affords us more freedom in the permuted configurations ($|\mathcal{P}| > |\mathcal{E}|$). The choice of whether to use exchanges or permutations depends strongly on how well this additional freedom can be exploited to perform cancellations in an algorithmic setting. Indeed, in the current work, evaluating Eq. 3.17 for
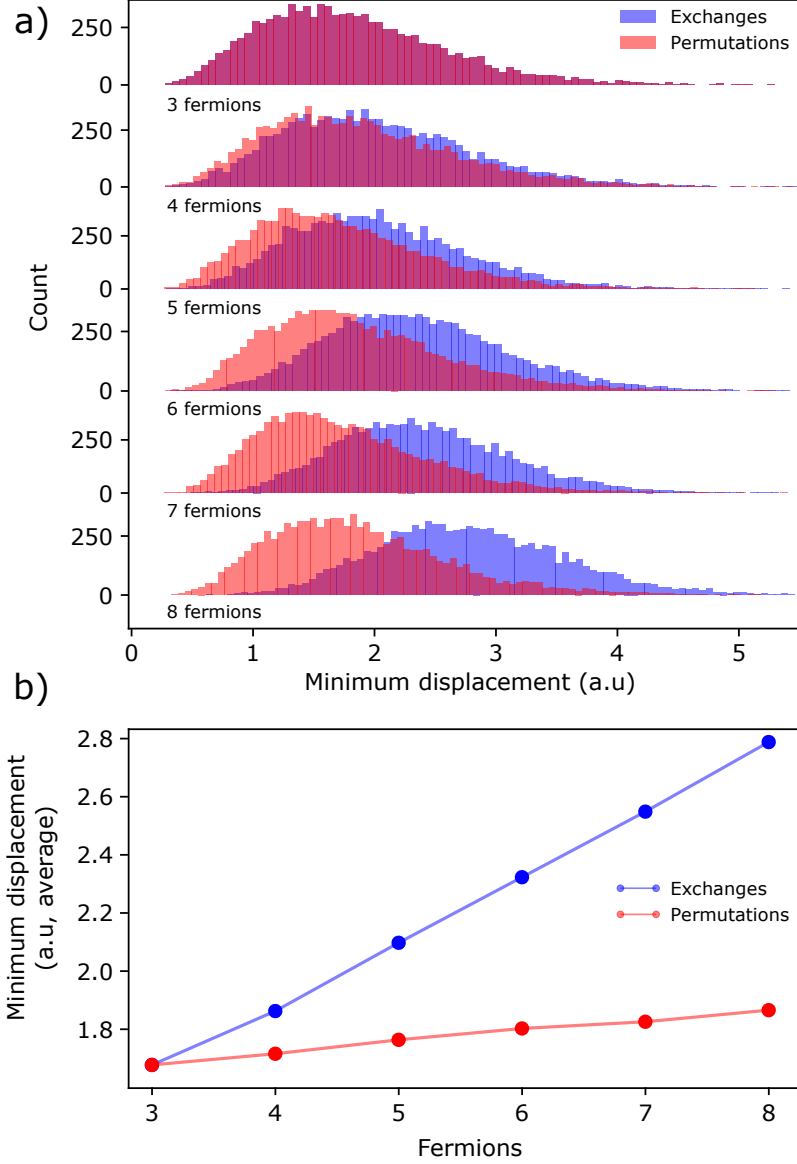
Figure 3.12: a) The distribution of D, given in Eq. 3.36. b) The average of the distributions in a), plotted against fermion number, showing the linear trend.

the purposes of cancellations is the rate-limiting step. For this reason, it is difficult to say in general whether permutations should be sampled directly, or via exchange-type moves. Further work on the implementation details of the permutation case is needed to resolve this question.

3.5.2 Scaling

The sign problem manifests itself as an exponential increase in the computational effort required to keep the bias in the energy estimator small as the number of fermions increases. In the method described in the preceding sections, the scaling is determined by the population of walkers required to obtain a stable fermionic

ground state, and how much this population can be reduced by increasing $\delta\tau_{\mathrm{eff}}$.

In Ref. [20], it is shown that the convergence of the energy to the infinite population limit can be sped up by reducing the Bose-Fermi gap (the difference in energy between the bosonic and fermionic ground states). Typically the Bose-Fermi gap is a constant property of the Green's function being sampled. However, in the present method, the Green's function is itself constructed from the entire walker population via inter-walker cancellations in Eq. 3.17 and the approximate nodal surface of Eq. 3.23. For small populations (without a large value of $\delta\tau_{\mathrm{eff}}$ to compensate) the cancellations due to exchanges become vanishingly probable and the we sample the bosonic dynamics of H rather than fermionic dynamics of H_X (this is the cause of bosonic collapse as discussed earlier). As the population increases, we approach the dynamics of H_X and the Bose-Fermi gap decreases (eventually to zero). This leads to a departure from fixed Bose-Fermi gap (power-law [20]) behaviour, as can be seen in Fig. 3.13.

To decrease the population required to describe a particular fermionic system we can increase the effective timestep $\delta\tau_{\mathrm{eff}}$. The improvement in convergence as a function of population obtained by doubling $\delta\tau_{\mathrm{eff}}$ can be seen in Fig. 3.13, allowing us to use around a quarter of the population for the same level of convergence. However, the amount that $\delta\tau_{\mathrm{eff}}$ can be increased is bounded by the length scale needed to resolve the analytic nodal surface, as can be seen in Fig. 3.8 (inset), where large values of $\delta\tau_{\mathrm{eff}}$ lead to a positive bias in the energy estimator. The optimal value of $\delta\tau_{\mathrm{eff}}$ can be estimated using Eq. 3.24 and, as can be seen in Fig. 3.15, using this value allows the description of larger systems than would otherwise be possible; we note none of the calculations in Fig. 3.15 would lead to a fermionic result in the limit $\delta\tau_{\mathrm{eff}} \to \delta\tau$. However, as the number of fermions increases, the number of walkers required to describe the nodal surface eventually also increases, regardless of the choice of $\delta\tau_{\mathrm{eff}}$ (forcing one to perform infinite population extrapolations as in Figs. 3.13 and 3.14). This can also be seen in Fig. 3.15 where, on increasing the number of fermions, (partial[2]) bosonic collapse eventually occurs leading to a large underestimation of the energy. As a result, even though increasing $\delta\tau_{\mathrm{eff}}$ enables a finite set of walkers to describe larger systems than would otherwise be possible, it cannot be increased fast enough with system size to completely overcome the sign problem.

---

[2]Cancellations still occur, and increase the energy estimator, just not at a sufficient rate to stabilise the fermionic ground state.
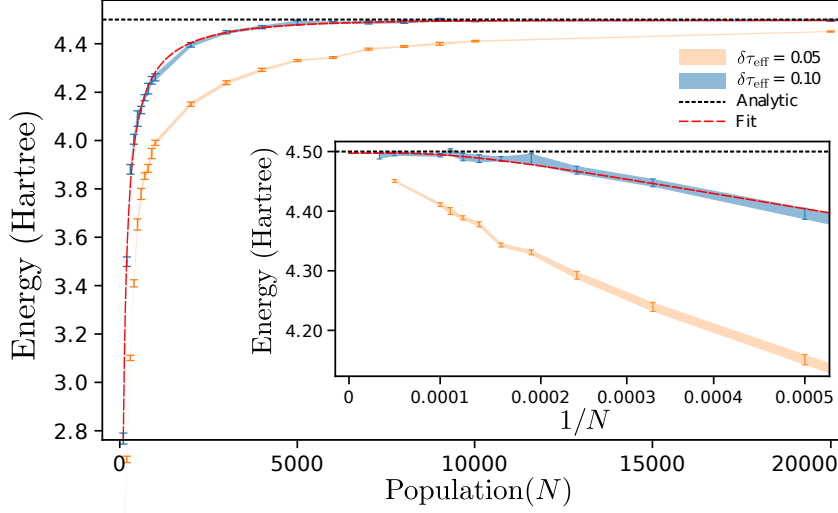
Figure 3.13: The DMC energy as a function of target population for three non-interacting fermions in a 1D harmonic well. The result for each population was calculated using 5×10^4 iterations with a timestep of $\delta\tau = 10^{-3}$ atomic units. Shown are calculations using two different values of $\delta\tau_{\text{eff}}$ (0.1 is shown in blue and converges faster than 0.05, shown in orange). The inset shows the same data plotted against the inverse population. For large populations, we found a deviation from power-law behaviour where the convergence is instead exponential. For $\delta\tau_{\text{eff}} = 0.1$ the best fit converges as $N^{-0.79} \exp(-N/3542)$ and gives an energy of 4.497 ± 0.003 Hartree in the infinite population limit. The analytic energy is 4.5 Hartree.
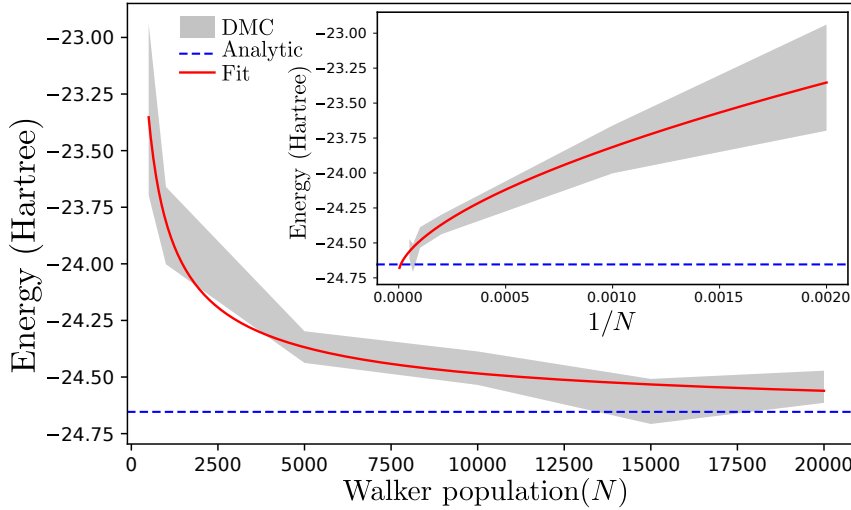
Figure 3.14: The DMC energy as a function of the target walker population for a Boron atom. The result for each population was calculated using 5×10^4 iterations with a timestep of $\delta\tau = 10^{-3}$ atomic units. $\delta\tau_{\text{eff}}$ was set to 1.35 atomic units, to facilitate comparison with our calculations of the beryllium atom. The inset shows the same data, plotted against the inverse population. The exact energy shown is at -24.65386608 $\pm 2 \times 10^{-9}$ Ha, which is the result obtained in the infinite-basis limit of an explicitly-correlated Gaussian basis set expansion [37]. The best fit power law converges as $N^{-0.6}$ and gives an energy of -24.67 ± 0.1 Hartree in the infinite population limit.
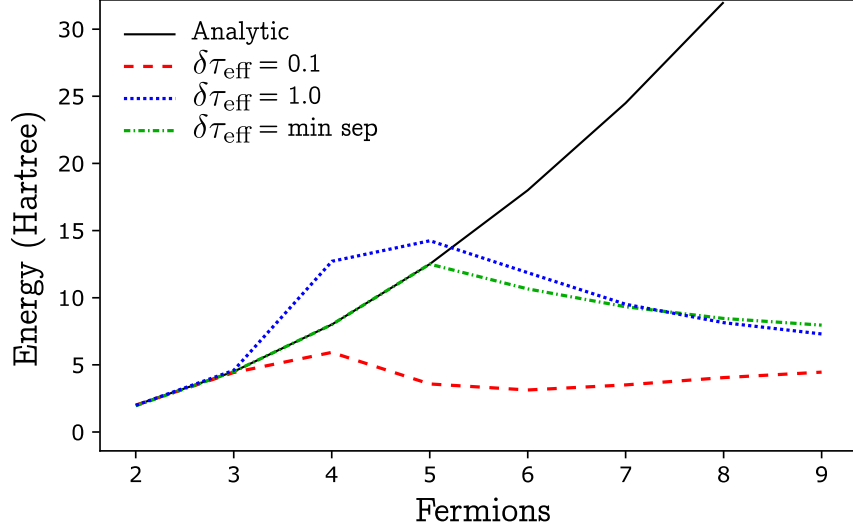
Figure 3.15: The energy of a system of non-interacting fermions in a 1D harmonic well, as a function of the number of fermions. The DMC calculations were carried out using 5000 walkers for 10^4 iterations with a timestep of $\delta\tau = 10^{-3}$ atomic units. The statistical errors in the energy are smaller than the widths of the lines.

3.6   Limitations and future direction

As discussed in the preceding sections, the major limitation of this method is the inability to treat larger systems. This arises from the sign problem, and manifests itself as the amplification of bosonic components in the wavefunction, which eventually dominate the fermionic component. The bosonic components in the current method are introduced by the stochastic nature of the walker moves, which do not preserve exchange symmetry on a per-timestep basis (so even using a purely fermionic initial state, bosonic components eventually appear). When the method is running successfully, these components are removed, by a cancellation process, faster than they are created. However, the increased degrees of freedom of larger systems lead to more routes in configuration space by which bosonic components appear - leading to an exponentially increasing rate of bosonic component generation, which will eventually overcome any cancellation method that does not itself become exponentially more efficient with system size. This is a fate shared by all current methods that rely on cancellation processes [134, 14, 270] and alternative methods to generate exact many-body wavefunctions also exhibit exponential scaling with system size in the general case (e.g configuration-interaction in the full-CI limit [209]). A possible route to overcome this limitation is to devise a scheme that preserves exchange symmetry on a per-timestep basis. Such schemes have recently shown promise in grid-based Monte Carlo calculations [150]. The extension of such symmetry-preserving methods to continuous space would be an interesting direction

of future work; a successful continuous-space, symmetry-preserving method would allow the evaluation of the nodal structure of many-body wavefunctions to an arbitrarily small precision for larger systems than the implementation in this work. Understanding the structure of many-body wavefunctions using such a method would provide valuable data to improve current approximate methods, in particular density functional theory. Currently, more accurate, but still approximate, methods are used for this purpose (such as coupled-cluster theory or configuration interaction). The ability to double-check these so-called "*gold standard*" methods using an exact many-body treatment would allow us to validate their gold-standard status.

Besides the system size limit, there are several ways in which the method (as described in this chapter) could be improved. These are all available as standard in production QMC codes and would be straightforward to implement in the present method

1. Allowing the treatment of spin-polarized systems, which would be required to treat systems in a magnetic field. The spin degree of freedom is already implemented in the current code, but is only used to determine the exchange symmetry of the particles. The derivation of the Green's function could be readily generalized to include spin, allowing the derivation of the propagation scheme for spin-polarized systems.

2. The ability to specify a trial wavefunction would allow the method to take advantage of importance sampling to increase sampling efficiency [90]. It would also allow the use of projection estimators for observable properties, reducing the bias introduced when bosonic components are present (see sec 3.4.3).

3. The aforementioned trial wavefunctions could be used to enforce the nodal condition upon the walkers. Strictly enforcing the nodal condition would lead to a normal FN-DMC calculation, but one can easily imagine a scheme whereby the nodal condition were only softly enforced (e.g by only removing a walker from the simulation once it has entered a region of the opposite sign, but with a sufficient *magnitude*). Either way, this should be strictly optional functionality as trial wavefunctions used for this purpose add a fixed-node bias to estimated quantities.

4. The estimation of expectation values besides the total energy (e.g density, spin-density, one-body and two-body density matrices, structure factors etc.) - formulae for calculating these from QMC trajectories are readily available in the literature [90]. This would allow prediction of various experimental quantities, such as X-ray diffraction patterns, as well as investigation of theoretical problems such as density and density-matrix functional theories.

3.7 Conclusion

We have constructed a scheme for fermionic diffusion Monte Carlo that makes no reference to a trial wavefunction. We have shown how the resulting propagation scheme can be interpreted as the formation of a *stochastic* nodal surface, which is free to vary and minimize the energy. We go on to derive a diffusion scheme that maximally stabilizes the nodal surface and show that stable fermionic ground states for simple harmonic systems and light atoms can be obtained. We have demonstrated that the number of walkers required to resolve the nodal surface can be reduced, without introducing significant bias, by introducing an approximate long-range influence on the nodal surface and have provided a method for estimating a sensible choice for the associated parameter ($\delta\tau_{\text{eff}}$). Extending the method to employ a guiding wavefunction for the purposes of importance sampling and to allow the use of projection estimators should allow the study of larger systems, but the sign problem still persists for the method in its current form. We hope that methods based on the constrained-optimization formalism of DMC introduced in this thesis will enable studies to improve the understanding of nodal surfaces in electronic wavefunctions. In future, this method could be applied to the study of exchange and correlation in periodic systems, with the potential to generating exchange-correlation functionals for DFT calculations that do not depend on a choice of trial wavefunction at the DMC level. An open-source C++ implementation of the methods described in this chapter is available [2].

Chapter 4

Mixed phases in Lithium metal

4.1 Motivation

Lithium is the lightest simple metal in the periodic table. It has two stable isotopes, Li-7 (which we focus on in this chapter) and the much rarer Li-6, both of which exhibit a Martensitic (diffusionless) phase transition from BCC to close-packed upon cooling [9, 233, 54]. Such phase transitions are of significant practical importance, for example in the martensite-austenite transition in steel, upon which many thermomechanical treatments and alloying methods are based [57], or in the shape-memory of materials [129]. Li is the simplest metal that undergoes such a transformation and, because it is so light, (potentially anharmonic) quantum-nuclear effects contribute significantly to the free energies of different structures [19, 252, 54]. In particular, at low temperatures Li exhibits close-packed polytypism, whereby the Martensitic transition results in a mixture of metastable close-packed phases that can be described by hexagonal layers. Thanks to its simplicity, Li is a useful model system for understanding transformation kinetics and characterization of such mixed phases, especially where vibrational effects are important [9, 38, 8]. The work in this chapter weighs in on a recent dispute over the ground state structure of Lithium metal [9], with a particular focus on anharmonicity and mixed-phase behaviour.

4.2 Vibrational energies

Thanks to Li being the lightest metal, vibrational contributions are of crucial importance to the thermodynamics of Li phases [252, 9, 104]. In particular, we extend the Harmonic theory of phonons described in Sec. 2.5.1 to the anharmonic approach described in Sec. 2.5.2 leading to the Hamiltonian

$$H = \left(\sum_{\mathbf{q},\sigma} -\frac{1}{2} \frac{\partial^2}{\partial p_{\mathbf{q},\sigma}^2} \right) + V(\{p_{\mathbf{q},\sigma}\}, \beta). \tag{4.1}$$

where $V(\{p_{\mathbf{q},\sigma}\}, \beta)$ is the Born-Oppenheimer energy at the perturbation given by phonon coordinates $\{p_{\mathbf{q},\sigma}\}$. The inverse temperature β can affect the Born-Oppenheimer energy through entropic effects (Mermin entropy) and smearing of the electronic Fermi surface. However, we found these effects to be negligible in our calculations, helped by the use of a dense electronic k-point sampling which provides an accurate resolution of the Fermi surface.

Following Sec. 2.5.2, we calculate an anharmonic correction to the phonon dispersion at each phonon q-point independently by diagonalizing H in the basis of single phonon states at that q-point, $\{|\mathbf{q}, \sigma\rangle\}$,

$$H_{\sigma,\nu}(\mathbf{q}) = \langle \mathbf{q}, \sigma | H | \mathbf{q}, \nu \rangle \tag{4.2}$$

Clearly there is an infinite number of excited states (σ's) at each q-point; in order to evaluate the matrix elements we truncate this set to the first 20 exited states, which is more than sufficient for our purposes. The anharmonic potential $E_{el}(\{p_{\mathbf{q},\sigma}\}, \beta)$, is mapped using DFT calculations for a discrete set of 31 amplitudes for each mode. The potential is then interpolated to a large number of points (5000) along each amplitude using a quadratic spline. The integral in Eq. 4.2 is then carried out over these points to obtain the matrix elements of the Hamiltonian, which is then diagonalized. The resulting eigenvalues are the anharmonic excitation energies at a particular q-point, giving an anharmonic correction to the phonon dispersion. These excitation energies can then be used to construct an anharmonic phonon free energy at any given temperature (including entropic and zero-point effects) [93, 25]. We find that this method is around 50 times more expensive than traditional harmonic phonon calculations.

Application of a Legendre transform to the free energies provides the Gibbs free energy at a given temperature and pressure via the following minimization:

$$G(T, P) = \min_V [F(T, V) + PV] \tag{4.3}$$

The thermodynamically most stable phase has the lowest Gibbs free energy at a given temperature and pressure. We perform this minimization by fitting $F(T, V)$ data to a suitable equation of state (see results section 4.3.2). Within this method the effects of thermal expansion and any anharmonic contributions to the equilibrium volume of the system are included.

4.3 Results and discussion

4.3.1 Structure searching

The FCC, BCC, HCP and 9R structures have all been proposed in the past to explain the experimental data for Li [9, 211, 26, 27]. We find that these struc-

Space group	F_{min} (meV/atom)	V (Å^3/atom)	N
$Fm\bar{3}m$ (FCC)	0.0	18.96827	39
$P6_3/mmc$ (HCP)	0.8115	18.97343	17
$Im\bar{3}m$ (BCC)	1.3684	18.94680	26
$R\bar{3}m$ (9R)	2.0593	18.94582	3
$I4/mmm$	2.1320	18.94480	30
$C2/m$	2.5273	18.97458	10
$Cmcm$	2.7300	18.93219	5
$Immm$	5.2318	18.97877	2
$P2_1/m$	11.781	18.99312	3
$P6/mmm$	14.167	19.05003	3

Table 4.1: The lowest-energy results of a simple Li AIRSS search. For each space group the number of times it was found (N) is shown. We also report, for each space group, the lowest Helmholtz free energy found (F_{min}, relative to the FCC phase) and the corresponding volume (V). The space groups are reported in order of increasing F_{min}. Only electronic energies are calculated, vibrational energies are neglected in these calculations. The conventional names for the four phases which are investigated more deeply in this chapter are shown in brackets next to the space group. In total 469 structures were generated.

tures can be recovered very quickly from first principles calculations using *ab initio* random structure searching (AIRSS [220, 204, 221], see Sec. 2.7.2) using only very simple constraints on the initial structures. In particular we constrain the volume per atom to be within 50% of the known value (based on the density of solid BCC Li) and require that no two atoms are within 1Å of one another (approximately 1/3 of the nearest neighbour distance in BCC Li). The unit cells are generated to contain between 1 and 6 atoms. The results of an AIRSS search with these constraints are shown in Table 4.1. We see that it is possible to obtain a good heuristic understanding of the energy landscape, even when looking at these static stable structures, where the effects of vibrations are neglected. The electronic energies of these phases are extremely similar to one another, with differences on the order of meV/atom. Several other searches were also carried out with additional constraints on the symmetry of the initial structures, but this was not found to be useful in this case.

Close packed structures

In the previous section we saw that the AIRSS searches quickly recovered the FCC, BCC, HCP and 9R structures. Most of the other structures found in the searches

consist of various close-packed polytypes. Any repeating sequence of hexagonal layers, where no two adjacent layers are of the same type (the layer types are labeled A, B and C; see Fig. 4.1), gives a close-packed structure. As a result, there is an infinite number of different close packed polytypes. Many of these are realized in the Martensitic phase of Li that is observed experimentally at ambient pressure and low temperatures; this phase consists of a mixture of close-packed phases and remnants of the high-temperature BCC phase [9]. This is an example of one of the most common forms of Martensitic transformation (BCC → close-packed) [135], in which anharmonic effects may play an important role [103]. Close packed crystal structures with a short repeating sequence of hexagonal layers are more likely to be realized in such a mixed phase, as several repeats are more likely to fit between defects. Layer-layer interactions also diminish rapidly with inter-layer separation [177]. This justifies the truncation of the infinite set to smaller sequences that are below a certain repeat length. In particular we investigate the energetics of all of the possible close-packed Li structures with a sequence of 15 or fewer layers. There are 10,922 such structures[1], 489 of which are unique, labeling equivalent structures following Ref. [177][2]. For a given unique structure i, we call the number of equivalent structures w_i.

The powder neutron diffraction pattern is simulated for each of these unique structures and the resulting overall mixed-phase pattern is obtained by thermodynamic weighting of the individual patterns. Neglecting defect energies the resulting neutron pattern at temperature T is given by (see Appendix. A.4):

$$I(2\theta) = aI_{\text{BCC}}(2\theta) + b\sum_i I_i(2\theta)\frac{w_i}{\exp\left(\frac{E_i-\mu}{k_BT}\right) - 1} \tag{4.4}$$

where $I_i(2\theta)$ is the pattern for the ith structure, E_i is the energy per atom and w_i is the multiplicity of the structure (as defined in the previous paragraph). a, b, μ and T are fitted to the experimental pattern. The resulting combined patterns are shown in Fig. 4.2.

In agreement with Ackland *et al.* the experimental data are well described by a Martensite consisting of a mixture of close-packed phases and BCC remnants [9]. We give the results for such weighted close-packed patterns both at a fixed volume per atom, and after optimizing the volume of each phase. The latter case is equivalent to allowing density variations across the sample based on the local

---

[1]This number is $\sim \frac{2}{3}2^{14}$. In general the number of close-packed structures with $\leq n$ hexagonal layers is $\sim \frac{2}{3}2^n$. The factor of 2^n arises from the number of leaves in a tree such as Fig. 4.1 and the $\frac{2}{3}$ arises from the $\frac{1}{3}$ of such structures which have the same initial and final layer in the sequence (meaning they do not correspond to a close packed structure).

[2]We note a minor error in [177], where it is stated that the number of unique close-packed sequences with up to 10 atomic layers is 43. There are in fact only 38.
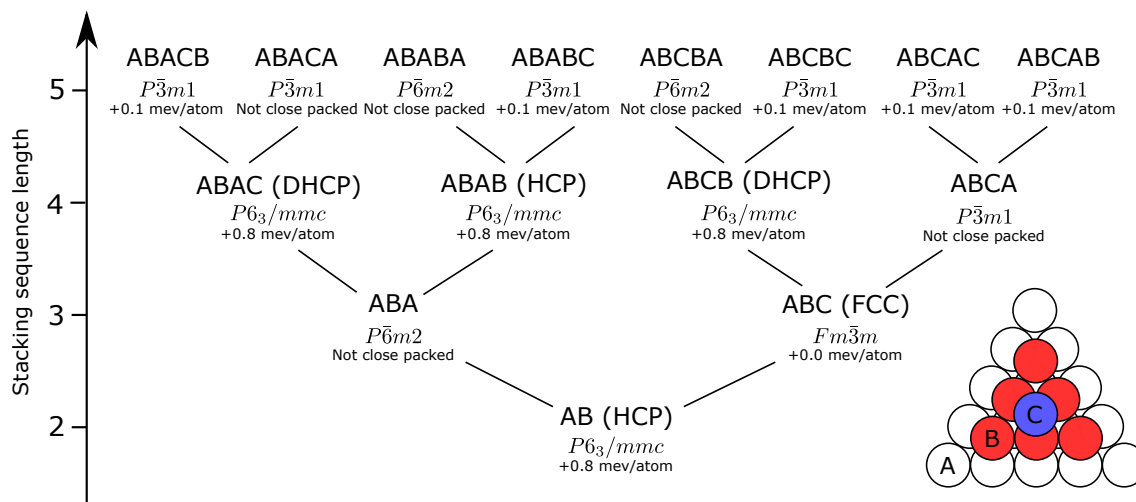
Figure 4.1: The possible hexagonal-layer stacking sequences from 2 to 5 layers, showing the stacking sequence, the space group and the electronic energy per atom relative to the FCC (ABC stacking) electronic ground state. Note that the energy differences are on the meV scale. The inset is an illustration of the A, B and C type layers. Structures which are not close packed as a result of having the same first and last layer in the sequence are labeled as such.

Figure 4.2: Simulated and experimental neutron diffraction patterns for Li. The experimental pattern (< 20 K at ambient pressure) from [30] is shown, as well as patterns derived by thermodynamically weighting close-packed Li structures and AIRSS structures using DFT energies (see Eq. 4.4). For the close-packed patterns all unique close-packed sequences consisting of fewer than 16 hexagonal layers are included. A Gaussian broadening was applied to the derived patterns with a width derived from a best-fit to the experimental data.

crystal structure. We see little difference in the resulting patterns, likely because the optimized volumes are all very similar.

Simply weighting the close packed phases thermodynamically in this way reproduces many of the non-BCC characteristics of the experimental pattern, but not as accurately as the large-scale molecular dynamics calculations performed in Ref. [9]. This is because transformation kinetics are not fully described by a simple thermodynamic weighting of phases, as energy barriers along transition pathways between phases also play an important role [38]. Such kinetic effects can be captured by molecular dynamics simulations using suitable interatomic potentials [9, 141]. For completeness, the pattern that results from combining the structures found in our AIRSS search using Eq. 4.4 is also given in Fig. 4.2. We see that this actually gives a reasonably close pattern to the experiment, suggesting that a decent first approximation to a mixed phase material might be obtained by simply mixing together the structures that appear in an AIRSS search for that material.

Absence of the DHCP structure

It is interesting to note that our AIRSS searches did not return the DHCP structure (with hexagonal stacking sequence ABCB, space group $P6_3/mmc$ - see Tab. 4.1). In order to investigate this further, we took the DHCP structure and displaced the locations of each of the atoms by a small random amount (up to 0.1Å in any direction), followed by a geometry optimization. In the majority of cases the structure failed to relax back down to the DHCP structure; suggesting it has a small basin of attraction in configuration space. The resulting structures are shown in Table 4.2. The $Cmcm$ structure appears most frequently and the Cm structure is also prevalent. These structures both appear in the search used to generate Tab. 4.1, in lieu of the DHCP structure. As a result, we are led to believe that the basin of attraction of the DHCP structure is small, and closely surrounded by the other minima appearing in Table 4.2. All of these structures are very close in energy to the DHCP structure and many have space groups which are subgroups of $P6_3/mmc$, suggesting that they block the DHCP structure in our AIRSS search.

F_{min} (meV)	Space group	N_{sg}
0.0	$P6_3/mmc$	12
0.009	$Cmcm$	19
0.039	Pm	4
0.040	$Amm2$	5
0.049	$Cmc2_1$	5
0.068	Cm	11
0.086	$P2_1/m$	5
0.095	$C2$	9
0.104	$P1$	5
0.122	$C2/m$	2
0.194	Cc	2
0.201	$C222_1$	1
0.207	$C2/c$	1
0.210	$P2_1$	3
0.245	$P6_3mc$	1
0.263	$P\bar{6}m2$	1

Table 4.2: The space groups resulting from perturbing the atoms of the DHCP (ABCB, $P6_3/mmc$) structure by up to 0.1Å and performing a subsequent geometry optimization. N_{sg} is the number of structures found with each space group.

4.3.2 Free energy calculations

Quasi-harmonic results

Quasi-harmonic calculations of the Gibbs free energy are performed for the FCC and BCC phases of Lithium, using a parameter set chosen to give well-converged physical properties. We use plane-wave DFT with a PBE functional, a plane-wave cutoff of 3 keV and an electronic k-point grid with a spacing of 0.02 Å^{-1}. The phonon dispersion is calculated on a $10 \times 10 \times 10$ grid in reciprocal space using the highly-efficient non-diagonal supercells method [176] and then Fourier interpolated to a $40 \times 40 \times 40$ grid. The Helmholtz free energy is calculated for a range of volumes around the equilibrium volume and fitted to the Birch-Murnaghan equation of state [33] to extract the Gibbs free energy. An example of such a fit at 300 K is shown in Fig. 4.3.

The Gibbs free energy derived from the Birch-Murnaghan fits is shown in Fig. 4.4 for a range of temperatures between 0 and 350 K. The Murnaghan, Rose-Vinet and Poirier-Tarantola equations of state [201, 271, 224] give essentially identical results. A transition from the FCC to a BCC phase occurs upon heating at 217±13 K. This phase transition is observed on isobaric heating of the FCC phase experimentally; however, the calculated transition temperature is somewhat above the experimental range of 110-200 K [9, 216, 246, 190, 235, 52]. The reverse BCC → FCC transition
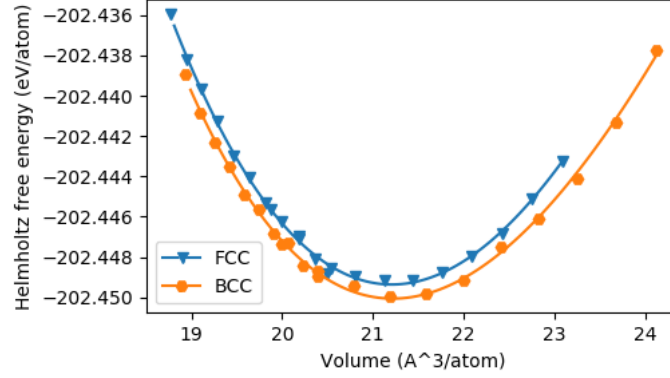
Figure 4.3: The Helmholtz free energy per atom verses volume per atom at 300 K for FCC and BCC Lithium. The solid lines show the Birch-Murnaghan fit used to derive the 300 K data in Figs. 4.4 and 4.5.

System	Parameter	Value	Standard fitting error
FCC	E_0	-199.6000 eV	0.6 meV
	V_0	19.36 Å^3	0.05 Å^3
	K_0	0.08993 GPa^{-1}	0.0009 GPa^{-1}
BCC	E_0	-199.6011 eV	0.4 meV
	V_0	19.39 Å^3	0.02 Å^3
	K_0	0.09100 GPa^{-1}	0.0003 GPa^{-1}

Table 4.3: Results of the fit to the Birch-Murnaghan equation of state at 0 K.

is not seen experimentally upon isobaric cooling at ambient pressure. The FCC phase is instead prepared via a high-pressure route [9] to avoid formation of the Martensitic phase investigated in section 4.3.1. Isothermal compressibilities and thermal expansion coefficients are given in Tables. 4.3 and 4.4 (calculated from fits at 0 and 300K, respectively) and match experimental results closely [139, 50].

Quasi-anharmonic results

Using the method outlined in section 4.2, we calculate an anharmonic correction to the Helmholtz free energy, effectively moving from the quasi-harmonic regime into the "quasi-anharmonic". The resulting correction to the Gibbs free energy is shown in Fig. 4.6. The effect on the FCC → BCC transition is included in Fig. 4.4. Anharmonic effects are much stronger in the BCC phase, with the FCC → BCC transition temperature increasing by 15 K (to 232 K) as a result.

It is also interesting to note the difference between the lowest-lying eigenvalues of the harmonic and anharmonic Hamiltonians, shown in Figs. 4.7 and 4.8. These corrections are present even at zero temperature as they result directly from the

System	Parameter	Value	Standard fitting error
FCC	E_0	-199.6383 eV	0.9 meV
	V_0	19.78 Å^3	0.07 Å^3
	K_0	0.0822 GPa^{-1}	0.001 GPa^{-1}
BCC	E_0	-199.6420 eV	0.7 meV
	V_0	19.85 Å^3	0.03 Å^3
	K_0	0.0802 GPa^{-1}	0.0006 GPa^{-1}

Table 4.4: Results of the fit to the Birch-Murnaghan equation of state at 300 K from Fig. 4.3.
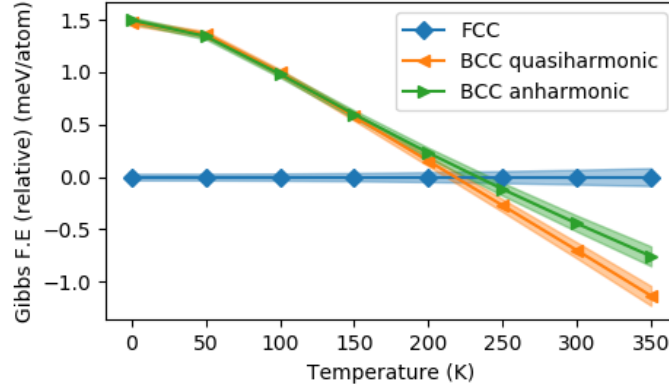
Figure 4.4: The Gibbs free energy of the FCC and BCC Li phases at a range of temperatures. This data was obtained by fitting the Birch-Murnaghan equation of state to data for $F(V, T)$ from DFT calculations (see main text). The standard fitting error is shown as a shaded region. The result of including the anharmonic correction from Fig. 4.6 is also shown.
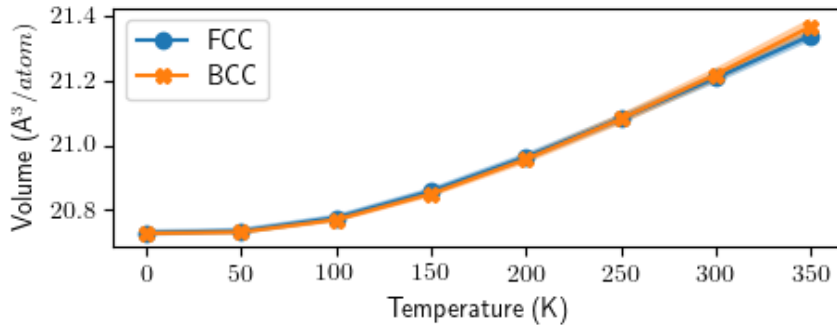
Figure 4.5: From the same calculation as Fig. 4.4, but showing the derived equilibrium volume and thermal expansion.
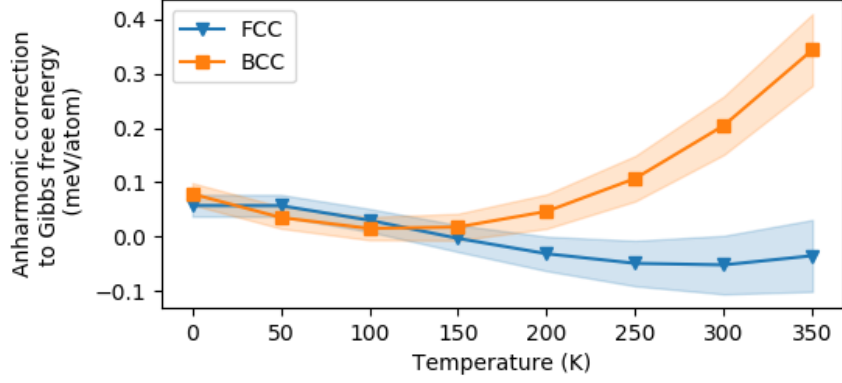
Figure 4.6: The correction to the Gibbs free energy of BCC and FCC Li that results from including the effects of anharmonic vibrations via Eq. 4.2.

anharmonicity of the 0 K Born-Oppenheimer surface. The most significant effect is a stiffening of the N-point $(0, 0.5, 0)$ phonon in BCC Li by ~ 1 meV (around 30 % stiffer than the harmonic case). It has been suggested that stiffening of this mode could be important in stabilizing BCC phases [46], however we find that this mode is dynamically stable to begin with (contrary to BCC Zr in [223, 46]).

Functional dependence, including the 9R and HCP structures

We perform a second set of similar calculations which include the 9R and HCP structures as well as the BCC and FCC structures. These calculations are carried out with an LDA functional and with less intensive convergence parameters (1 keV plane-wave cutoff, 0.035 Å^{-1} electronic k-point grid spacing and an 8x8x8 phonon q-point grid) in order to investigate the sensitivity of our results with respect to convergence parameters and the functional used. The resulting Gibbs free energies are shown in Fig. 4.9. The sequence of phase transitions with increasing temperature is the same as in the PBE calculations in Fig. 4.4. However, the FCC → BCC transition temperature in the LDA calculations is 47 K higher at 264±26 K. This is due to the fact that the transition temperature is extremely sensitive to the convergence of the differences in the Gibbs free energy; we estimate an increase of 117 K in transition temperature per meV increase in the BCC-FCC energy difference. The appearance of a metastable BCC → 9R transition upon cooling through 120 K is compatible with the Martensitic transition, however the predicted transition temperature is once again somewhat above the experimental range [246, 235, 247, 190, 216, 52].

We go on to show that the FCC → BCC transition is driven by entropic contributions to the phonon free energy. The phonon free energy is given by:

$$F_{ph}(T, V) = \frac{1}{2} \int \omega D(\omega, V) d\omega + \int \frac{\omega}{\exp(\omega/T) - 1} D(\omega) d\omega - TS_{ph}(T, V) \quad (4.5)$$
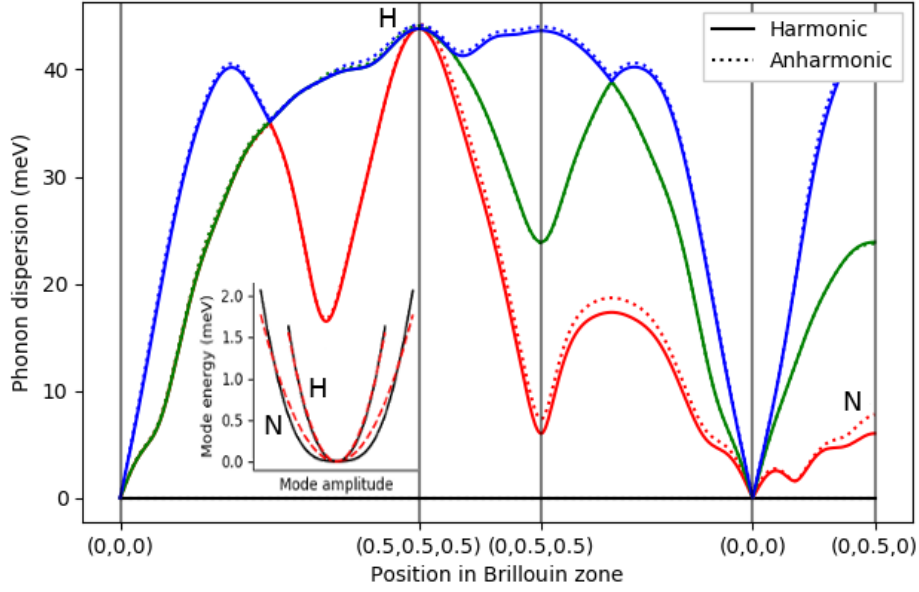
Figure 4.7: The 0 K vibrational band structure of BCC Li. The solid line shows the normal harmonic phonon dispersion. The dotted line shows the lowest eigenvalue of the anharmonic Hamiltonian given in Eq. 4.2. The Brillouin zone positions are given in terms of primitive reciprocal lattice vectors. The shapes of the mode potentials at points N and H in the Brillouin zone are shown in the inset with quadratic fits (red dashed lines) to illustrate the anharmonicity at the N point.
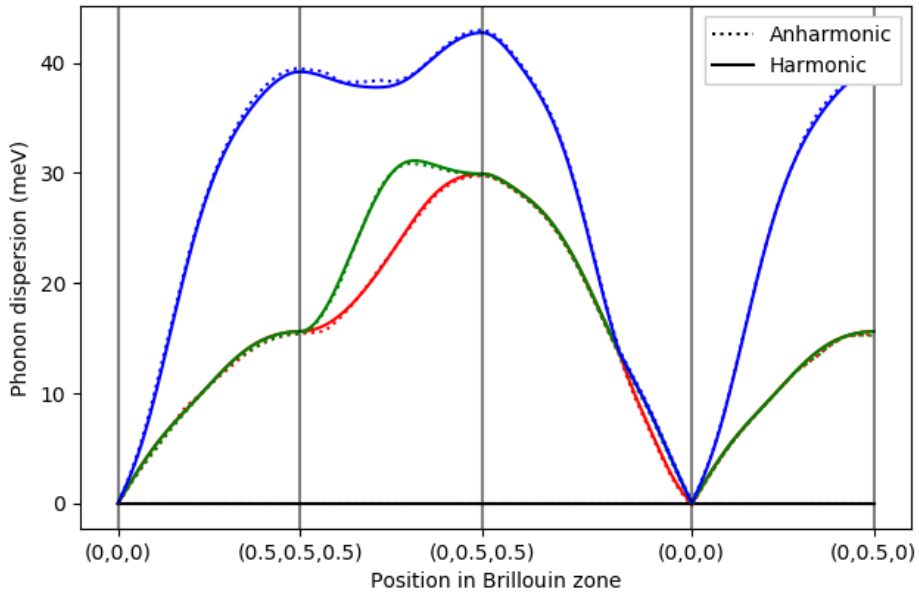
Figure 4.8: As Fig. 4.7, but for FCC Li. In this case the anharmonicity is negligible.
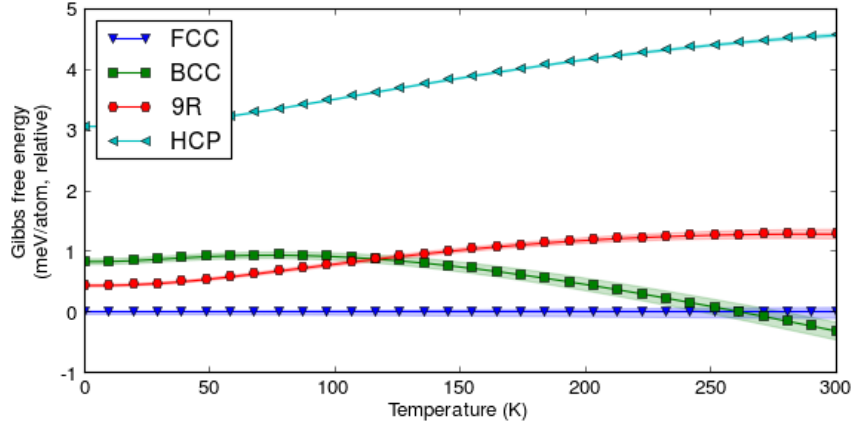
Figure 4.9: The Gibbs free energy of Li phases at a range of temperatures. The data was obtained by fitting the Birch-Murnaghan equation of state to data for $F(V, T)$, obtained using DFT calculations (see main text). The standard fitting error is shown as a shaded region.

where $D(\omega)$ is the phonon density of states. The first term is the zero-point energy, the second arises from the thermal occupation of phonon states and the entropic term, S_{ph}, is given by [93, 25]:

$$S_{ph}(T, V) = \int \frac{\omega/T}{\exp\left(\omega/T\right) - 1} D(\omega, V) d\omega - \int \ln\left[1 - \exp\left(-\omega/T\right)\right] D(\omega, V) d\omega \quad (4.6)$$

The importance of the entropic term becomes apparent when its contribution is explicitly neglected, resulting in the Gibbs free energy landscape shown in Fig. 4.10 in which phase transitions are no longer present. The individual contributions to the phonon free energy for BCC Li are shown in Fig. 4.11, in which it can be clearly seen that the entropic contribution dominates the zero-point and occupational terms. The FCC, HCP and 9R phases show similar behaviour. Because the low-lying phonon modes in BCC Li are softer than those of FCC (see Figs. 4.7 and 4.8), the entropic effects are stronger, which leads to the FCC $\rightarrow$ BCC phase transition upon heating.

4.4 Conclusions

We have investigated the application of a general framework for deriving the phase behaviour of materials with large vibrational free-energy contributions to metallic Li. The AIRSS approach [220, 204] is found to quickly pick out relevant low-energy phases with little effort, including both the BCC phase and numerous close-packed phases. We have found that DFT predicts a zero-temperature FCC phase and a room temperature BCC phase; a metastable transition from BCC to the 9R phase is also predicted in the region of FCC stability. These results are consistent with
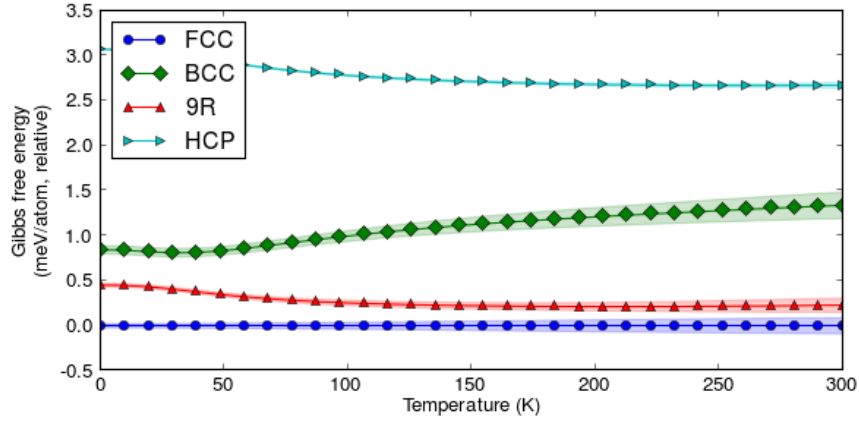
Figure 4.10: As Fig. 4.9, but neglecting the entropic contribution to the phonon free energy. With this modification we see that the phase transitions in the 0–300 K range disappear.
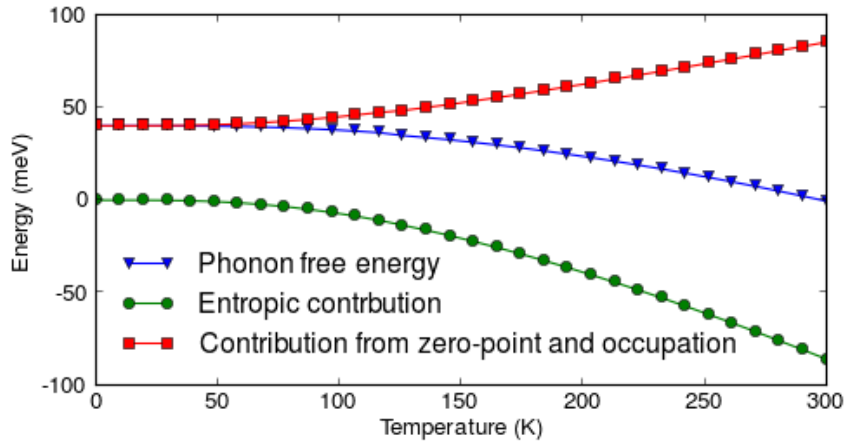
Figure 4.11: Phonon contributions to the Gibbs free energy of BCC Li.

experimental results, which show a (partial) Martensitic transition from BCC to various close-packed forms on cooling [9, 246, 190, 235]. Experiments also show that the FCC phase is stable at low temperatures, but may be formed via a high-pressure pathway [9]. The predicted FCC $\rightarrow$ BCC transition temperature is found to be extremely sensitive to changes in the Gibbs free energy landscape ($\sim$ 117 K/meV). As a result, effects which would normally be considered negligible become potentially important, including the effects of anharmonic vibrations. We calculate an anharmonic correction of up to 0.4 meV/atom, which increases the predicted FCC $\rightarrow$ BCC transition temperature by 15 K to 232 K, but find that it does not qualitatively change the phase diagram. Our calculations show that the N-point phonon in BCC Li is dynamically stable, contrary to the case for BCC Zr in Ref. [223, 46] and is around 30% stiffer when anharmonic effects are included. We calculate the various contributions to the Gibbs free energy and find that entropic contributions to the phonon free energy are of crucial importance, without which phase transitions disappear entirely.

Chapter 5

Stability and superconductivity of rare earth decahydrides

5.1 Motivation

Hydrogen was predicted to be a room-temperature superconductor at very high pressure in 1968 [18], but the pressures required to metallise hydrogen are difficult to obtain [191, 187, 74, 53, 56, 181]. Hydrides have been suggested to have lower metallisation pressures than pure hydrogen due to *chemical pre-compression* [16] and therefore might become superconducting at more readily accessible pressures. This idea has motivated a surge of research examining potential superconductivity in high-pressure hydrides, with several reviews summarising recent developments [66, 304, 87, 35, 206, 203, 218].

Theoretical studies of ScH_3, LaH_3 [70], YH_3, YH_4 and YH_6 [136, 162, 110] identified hydrides of rare-earth elements as potential high-temperature superconductors. First-principles structure searching studies of rare-earth hydrides have reported structures with high hydrogen content adopting cage-like structures [165, 213]. Of particular note, a T_c of 264-286 K was calculated for $Fm\bar{3}m$ LaH_{10} at 210 GPa [165], while the analogous YH_{10} structure was calculated to have $T_c = 305$-326 K at 250 GPa. Slight distortions of the cubic LaH_{10} phase were found to lead to $C2/m$ and $R\bar{3}m$ structures at lower pressures [98, 164], though Ref. [79] showed that quantum effects render $Fm\bar{3}m$ as the true ground state. These predictions were followed by experimental measurement of critical temperatures reaching 260 K in LaH_{10} at 170-200 GPa [249, 62]. The high-T_c phase was determined to be a structure with an fcc arrangement of La atoms, lending support to theoretical predictions.

In addition to the aforementioned studies, others have focused on heavier rare-earth hydrides, exploring the synthesis and superconducting properties of cerium [231, 161], praseodymium [300] and neodymium [298] hydrides. Here, within the

framework of density functional theory (DFT) [115, 142], we revisit LaH_{10} and YH_{10} using crystal structure prediction methods. We find a phase transition to a new hexagonal phase in LaH_{10} at high pressures, with the metastability of this phase at low pressures offering an explanation for the experimental observation of hcp impurities in fcc samples [62]. We go on to predict the phases and corresponding critical temperatures that may be observed in YH_{10}.

5.2 Theory and methodology

5.2.1 Phonons and superconductivity

The Hamiltonian of a coupled electron-phonon system [102] can be written (see Sec. 2.6) as

$$H = \underbrace{\sum_{kn} \epsilon_{nk} c_{nk}^{\dagger} c_{nk}}_{\text{electronic dispersion}} + \underbrace{\sum_{q\nu} \omega_{q\nu} \left(a_{q\nu}^{\dagger} a_{q\nu} + \frac{1}{2} \right)}_{\text{phonon dispersion}} + \underbrace{\frac{1}{\sqrt{N_p}} \sum_{kqmn\nu} g_{mn\nu}(k,q) c_{m,k+q}^{\dagger} c_{nk} \left(a_{q\nu} + a_{-q\nu}^{\dagger} \right)}_{\text{electron-phonon coupling}}. \tag{5.1}$$

In this work, we calculate the electronic Kohn-Sham eigenvalues ϵ_{nk}, phonon frequencies $\omega_{q,\nu}$, and electron-phonon coupling constants $g_{mn\nu}(k,q)$ appearing in H from first-principles using the QUANTUM ESPRESSO DFT code [100, 101]. The Hamiltonian in Eq. 5.1 can be treated within Migdal-Eliashberg theory [71], allowing us to define the electron-boson spectral function

$$\alpha^2 F(\omega) = \frac{1}{N(\epsilon_F)} \sum_{mnq\nu} \delta(\omega - \omega_{q\nu}) \sum_{k} |g_{mn\nu}(k,q)|^2 \times \delta(\epsilon_{m,k+q} - \epsilon_F) \delta(\epsilon_{n,k} - \epsilon_F). \tag{5.2}$$

From $\alpha^2 F$ we extract the superconducting critical temperature by solution of the Eliashberg equations using the ELK code [1]. From the quantities appearing in H we may also construct the electronic and vibrational densities of states, from which we can derive the Gibbs free energy as a function of temperature (as we did in Chapter 4). We do this at a range of pressures, allowing us to construct pressure-temperature phase diagrams.

To evaluate the double-delta sum in Eq. 5.2 for finite $\mathbf{k}$- and $\mathbf{q}$-point grids, we follow the method detailed in Appendix A of Ref. [279] and smear the delta functions with finite-width Gaussians. In order to best approximate the delta functions, the smallest sensible smearing should be used. However, the smearing must be large enough to accommodate the finite $\mathbf{k}$-point grids used. We identify the optimal choice of smearing from discrepancies in the results between different $\mathbf{k}$-point grids [279], as can be seen in Figs. 5.1 and 5.2.
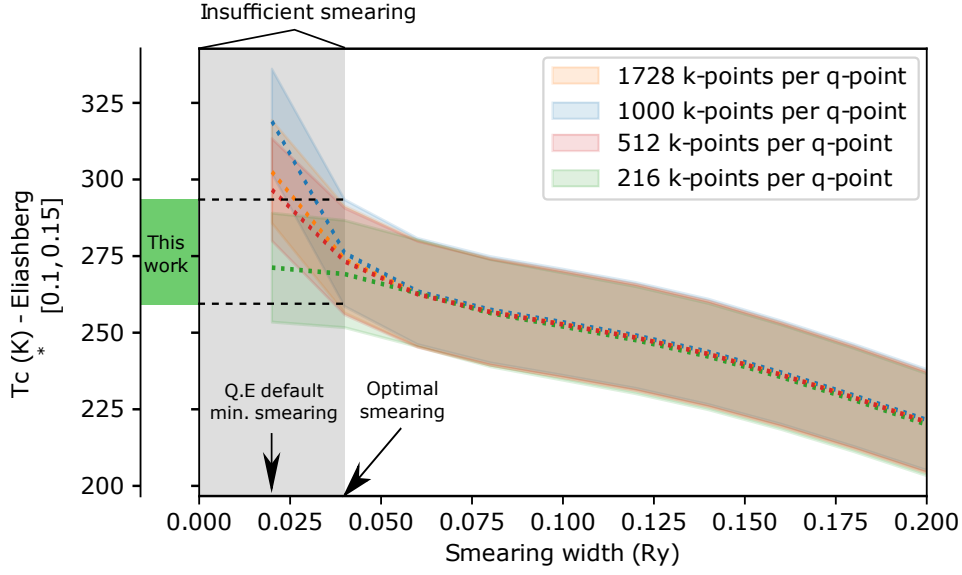
Figure 5.1: The dependence of T_c on the double-delta smearing width, σ, for $Fm\bar{3}m$-YH$_{10}$ at 350 GPa. The region of insufficient smearing is shown, along with our choice of σ for this structure and pressure. The smallest value used in an electron-phonon calculation with default QUANTUM ESPRESSO settings is shown.
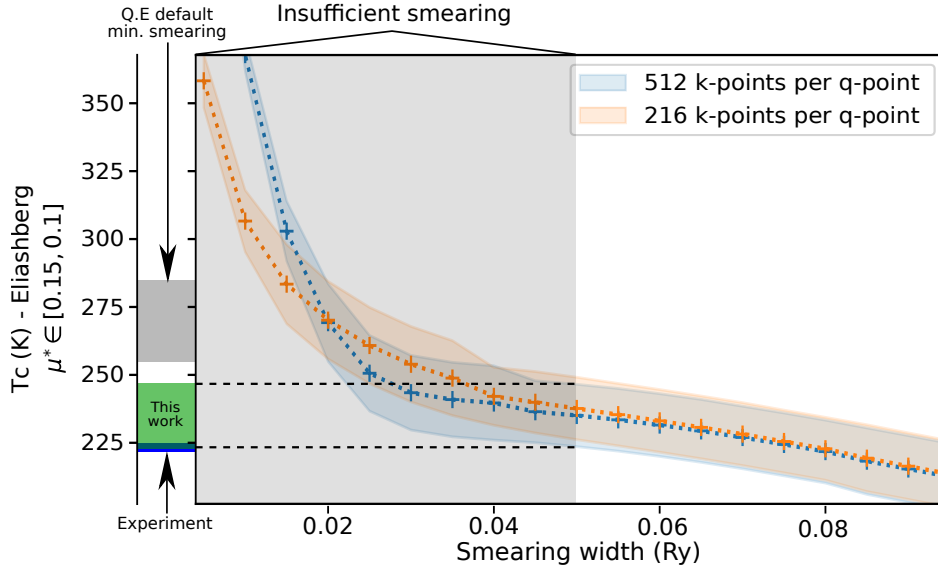
Figure 5.2: The dependence of T_c on the double-delta smearing width, σ, for $Im\bar{3}m$-YH$_6$ at 160 GPa. A recent experimental measurement at 166 GPa, falling just within our calculated T_c range, is also shown [265]. We note that Refs. [265, 145] highlighted that previous calculated T_c values were considerably higher than their experimental observations and that the results of Ref. [110], which used accurate Wannier interpolation techniques, are in agreement with ours.

Our electron-phonon calculations were carried out using the Perdew-Burke-Ernz-

erhof (PBE) generalised gradient approximation [215] and ultrasoft pseudopotentials, validated against the all-electron WIEN2k code [34]. Well-converged **k**-point grids with a spacing of at most $2\pi \times 0.015$ Å^{-1} and an 820 eV plane wave cut-off were used. The **q**-point grids used were typically 8 times smaller than the **k**-point grids and were Fourier interpolated to 10 times their original size. For the cubic systems studied, this corresponds to $\geq 24 \times 24 \times 24$ **k**-point grids and a $3 \times 3 \times 3$ **q**-point grid Fourier-interpolated to $30 \times 30 \times 30$.

5.2.2 Structure searching

Our structure searching calculations were performed using *ab initio* random structure searching (AIRSS) [219, 203] and CASTEP [48]. The PBE functional, CASTEP QC5 pseudopotentials, a 400 eV plane wave cut-off and a **k**-point spacing of $2\pi \times 0.05$ Å^{-1} were used in these searches unless otherwise stated. The C2X software [228] was used for converting between CASTEP and QUANTUM ESPRESSO file formats, and also for reporting the space groups of structures at various tolerances.

5.3 Results and discussion

In the following sections, we report results in terms of phonon-corrected pressures, obtained by fitting the Birch-Murnaghan equation of state [33] to our data. Where static DFT pressures are reported instead, they are labelled as P_{DFT} - this second set of pressures facilitates comparison with previous calculations as they are simply an input parameter to the DFT geometry optimisation.

5.3.1 LaH$_{10}$

Low enthalpy candidates found by AIRSS for LaH$_{10}$ include the space groups $Fm\bar{3}m$, $R\bar{3}m$, and a 2-formula-unit $C2/m$, which had been identified previously. The searches also revealed a new structure of $P6_3/mmc$ symmetry. These structures are shown in Fig. 5.3. A $C2/m$ structure with 3 formula units per unit cell was also found to be energetically competitive, but was not considered further in this work as it behaves similarly to the 2-formula unit phase over the pressure range of interest. We also found several previously unreported structures at low pressures with space groups $Pnnm$, $C2$ and $P2_12_12_1$ and unit cells containing 2, 3 and 4 formula units, respectively. These are the lowest enthalpy structures in the low pressure region, but are distortions of the high-symmetry $Fm\bar{3}m$ structure and, similarly to the case of $R\bar{3}m$ noted in Ref. [79], it is possible that anharmonic effects may remove them from the potential energy surface. In addition to this, the low symmetry and
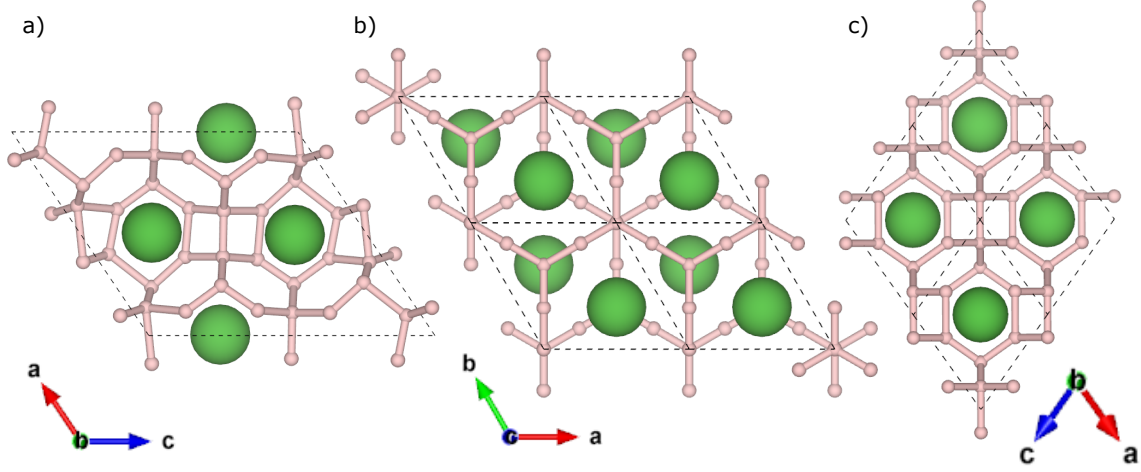
Figure 5.3: Structures of LaH$_{10}$. (a) 2 formula unit/cell $C2/m$, (b) 2 formula unit/cell $P6_3/mmc$, (c) 1 formula unit/cell $Fm\bar{3}m$. The $R\bar{3}m$ structure is not shown as it is visually indistinguishable from the $Fm\bar{3}m$ structure at the pressures of interest.

large unit cells of these structures make converged phonon calculations prohibitively expensive; they are therefore not considered further in this work.

The calculated LaH$_{10}$ phase behaviour is shown in Fig. 5.4 and the corresponding critical temperatures are shown in Fig. 5.5. Our calculations for the $Fm\bar{3}m$ phase include unstable phonon modes for $P_{DFT} \leq 210$ GPa. In the harmonic picture, explicitly taking into account this dynamical instability leads to a window of stability for the $C2/m$ phase, in agreement with previous calculations [98, 164]. However we note that, under the assumption that the unstable modes can be neglected in the calculation of the Gibbs free energy, we obtain the same behaviour as the anharmonic calculations of Ref. [79], i.e., $Fm\bar{3}m$ is the only phase with a predicted region of stability at lower pressures. With increasing pressure, as noted in previous theoretical work [79], the $R\bar{3}m$ structure approaches $Fm\bar{3}m$ symmetry. We therefore expect that these phases will not be distinguishable at high pressures.

At 300 K, the $P6_3/mmc$ structure becomes thermodynamically favourable at pressures above $\sim$420 GPa. More importantly, this hexagonal phase is also metastable at low pressures, lying within 20 meV/atom of the cubic phase down to 150 GPa, and therefore provides an explanation for the experimental observation of hcp impurities in fcc-LaH$_{10}$ samples at 170 GPa in Ref. [62]. A low-energy hexagonal LaH$_9$ structure predicted previously in similar pressure regions [148] could offer an alternative explanation for the observation of these impurities. However, the authors of Ref. [62] determined that the two kinds of hcp impurities in their fcc-LaH$_{10}$ samples possessed LaH$_{10}$ stoichiometry. We also calculated a high-quality La-H convex hull at 150 GPa using AIRSS [219] and qhull [23] (see Fig. 5.6). It shows that the

93
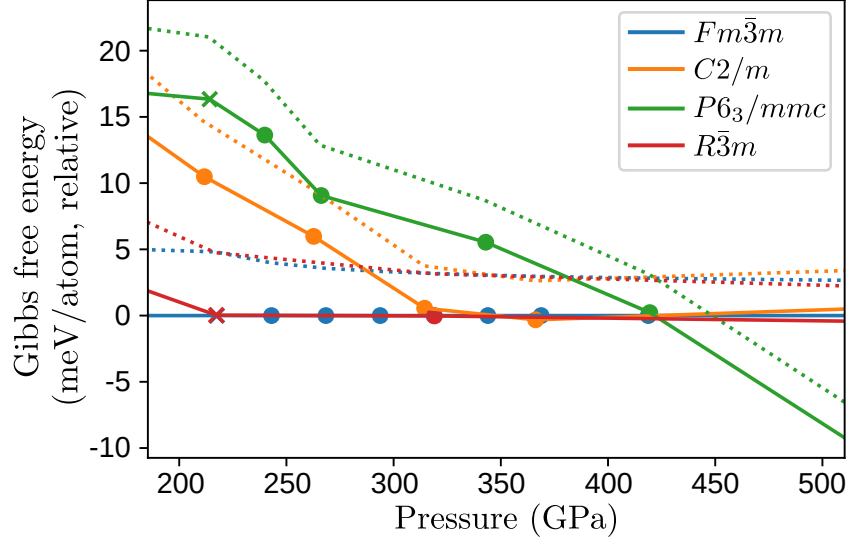
Figure 5.4: The Gibbs free energy as a function of pressure for energetically competitive phases of LaH$_{10}$, plotted relative to a third-order Birch-Murnaghan fit of the $Fm\bar{3}m$ data. Crosses represent calculations with unstable phonon modes - these points are not included in the Gibbs free energy fit. Solid lines are at 300 K, dashed lines are at 0 K.
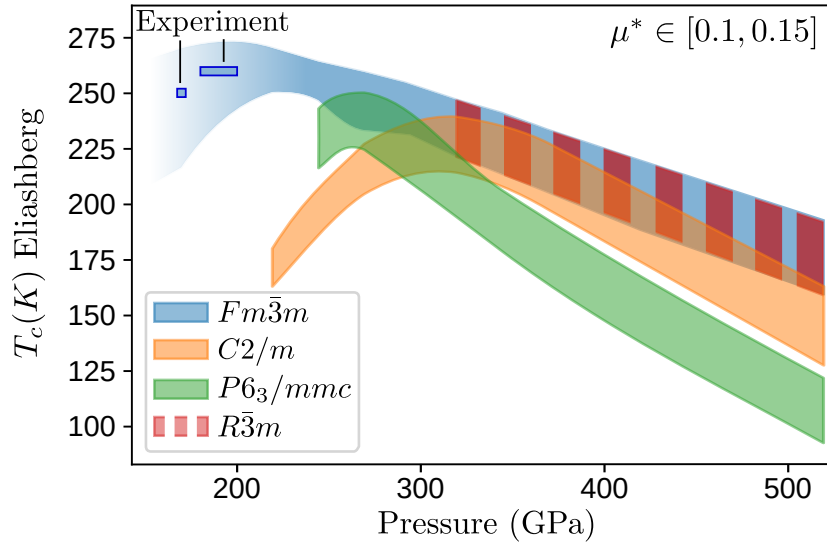
Figure 5.5: Calculated $T_c(P)$ for dynamically stable phases of LaH$_{10}$ from direct solution of the Eliashberg equations. The width of the lines arises from our treatment of the Morel-Anderson pseudopotential, μ^*, [200] as an empirical parameter with values between 0.1 and 0.15. The $Fm\bar{3}m$ result has been extended into the region where it is dynamically unstable (shaded according to unstable fraction of the phonon density of states) in order to facilitate comparison with the experimental results of Refs. [62, 249]. This extension was achieved by removing the contribution of unstable phonon modes, in their entirety, to the Eliashberg function while maintaining its normalisation.

94

$P6_3/mmc$-LaH$_{10}$ structure predicted in this work lies closer to the hull than the $P6_3/mmc$-LaH$_9$ structure of Ref. [148]. It is therefore likely that the hcp impurities originate from our new $P6_3/mmc$-LaH$_{10}$ phase.

To facilitate comparison with experiment, we have calculated powder X-ray diffraction patterns for cubic and hexagonal LaH$_{10}$ and hexagonal LaH$_9$ at 150 GPa - these are shown in Fig. 5.7. At 150 GPa, we find a c/a ratio of 1.526 for $P6_3/mmc$-LaH$_{10}$ and of 1.564 for $P6_3/mmc$-LaH$_9$. At the same pressure, the volume per formula unit for $P6_3/mmc$-LaH$_{10}$ is 33.15 Å^3 and for $P6_3/mmc$-LaH$_9$ is 31.73 Å^3.

We calculate T_c = 232-259 K for $Fm\bar{3}m$-LaH$_{10}$ at 269 GPa (P_{DFT}=250 GPa), which is lower than the previous theoretical result of T_c = 257-274 K [165]. However, we observe an increase in T_c on reduction of the double-delta smearing parameter to below our calculated optimal value, potentially explaining this discrepancy. Careful choice of smearing has previously been noted as important in other hydride systems [107]. We also note a previous calculation of T_c for this structure at 200 GPa [213], however, in agreement with other calculations [164, 98] we find $Fm\bar{3}m$ to be dynamically unstable at this pressure. This dynamical instability means we cannot directly compare with experiment, which found T_c = 250 K at around 170 GPa [62] and T_c = 260 K at 180-200 GPa [249]. However, ignoring the contribution of the unstable phonon modes to the Eliashberg function at pressures ≤ 210 GPa allows for a rough estimation of T_c in these regions; this is depicted as the faded-out section in Fig. 5.5 and the results obtained are in agreement with experimental results. For the $C2/m$ phase, using an optimal value of smearing we calculate T_c = 205-225 K at 262 GPa (P_{DFT}=250 GPa), compared to T_c = 229-245 K in Ref. [164].

5.3.2   YH$_{10}$

Low-enthalpy candidates for YH$_{10}$ found using AIRSS include $Fm\bar{3}m$, which had been identified previously, a slight distortion of this phase, $R\bar{3}m$, and structures of $P6_3/mmc$ and $Cmcm$ symmetry. These structures are shown in Fig. 5.8. The calculated YH$_{10}$ phase behaviour is shown in Fig. 5.9 and the corresponding critical temperatures are shown in Fig. 5.10. We do not predict any phase transitions within the range of stability of the YH$_{10}$ stoichiometry [213]. However, the difference in Gibbs free energy between the $Fm\bar{3}m$ and $R\bar{3}m$ phases is exceedingly small (see Fig. 5.9), reflecting their structural similarity.

Previous calculations for $Fm\bar{3}m$-YH$_{10}$ found T_c = 305-326 K at 250 GPa [165] and T_c = 303 K at 400 GPa [213]. Here, we calculate T_c = 270-302 K at 324 GPa (P_{DFT}=300 GPa) and T_c = 250-280 K at 425 GPa (P_{DFT}=400 GPa). Our more conservative T_c results can again be explained by considering the smearing parameter used to approximate the double-delta integral in Eq. 5.2. We were able to reproduce
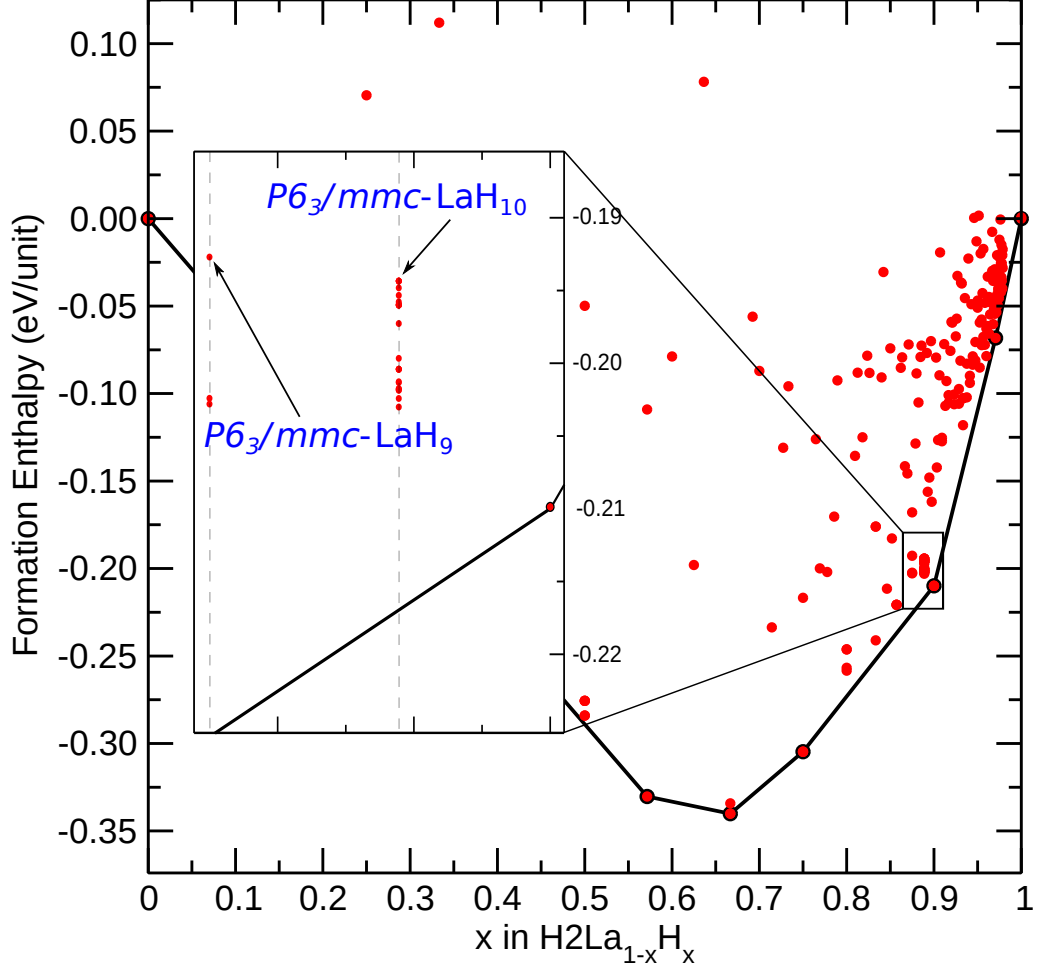
Figure 5.6: A convex hull for the La-H system at 150 GPa, accurately calculated using CASTEP, **k**-point spacing of $2\pi \times 0.03$ Å^{-1} and a 700 eV plane-wave cut-off. Default on-the-fly pseudopotential strings were used for H, but the inclusion of a fraction of a 4f electron in the generation of the La pseudopotential was found to be crucial leading to the modified string 2|2.3|5|6|7|50U:60:51:52:43{4f0.1}(qc=4.5)[4f0.1]. A pseudopotential without this addition was unable to reproduce the all-electron $Fm\bar{3}m$-LaH$_{10}$ PV curve and led to a qualitatively different convex hull. In agreement with Ref. [148], we find that LaH$_9$ is not on the hull at this pressure. However, we also find that LaH$_{16}$ does not lie on the hull at 150 GPa, despite finding the $P6/mmm$-LaH$_{16}$ structure studied in that work.
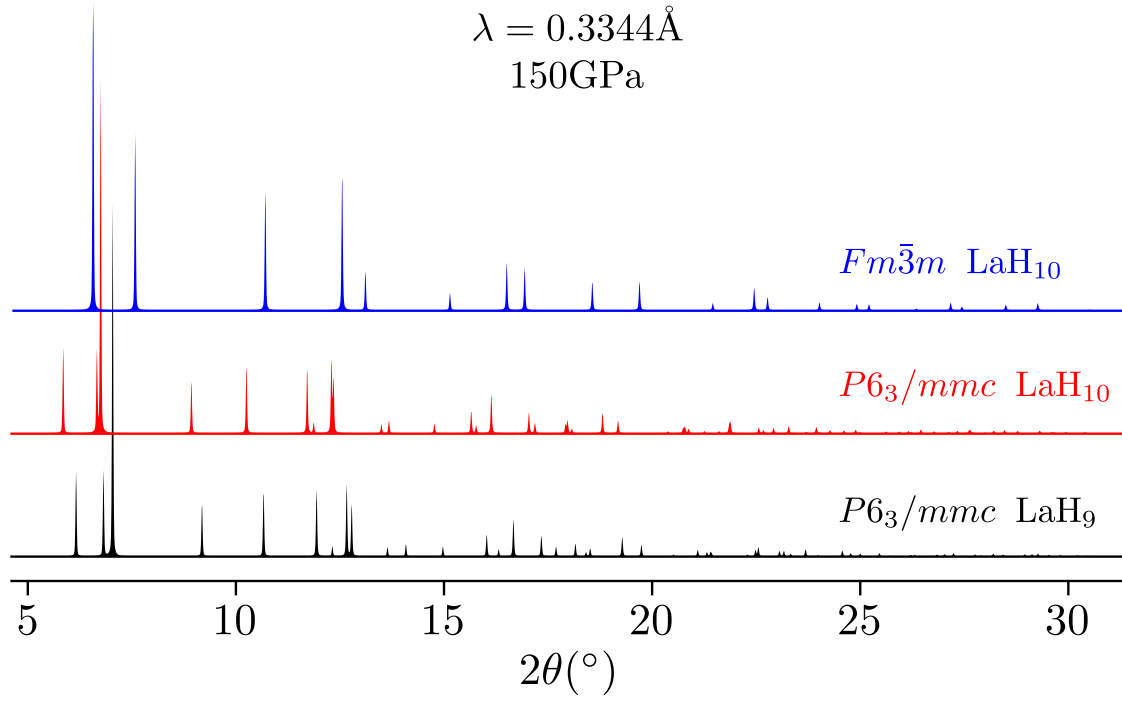
Figure 5.7: Simulated X-ray powder diffraction patterns for the $Fm\bar{3}m$ structure of LaH_{10} and potential hexagonal impurities.
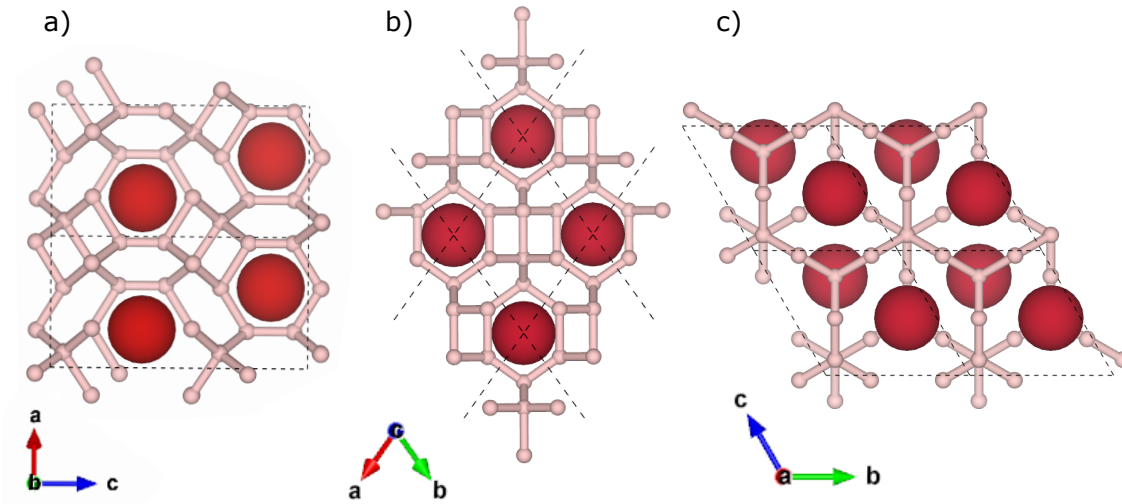
Figure 5.8: Structures of YH_{10}. (a) 2 formula unit/cell $Cmcm$, (b) 1 formula unit/cell $Fm\bar{3}m$, (c) 2 formula unit/cell $P6_3/mmc$. The $R\bar{3}m$ structure is, again, not shown.
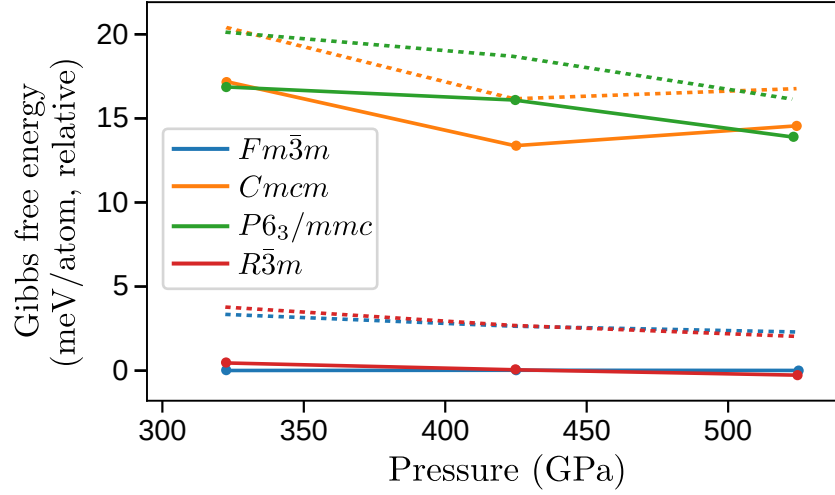
Figure 5.9: The Gibbs free energy as a function of pressure for energetically competitive phases of YH_{10}, plotted and interpolated relative to a third-order Birch-Murnaghan fit of the $Fm\bar{3}m$ data. Solid lines are at 300 K, dashed lines are at 0 K.
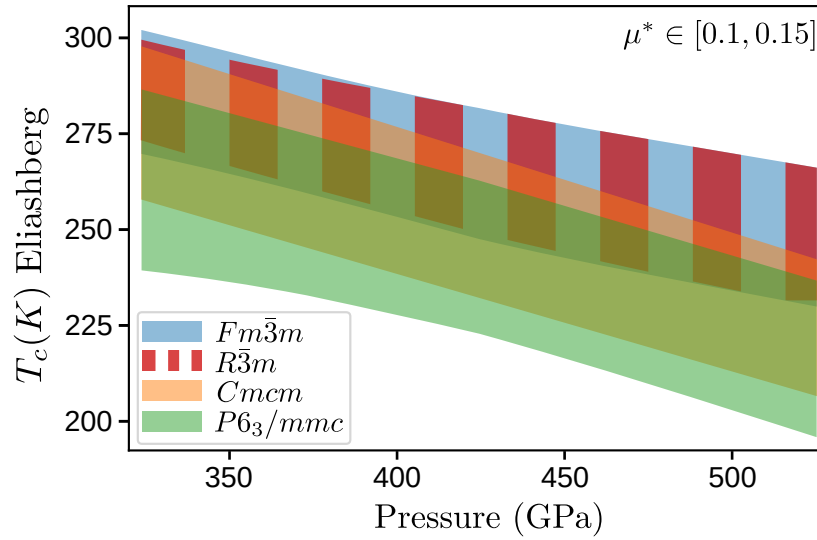
Figure 5.10: Calculated $T_c(P)$ for dynamically stable phases of YH_{10} from direct solution of the Eliashberg equations. μ^* is, again, taken to have a value between 0.1 and 0.15.

the results of previous calculations by using the minimum default smearing employed in QUANTUM ESPRESSO, which in this case overestimates T_c by $\sim$30 K (see Fig. 5.1) compared to optimal smearing. We note that our results agree with those obtained using Wannier interpolation techniques [110]. Using the same method to calculate an optimal smearing also provides results in agreement with recent experimental measurements for $Im\bar{3}m$-YH$_6$ [265], as shown in Fig. 5.2.

5.4   Conclusions

We have identified a new hexagonal phase of LaH$_{10}$ with $P6_3/mmc$ symmetry. Our calculations show a pressure-induced phase transition into this new phase from the cubic phase believed to be observed in experiment [249, 62]. The overall phase behaviour predicted within the harmonic picture is $C2/m \to Fm\bar{3}m \to P6_3/mmc$ with all three of these phases predicted to be high-T_c superconductors. Making the assumption that unstable modes can be neglected, however, gives the same picture as the anharmonic calculations of Ref. [79] where $Fm\bar{3}m$ is the true ground state at lower pressures. The new hexagonal phase predicted here offers a direct explanation for the observation of hcp impurities in recent experiments [62].

We found that YH$_{10}$ adopts very similar structures to LaH$_{10}$, with one of $P6_3/mmc$ symmetry again amongst the most energetically competitive candidates. Over the pressure range considered the $Fm\bar{3}m/R\bar{3}m$ phase remains the most stable. The difference in Gibbs free energy between these two structures is extremely small, meaning synthesis of a pure sample of either could be difficult.

We found the double-delta smearing employed in superconductivity calculations to be of particular importance. Its effect on calculated T_c changes from system to system; in particular, in our calculations the default minimum smearing employed by QUANTUM ESPRESSO overestimates T_c for LaH$_{10}$ by $\sim$20 K and YH$_{10}$ by $\sim$30 K when compared to optimal smearing.

Chapter 6

Machine learning for novel superconducting hydrides

6.1 Motivation

In the previous chapter, we demonstrated a methodology to investigate the stability and superconductivity of a particular target system. In particular, we looked at rare earth hydrides X in the stoichiometry XH_{10}. However, in recent years, potential superconductivity has been investigated in many compressed hydrides, including scandium [70], sulfur [65, 59, 76], yttrium [136, 162, 165, 213, 110, 265, 145, 243], calcium [274], actinium [236], thorium [151], pnictogen [91], praseodymium [300], cerium [231, 161], neodymium [298], lanthanum [165, 213, 249, 62, 243, 148] and iron hydrides [184, 152, 107]. Several reviews summarising recent developments in the field are available [66, 304, 87, 35, 218, 31]. Inspired by known superconductors, researchers have also attempted to increase T_c by chemical means; replacing atoms in known structures and assessing stability and superconductivity [44], doping known binaries with more electronegative elements to make ternary hydrides [255], and mapping alchemical phase diagrams [108]. However, the search for superconducting hydrides has, so far, largely focused on finding materials exhibiting the highest possible critical temperatures (T_c). This has led to a bias towards materials stabilised at very high pressures, which introduces a number of technical difficulties in experiment.

In this chapter, we train machine learning models on a set of literature data for superconducting binary hydrides. Machine learning has previously been used in modelling hydride superconductors, with a focus on predicting the maximum obtainable critical temperature for a given composition [238]. However, on examination of the literature (see Fig. 6.1), it becomes apparent that the pursuit of superconductivity close to ambient conditions is as much about reducing the required pressure as it
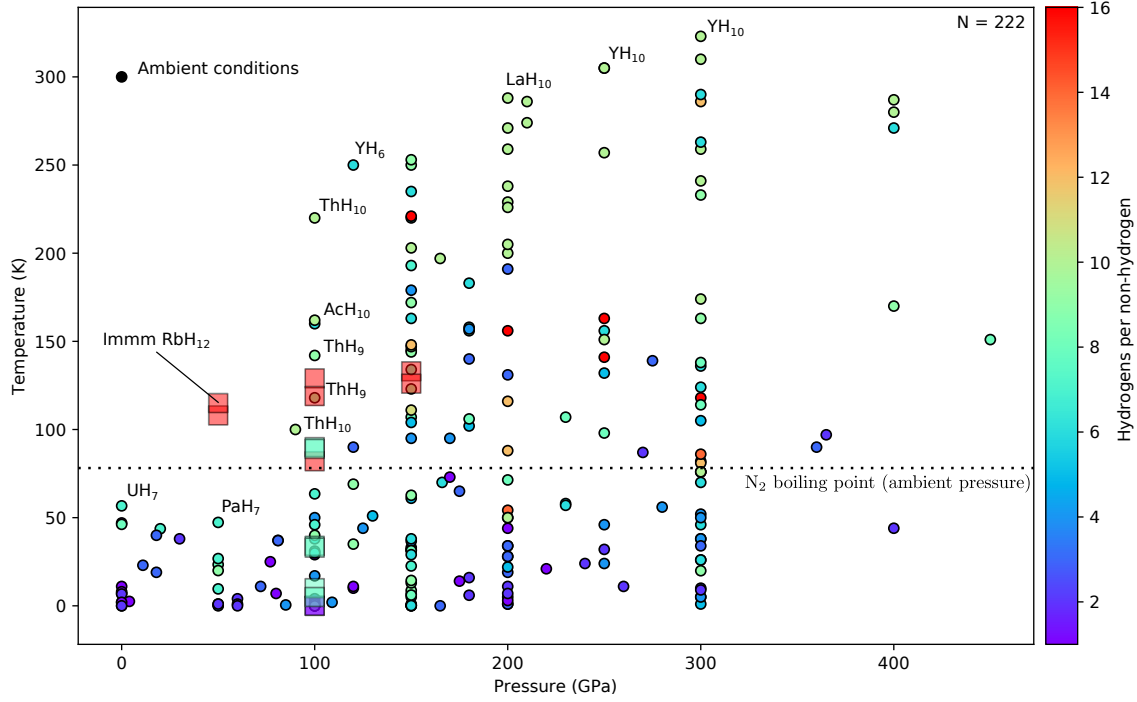
Figure 6.1: The critical temperatures of binary hydrides at various pressures found in the literature are shown as circles. Materials on the frontier towards ambient conditions are labeled. Multiple points with the same stoichiometry arise from T_c for a particular phase calculated at different pressures, or from different structural phases of the same material. New structures found in this work, with T_c calculated using DFPT (as reported in Table 6.1), are shown as translucent squares; of note is $Immm$-RbH$_{12}$ (labeled, see also Fig. 6.6), which extends the frontier significantly.

is about increasing the critical temperature. This is especially important given that working at high pressure can often present a far greater experimental challenge than working at low temperature. In this work, we therefore model critical temperature and operational pressure on an equal footing. Our models are used to inform the choice of composition for crystal structure searches and subsequent electron-phonon calculations, with the aim of extending the operation of hydride superconductors towards ambient conditions.

6.2 Trends in Hydrides

A large amount of computational - and some experimental - data for the binary hydrides is available in the literature [98, 265, 79, 110, 145, 281, 297, 285, 178, 267, 236, 151, 274, 290, 149, 84, 106, 70, 82, 278, 300, 168, 239, 213, 165, 289, 242, 157, 147, 238, 171, 208, 45, 95, 302, 287, 251, 156, 152, 173, 171, 284, 245, 160, 77, 258, 123, 277, 121, 159, 257, 118, 171, 283, 303, 73, 163, 131, 164, 60, 249, 62, 162]

(values from these references form our dataset, shown in Fig. 6.1). In some subsets of hydrides certain material properties show a simple dependence on the properties of the non-hydrogen element. For example, in the alkaline earth hydrides the van der Waals radius of the ion is well correlated with the metallization pressure [291]. However, obtaining strong electron-phonon coupling at low pressures is, in general, a more complicated process; simple correlations between composition and operational pressure or critical temperature are therefore absent in the dataset as a whole. We look at more complicated trends by constructing machine learning models of critical temperature and operational pressure which take as input a set of easily-obtained material descriptors. For a particular element E and corresponding binary hydride EH_n these descriptors are

- hydrogen content (n)
- van der Waals radius of E
- atomic number of E
- mass number of E
- numbers of s, p, d and f electrons in the (atomic) electron configuration of E

Once constructed, we apply the model to all materials with the chemical composition EH_n, where E is any element in the periodic table and $n \in [1, 2, \ldots, 32)$[1]. From these, the materials which are predicted to exhibit superconductivity closest to ambient conditions serve as a guide for searches for new binary hydrides.

6.2.1 Neural network

We train a fully-connected neural network (using the Keras frontend to the Tensorflow machine-learning library [47, 189]), with the topology shown in Fig. 6.2, on the dataset shown in Fig. 6.1. The squared absolute error $|(\Delta T_c, \Delta P)|^2$ between the predicted and literature values serves as our cost function, which we minimize using the *Adam* stochastic optimizer [137]. The input (and expected output) data is positive definite and therefore has a non-zero mean and is not normally distributed, prompting the use of self-normalizing activation functions [140, 49] to improve training behaviour. Since the number of data points is comparable to the number of parameters in our network, the risk of over-fitting becomes significant. To mitigate this, we split the data into a randomly selected validation set (consisting of 25% of the initial data points) and a training set (consisting of the other 75%). Once the model starts over-fitting to the training data the validation set error starts to increase, allowing us to choose the model parameters from the training epoch for

---

[1]A maximum of 31 hydrogens per atom was chosen to avoid over-extrapolation from the dataset (where the maximum is 16).
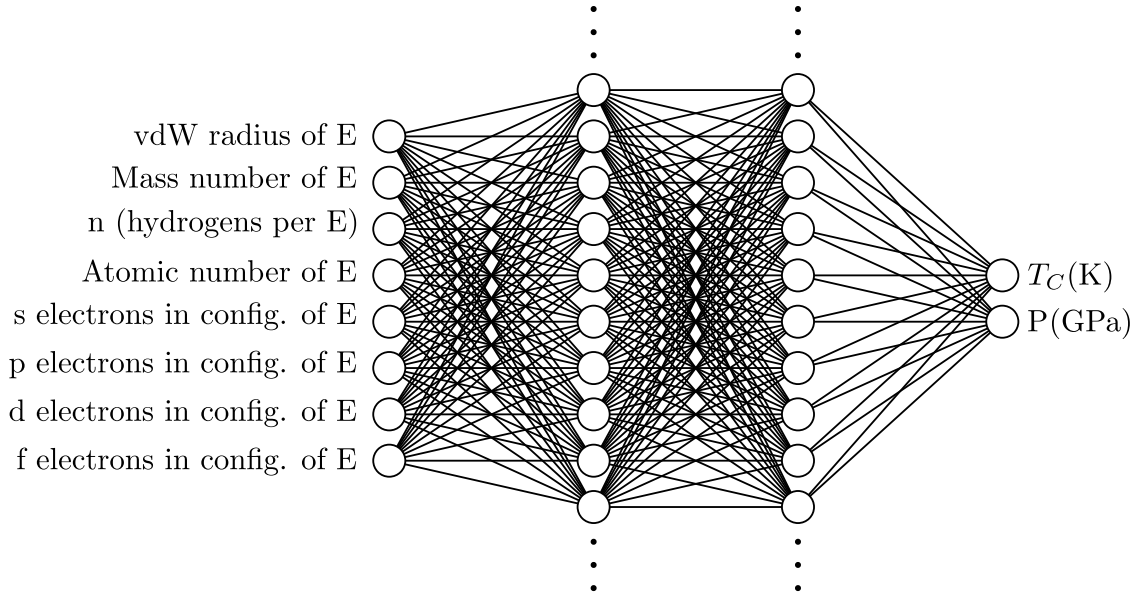
Figure 6.2: Topology of our neural network model. An input layer is fed to the material descriptors for the hydride EH_n, one per input node. This layer then feeds two densely-connected intermediate layers (of 32 nodes each), the last of which feeds the output layer with one temperature node and one pressure node.

which the validation set error is minimal. We cross-validate the results by repeating this process 64 times and averaging the predictions - this is an approximation of leave-p-out cross-validation with $p = 25\%$ of the dataset. We also apply L_2 regularization to the parameters in the intermediate dense nodes to decrease the propensity for over-fitting, improving the convergence of this cross-validation scheme.

6.2.2 Model behaviour

The basic behaviour of the machine learning model is shown in Fig. 6.3. We see that it achieves reasonable correlation with the literature values and predicts sensible pressures and temperatures for unseen materials. To gain insight into properties which favour ambient-condition superconductivity we define a measure of distance $D = |(P, T_c - 293)|_1$. This distance decreases as we move towards ambient conditions from the pressure-temperature region containing the known hydrides (see Fig. 6.1). In Fig. 6.4 we plot the distribution of material properties for the 10% of hydrides predicted to exhibit superconductivity closest to ambient conditions (i.e., the 10% with lowest D). We can see that the model predicts the heavy alkali and alkaline earth metal hydrides to be the best candidates, with the number of close-to-ambient materials then decreasing as we go across each period. The distribution of the number of hydrogen atoms is more uniform, suggesting it is necessary to consider a range of different stoichiometries for each composition. These conclusions are

reinforced by the construction of a generalised linear regression model (see Appendix. A.3), which reproduces the general trends exhibited by the machine learning model (but, unsurprisingly, exhibits worse correlation with the literature values). The predicted optimal (minimum D) hydride compositions from the machine learning model are shown for each element of the periodic table in Fig. 6.5.

We note that the points included in our dataset will be of varying quality, come from different research groups, and are of both experimental and theoretical origin. The majority are theoretical and calculated within the harmonic approximation. Although it has been shown that anharmonicity can affect the calculated critical temperature for hydrides [75, 76], there is insufficient data in the literature to build a model exclusively from anharmonic results. However, since we only seek to extract general trends, which will serve simply to inform areas of focus for structure searching, the dataset is sufficient for our purposes.

6.3 Structure searching

The models constructed in the previous section point towards the alkali and alkaline earth metal hydrides as some of the best candidates for superconductivity near ambient conditions. From these, we studied caesium and rubidium hydrides; these systems were chosen due to their predicted proximity to superconductivity at ambient conditions (see Figs. 6.4 and 6.5) and the fact that they have not been studied extensively in the past, unlike the hydrides of other elements in these two groups. Caesium and rubidium polyhydrides have been studied previously using structure searching methods in Refs. [240] and [117], respectively, although potential superconductivity was not investigated in either case.

Our structure searching calculations were performed using *ab initio* random structure searching (AIRSS) [219, 203] and the plane-wave pseudopotential code CASTEP [48]. Since our models suggest that a wide range of stoichiometries should be considered, convex hulls were constructed using AIRSS and qhull [23] in order to identify those which are stable at 50, 100 and 200 GPa. The Perdew-Burke-Ernzerhof (PBE) generalised gradient approximation [215], CASTEP QC5 pseudopotentials, a 400 eV plane-wave cut-off and a $\mathbf{k}$-point spacing of $2\pi \times 0.05$ Å^{-1} were used in all searches. The Cs-H convex hulls calculated in this work at 100 and 200 GPa both partially agree with the hull calculated at 150 GPa in Ref. [240]. Once stable stoichiometries had been identified, additional AIRSS searches for RbH_3, RbH_5, RbH_9, RbH_{11}, RbH_{12}, CsH_5, CsH_7, CsH_{13} and CsH_{15} using the same parameters and pseudopotentials were performed at 100 and 200 GPa.
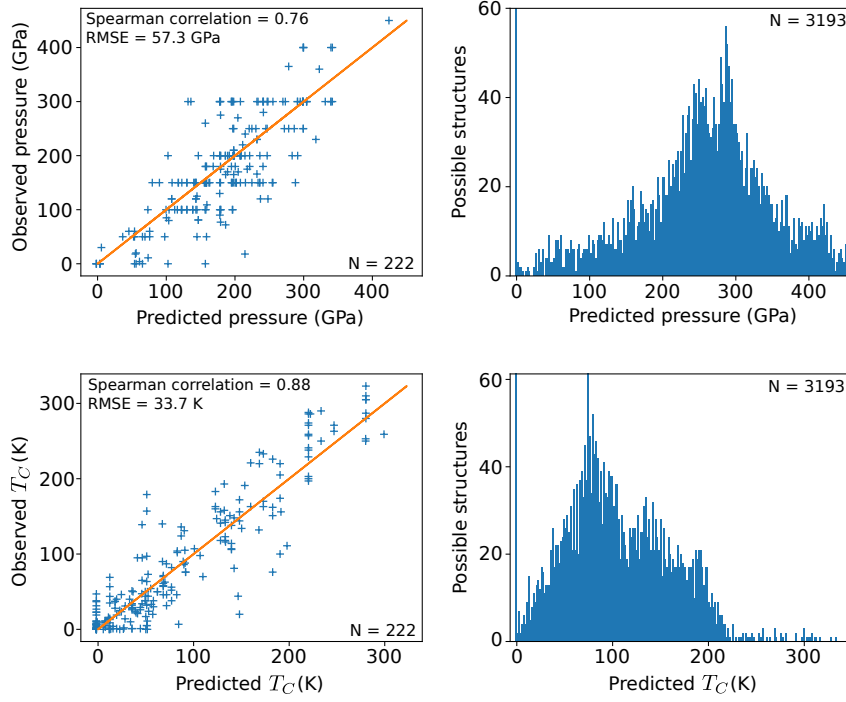
Figure 6.3: Behaviour of our machine learning model of critical temperatures and associated operational pressures for binary hydrides. The correlation between the predicted and observed values for the data in the literature is shown, as well as the resulting distribution of pressures and temperatures when the model is applied to the set of all possible binary hydrides as defined in Section 6.2.
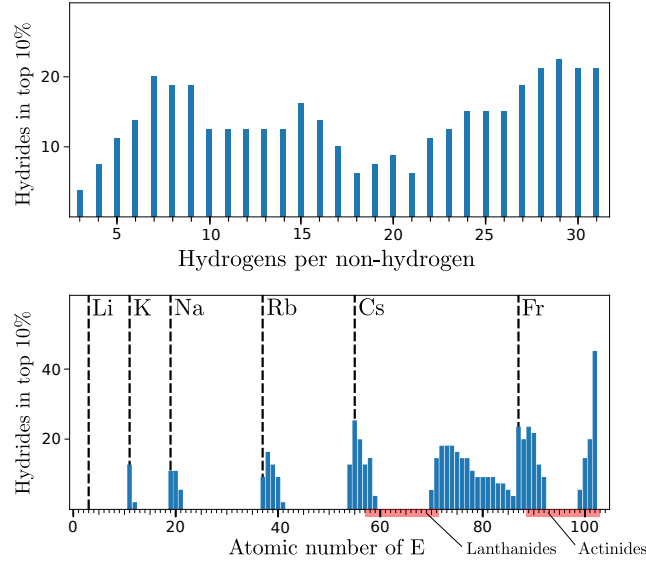
Figure 6.4: Distribution of hydrogen atoms per non-hydrogen atom and atomic number of the non-hydrogen element for the 10% of hydrides that our machine learning model predicted to exhibit superconductivity closest to ambient conditions (i.e., the 10% with lowest D). Black dashed lines indicate the atomic numbers of alkali metals.
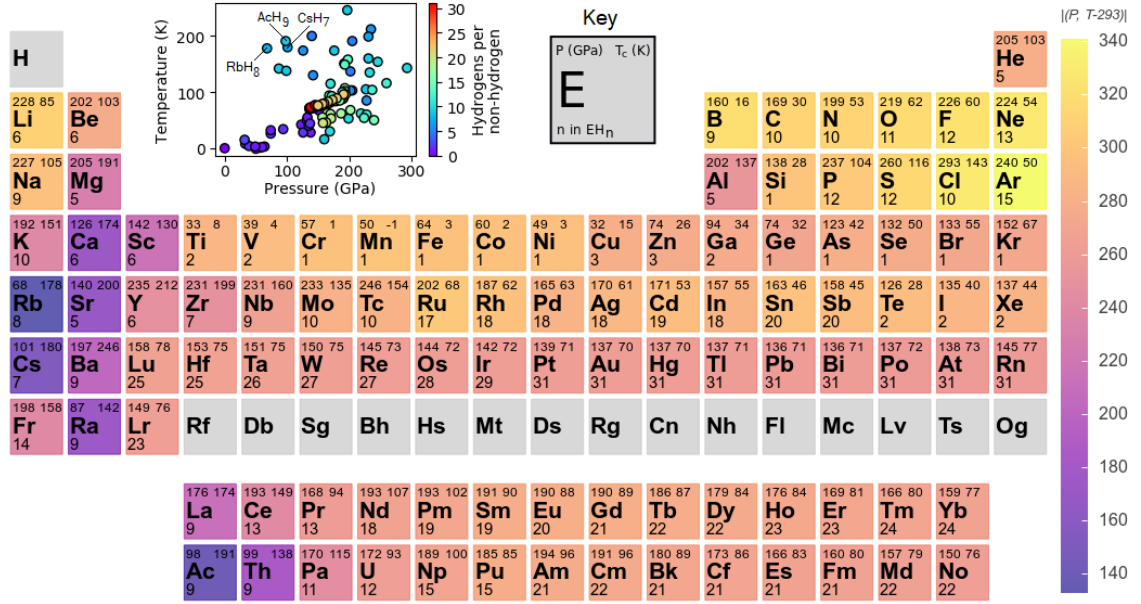
Figure 6.5: The periodic table of optimal binary hydrides according to our machine learning model. The predicted critical temperature, corresponding pressure and optimal hydrogen content is shown for each element. Elements are colored according to the predicted distance from ambient-condition superconductivity $D = |(P, T_c - 293)|$ (with pressure measured in GPa and critical temperature measured in K). Inset: the distribution in pressure-temperature space of these predictions. We note that we did not explicitly prevent the neural network from predicting negative critical temperatures and it does so for MnH. However, in general, the machine learning model has learned that critical temperatures should be positive (see the lower right panel of Fig. 6.3).

6.4 Selecting candidate structures

For each selected stoichiometry, the enthalpy was calculated as a function of pressure for the most stable structures arising from the AIRSS search. These geometry optimisations were performed using QUANTUM ESPRESSO [100, 101], the PBE functional, a 950 eV cut-off, ultrasoft pseudopotentials and a $\mathbf{k}$-point spacing of $2\pi \times 0.02$ Å^{-1}. The electronic density of states (DOS) at the Fermi energy was also evaluated for each structure at 50 GPa and 150 GPa in order to identify metallic structures. We were then able to limit our interest to structures which were both energetically competitive (according to the enthalpy plots) and had a considerable DOS at the Fermi energy in the low-pressure region (25-125 GPa). The remaining candidates, for which electron-phonon coupling calculations were performed, include $C2/m$-RbH$_{12}$, $Immm$-RbH$_{12}$, and various CsH$_7$ and RbH$_3$ structures (see Table 6.1).

6.5 Electron-phonon coupling and superconductivity

Having selected promising Cs and Rb candidates according to stability and metallicity, we go on to calculate their superconducting properties, using the same method as in the previous chapter (chapter 5). In carrying out the necessary DFPT calculations of the parameters appearing in the Hamiltonian, we use the PBE functional, the same ultrasoft pseudopotentials as in the stability screening, an 820 eV plane-wave cut-off, and a $\mathbf{q}$-point grid with a spacing of $\approx 2\pi \times 0.1$ Å^{-1} (e.g., a $2 \times 2 \times 2$ grid for a 26-atom unit cell of RbH$_{12}$). Two separate $\mathbf{k}$-point grids are used (of 6^3 and 8^3 times the size of the $\mathbf{q}$-point grid, respectively), allowing us to determine the optimal double-delta smearing width necessary to calculate the critical temperature following the method outlined in chapter 5.

The results of the critical temperature calculations are shown in Table 6.1. The highest-T_c results arise from structures with a cage-like arrangement of hydrogen atoms surrounding a central non-hydrogen element. The electronic states that originate from these cages are near the Fermi level, and are strongly coupled by cage vibrations. This provides the phonon-mediated pairing mechanism necessary for conventional superconductivity. Combined with a high average phonon frequency, owing to the light mass of the hydrogen atoms, this results in a high critical temperature (c.f the Allen-Dynes equation [11]). This can be seen directly by looking at the Eliashberg function, shown in Fig. 6.6, for two illustrative structures from Table 6.1. The enhanced high-frequency portion of the Eliashberg function for the high-T_c cage-like RbH$_{12}$ structure is apparent. In contrast, strong electron-phonon coupling
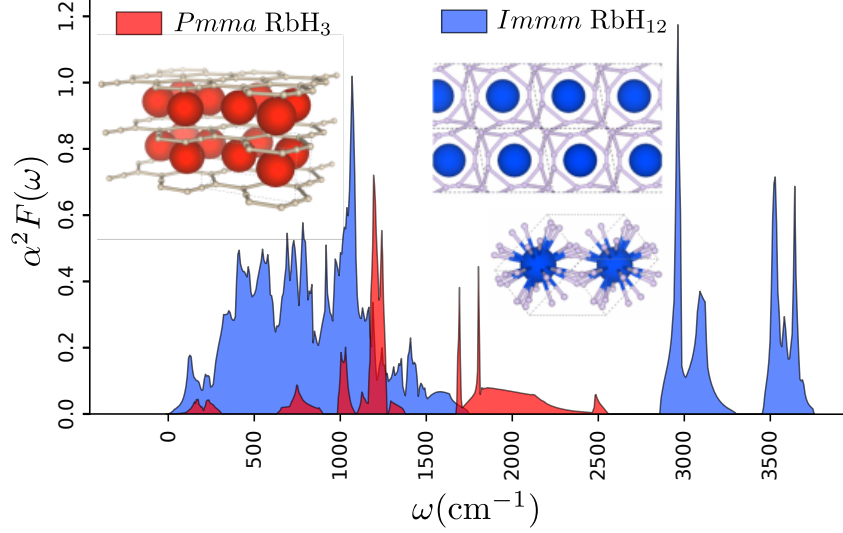
Figure 6.6: The Eliashberg function for $Immm$-RbH$_{12}$ and $Pmma$-RbH$_3$ found in our AIRSS searches (see Table 6.1). It is clear to see the enhanced high-frequency part of the Eliashberg function for Immm-RbH$_{12}$, arising from the hydrogen cage. In contrast, the Eliashberg function for the layered RbH$_3$ structure does not extend to such high frequencies. This effect can also be seen from the phonon linewidths, plotted along the phonon dispersion in Figs. 6.7 and 6.8.

is absent at high phonon frequencies for states near the Fermi level in the layered RbH$_3$ structure, leading to a negligible T_c. This can also be seen by examining the phonon linewidth broadening that arises due to electron-phonon coupling (see Figs. 6.7 and 6.8).

6.5.1   Limitations of the machine learning model

It is perhaps unsurprising that our machine learning model suggests low-hydrogen-content compositions, such as RbH$_3$, despite their resulting unfavorable structures for superconductivity, as it is trained on mostly cage-like structures. As a result, the model may implicitly assume that compositions it is given will behave as if they adopt cage-like arrangements, leading to an overestimation of T_c. Despite this, most of the structures found are high-T_c cage-like superconductors, of which $Immm$-RbH$_{12}$ is particularly interesting due to its location in Fig. 6.1.

In order to move beyond cage-like superconductors, it is necessary to generalize the model by either using a more diverse training set or by incorporating physical information that remains valid outside of the training set. Given the limited data in the literature, we explore the latter option in chapter 7.

Stoichiometry	Space group	Pressure (GPa)	T_c (K)
RbH$_{12}$	$C2/m$	50	108
RbH$_{12}$	$C2/m$	100	129
RbH$_{12}$	$C2/m$	150	133
RbH$_{12}$	$Cmcm$	100	82
RbH$_{12}$	$Immm$	50	115
RbH$_{12}$	$Immm$	100	119
RbH$_{12}$	$Immm$	150	126
CsH$_7$	$P1$	100	90
CsH$_7$	$I4mm$	100	34
CsH$_7$	$P4mm$	100	33
CsH$_7$	$I4/mmm$	100	10
CsH$_7$	Cm	100	5
CsH$_7$	$Cmc2_1$	100	89
RbH$_3$	$Pmma$	100	0
RbH$_3$	$Cmmm$	100	0

Table 6.1: Critical temperatures calculated using DFPT for promising hydride compositions - the structures listed here were found in this work using AIRSS and are available in an online repository [3]. The data in this Table is also shown in Fig. 6.1 for comparison with previous results in the literature.
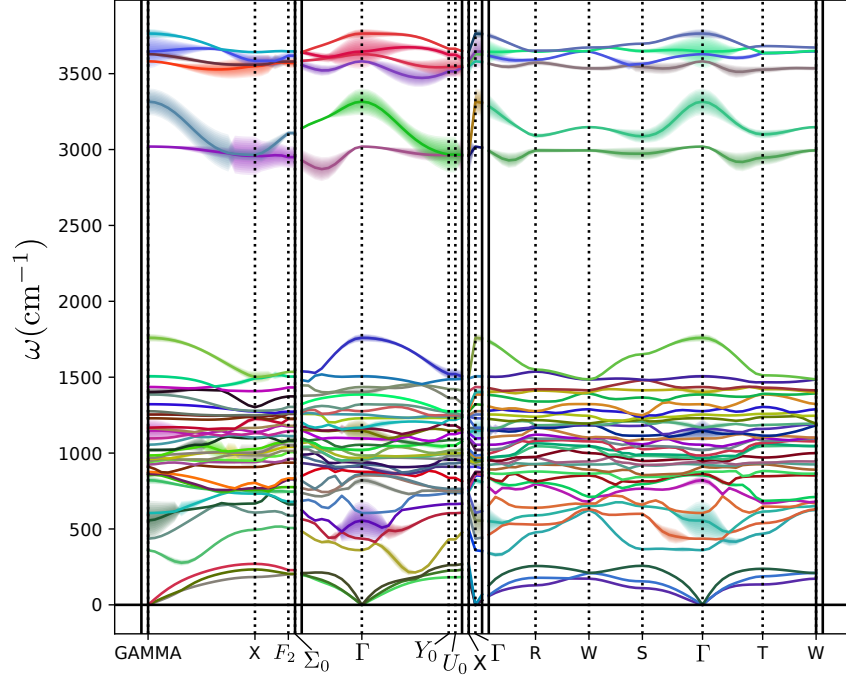
Figure 6.7: Phonon dispersion of the $Immm$ phase of RbH_{12} showing linewidths due to electron-phonon coupling, calculated using DFPT. Note that the high-frequency phonon modes due to vibrations of the hydrogen cage have large linewidths, resulting in a high critical temperature.
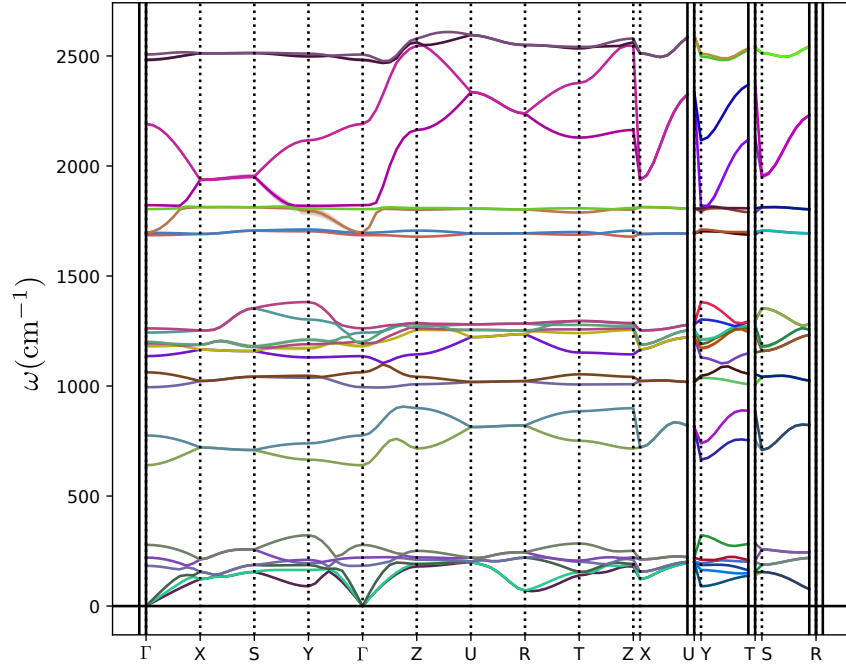
Figure 6.8: As Fig. 6.7, but for the $Pmma$ phase of RbH_3. Because of the layered structure, there are no high-frequency phonon modes which display strong electron-phonon coupling (in contrast to the case in Fig. 6.7).

6.6 Testing potential screening techniques for high-T_c candidates

In this work we also tested two potential methods for cheaply ranking potential superconductors. Good superconductivity in hydrides generally requires hydrogenic states close to the Fermi level, which (as exemplified by the findings of this work) often means favouring cage-like structures and avoiding structures with molecular-character H_2 units. It is therefore possible that the hydrogen-derived DOS normalised by the total DOS at the Fermi energy, $N_H(E_F)/N(E_F)$, may give some indication of whether a particular structure will exhibit high-T_c superconductivity. Here we also consider the hydrogen-derived electron-phonon coupling estimates (η_H) from Gaspari-Gyorffy theory [97] and test whether these two quantities could provide a method for ranking different structures (of the same stoichiometry and at the same pressure) before performing expensive electron-phonon calculations. We implemented Gaspari-Gyorffy theory within the ELK code [1]. The basics of this theory and its use here are explained in Appendix. A.9.

The calculated T_c values for the structures predicted and studied in this work allowed us to directly assess these potential screening methods. We observe that η_H correctly predicts the T_c ordering for the RbH_{12} structures at fixed pressure, as was the case for the LaH_{10} and YH_{10} systems from the previous chapter, on which preliminary tests were performed. The DOS ratio $N_H(E_F)/N(E_F)$ also demonstrates roughly the same general trends, but is cheaper to calculate since it can be obtained using a pseudopotential code. Unfortunately, looking instead at CsH_7, $N_H(E_F)/N(E_F)$ appears to be much less predictive and the performance of η_H is also mixed. We will revisit the use of these quantities for screening applications for a wider variety of systems in chapter 7.

6.7 Conclusions

Having identified the need to reduce the operational pressure of hydride superconductors, we searched for crystal structures which would exhibit superconductivity in novel regions of pressure-temperature space. We found that guiding structure searching techniques using a machine learning model allowed us to target the most promising regions. Specifically, we constructed models of critical temperature and operational pressure trained on the available theoretical and experimental results for binary hydride superconductors. Several novel systems were identified as promising superconductors closer to ambient conditions; here we focused on Cs and Rb hydrides, using AIRSS to identify stable stoichiometries and predict crystal struc-

tures. Other promising candidates included Ca, Sr, Ba, Ra, Ac, Th, La and Sc hydrides, most of which had already been theoretically studied to some extent [274, 258, 31, 118, 236, 151, 239, 165, 213, 164, 148, 70]. Critical temperatures of energetically-competitive candidate structures were then calculated from first principles using DFPT. A T_c of up to 115 K was calculated for RbH_{12} at 50 GPa, which represents a significant extension towards ambient-condition superconductivity from our dataset.

Chapter 7

High-throughput discovery of superconductors

7.1 Introduction

Having discussed how to calculate the structural and superconducting properties of hydrides in chapter 5 and investigated potential descriptors for screening candidate structures in chapter 6, we are well placed to carry out searches for superconductors on a larger scale. In order to apply these searches to as wide a range of materials as possible, we will need to refine the screening methods and DFPT calculations. In this chapter, we aim to develop a workflow that is capable of exploring the structure and superconductivity of binary hydrides composed from elements across the entire periodic table over a 10 to 500 GPa pressure range. Stable and metastable structures are screened using a model trained to predict T_c from inputs including electronic density of states (DOS) ratios and electron-phonon coupling estimates from Gaspari-Gyorffy theory [97] (see Appendix. A.9). High-throughput electron-phonon calculations are then performed for a large number of structures, selected based on the predictions of this model. This provides more T_c results which can be fed back into the model training set, allowing us to run the process iteratively. The results of this high-throughput model-training stage allow us to identify the most promising candidate systems at each pressure; more thorough structure searching is then performed for each of these systems and fully converged electron-phonon calculations are performed for the best predicted candidates. A large number of high-T_c superconductors are efficiently identified, including several not reported in previous work.

7.2 Optimized Electron-phonon DFPT

In this chapter we employ the method discussed in chapter 5 to calculate the super-conducting properties of candidate hydride structures. However, electron-phonon calculations are extremely expensive, reducing the breadth of materials that can be considered. One of the greatest assets of open-source code such as QUANTUM ESPRESSO is that it enables users to profile and optimise their specific use cases, rather than relying on developers to predict where bottlenecks may arise. At several stages of this study, we profiled the QUANTUM ESPRESSO code and implemented optimizations[1]. For example, in order to evaluate the double-delta smearing, terms of the form $\exp(-(\epsilon_1 - \epsilon_F)^2)\exp(-(\epsilon_2 - \epsilon_F)^2)$ need to be evaluated (for a calculation of $Fm\bar{3}m$-LaH$_{10}$, around 5 billion times). The exponentials were evaluated individually and via a lookup in a library of various smearing functions. By evaluating them directly (and combined into a single exponential), the whole calculation is sped up by around a factor of 4. We also found that the electron-phonon coupling calculation spent a significant amount of time in symmetrising the matrix of electron-phonon coefficients. Optimizing a single routine, called around one million times for the LaH$_{10}$ calculation, resulted in a 2 times speedup in the entire calculation. We also extended the electron-phonon calculation to parallelise over phonon modes, allowing better utilization of multi-node supercomputers. We emphasise that these modifications did not take a large amount of time to implement, and allowed us to study many more materials than would otherwise be possible. We strongly encourage other users of open-source code to profile their use cases, even if just to report bottlenecks to developers. Once we have the electron-phonon coupling constants from the expensive DFPT calculations, we obtain the superconducting critical temperature either by direct solution of the Eliashberg equations using the ELK code [1] or using the Allen-Dynes equation [12]. Both of these methods are essentially free (when compared to the preceding DFPT calculations) and we compare them in Sec. 7.6.6.

7.3 Automated electron-phonon calculations

An often-overlooked aspect of carrying out expensive calculations efficiently is the human cost. Setting up an electron-phonon calculation for a particular material is an involved process involving careful choice of parameters, accurate creation of input files, selection of quantities to calculate and management of the resulting data (which can be on the order of 100 GB/structure for electron-phonon calculations).

---

[1]The optimizations to QUANTUM ESPRESSO resulting from the work for this chapter have been submitted to the developers (see https://gitlab.com/miicck/q-e)

During the work involved in chapters 5 and 6, I developed a python library [126] to automate calculations using the QUANTUM ESPRESSO code; an overview of its structure is shown in Fig. 7.1. This library reduces the human input required to carry out an electron-phonon calculation from several hours of fiddling with input files and writing submission scripts to a few minutes of work at most, making the current high-throughput study possible.

7.4   Initial structure searching

The structure searching calculations for this chapter were performed using *ab initio* random structure searching (AIRSS) [217, 219] and the plane-wave pseudopotential code CASTEP [48]. The Perdew-Burke-Ernzerhof (PBE) generalised gradient approximation [215], CASTEP QC5 pseudopotentials, a 340 eV plane-wave cut-off and a $\mathbf{k}$-point spacing of $2\pi \times 0.07$ Å^{-1} were used. For the initial searches, sp-AIRSS [198] was utilised and structures with 8-48 symmetry operations were generated. This served the dual purpose of (1) reducing the computational cost of the searches and subsequent calculations during the model training phase and (2) allowing us to explore high-symmetry structures (which may be metastable or stabilised at nonzero temperatures). For each element, X, in the periodic table, a convex hull was produced at 10, 100, 200, 300 and 500 GPa in order to assess the stability of binary hydrides of the form X_nH_m. Structures on or near these static-lattice convex hulls were then selected for further investigation at the pressure of interest. In particular, in the training phase, stoichiometries within 80 meV/formula unit of the hull were selected (for each of these stoichiometries just the lowest energy structure was chosen). In order to discuss stability more clearly in this chapter, we will introduce two quantities: E_{stoic}, the distance of the given stoichiometry from the static-lattice convex hull, and E_{struc}, the distance of the given structure from the lowest energy structure of the same stoichiometry. Here, for example, we are selecting structures with $E_{stoic} \leq 80$ meV/formula unit and $E_{struc} = 0$.

7.5 T_c model and training phase

Our T_c model is a Gaussian Process Regression (GPR) model, initially trained using a set of 160 structures and corresponding T_c values from the literature (collected from Refs. [237, 120, 92, 7, 286, 124, 182, 275, 230, 288, 109, 82, 94, 294, 67, 241, 68, 78, 292, 170, 295, 296, 148, 96, 299, 174, 301, 172, 183, 284, 132, 83, 303, 159, 153, 239, 149, 160, 293, 169, 158, 5, 238, 106, 243, 127]). The inputs to our T_c model are the Gaspari-Gyorffy electron-phonon coupling estimates for hydrogen and element X
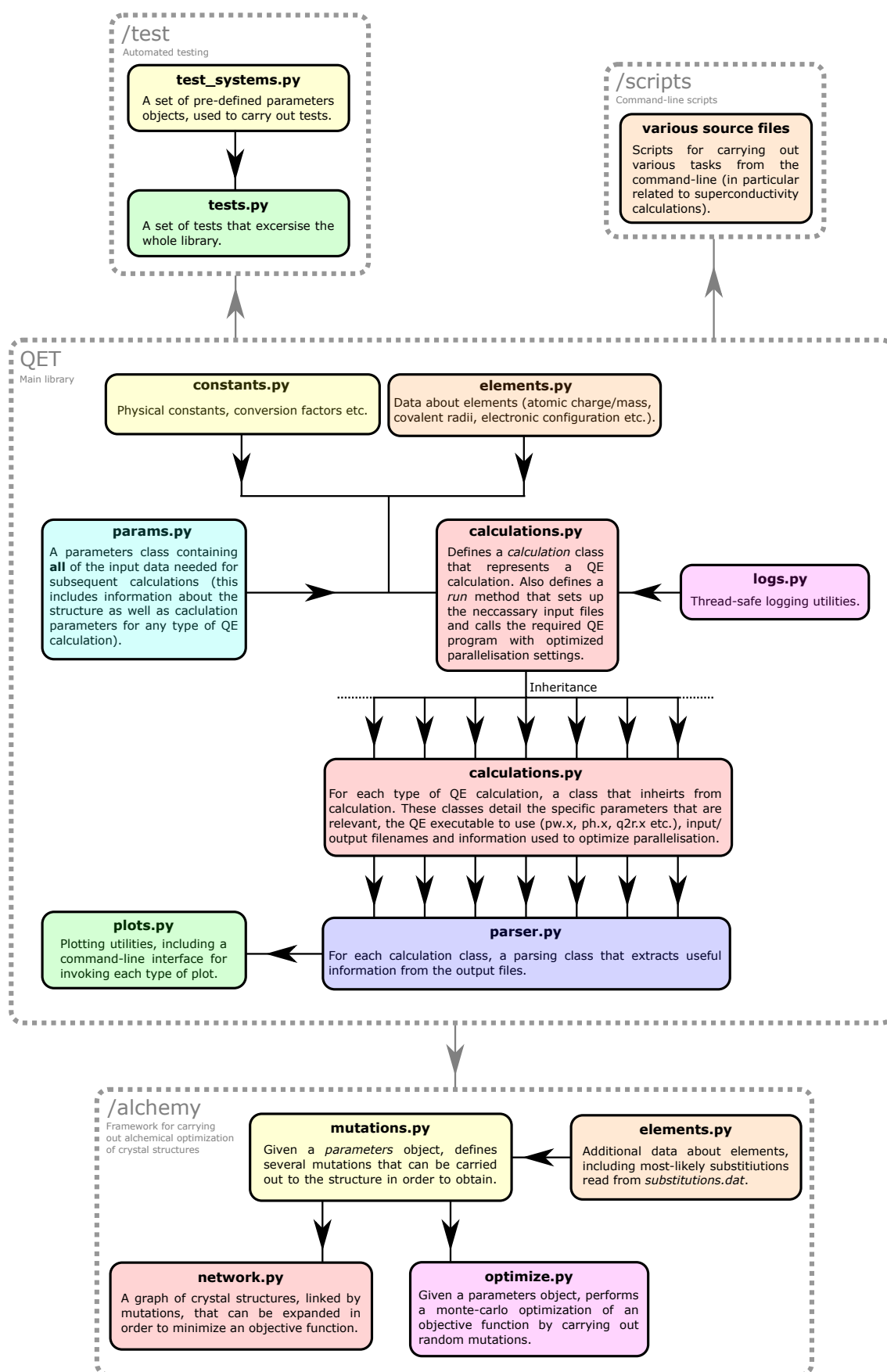
Figure 7.1: The QET python library.

(η_H and η_X), the mass of atom X in atomic units (M_X), the total DOS at the Fermi energy appropriately normalised by cell volume ($N(E_F)$), and the hydrogen DOS divided by the total DOS at the Fermi energy ($N_H(E_F)/N(E_F)$). Throughout this work, whether the structure originated from the literature or from our own searches, these inputs were computed using the modified version of ELK. For the literature data, where the structure was not given at the same pressure as T_c was reported, the structure was relaxed at the correct pressure using CASTEP [48]. Refs. [32, 69, 88, 305] and the data tables within were found to be helpful for identifying additional points to include in the original literature set.

The model was trained via optimisation of model parameters and hyperparameters in MATLAB [4] and was tested using nested cross-validation, with an inner loop used to optimise the parameters and an outer loop used to monitor the fit of the resulting model to unseen data. The correlation was evaluated in this way, repeated over several different random splittings of the data each time.

The overall process used in the training phase was iterative; structures at a given pressure were selected for further study based on the predictions of the model, with T_c values for the best predicted structures calculated explicitly using DFPT and fed back into the model's training set for use in the next iteration. The model was then retrained and predictions were made for a set of search structures at the next pressure until all pressures had been considered. To make the large number of electron-phonon calculations feasible, a relatively sparse $\mathbf{q}$-point sampling was used in the training stage, chosen to reproduce the known result for $Fm\bar{3}m$-LaH$_{10}$ [64, 243]. This corresponds to a $\mathbf{q}$-point spacing of $2\pi\times0.15\text{Å}^{-1}$ (a $2 \times 2 \times 2$ grid for $Fm\bar{3}m$-LaH$_{10}$ at 200 GPa). The Allen-Dynes equation was used to estimate T_c throughout the training phase.

Our method is summarised in Fig. 7.2. Fig. 7.3 shows the number of structures considered in total at each stage of the training process - given the relatively large cost of electron-phonon calculations (even in high-throughput operation), this figure highlights the importance of the stability filtering and model-based screening steps in our workflow. In total, 119 new DFPT data points were added to the training set in this work. Predictions for all previously-considered search structures were recalculated using the final model to ensure nothing of interest had been missed in earlier iterations. Most of the electron-phonon calculations performed in the training phase were for structures with a high predicted T_c according to our model, however, occasionally these calculations were performed for structures with mid-range or low T_c predictions in order to improve the behaviour of the model. The correlation of the predicted T_c values with the calculated T_c values across the training set decreased slightly on addition of more data to the training set. This is not surprising for two main reasons: (1) the original (literature-based) training set contained 160 entries

for binary hydrides of 45 different elements largely focused around 100-350 GPa, whereas the training set for the final model contains 279 entries for binary hydrides of 57 different elements from 10 to 500 GPa, and (2) as mentioned previously, the results we add to the training set are computed in a high-throughput manner and therefore will be somewhat under-converged compared to typical values found in the literature.

During the training of our model, we were able to efficiently rediscover a number of binary hydrides (with relatively high T_c values) which had been reported previously, including $Im\bar{3}m$-H_3S [68, 63, 78], $Im\bar{3}m$-LaH_6 [214], $I4/mmm$-AcH_{12} [237], $Im\bar{3}m$-SeH_3 [292], $R\bar{3}m$-SrH_6 [32], $R\bar{3}m$-LiH_6 [282, 122], $Fm\bar{3}m$-LaH_{10} [167, 214, 243, 64, 250, 148], $Fm\bar{3}m$-YH_{10} [214, 243], $Im\bar{3}m$-ScH_6 [284, 6, 214], $P6_3/mmc$-ThH_9 [239], $R\bar{3}m$-SrH_{10} [259], $Pm\bar{3}m$-SiH_3 [132], $C2/m$-LaH_7 [148], $Im\bar{3}m$-CaH_6 [275], $Im\bar{3}m$-MgH_6 [85], $Fm\bar{3}m$-ThH_{10} [153, 239], KH_6 [296], LaH_8 [167, 148], BaH_{12} [238], LaH_5 [148], AcH_{10} [237], LiH_8 [282], LaH_{11} [148], MgH_{12} [179], YH_9 [214, 146] and ScH_{12} [284]. Although found previously, only half of these 26 stoichiometries were in our original literature data set, highlighting the capabilities of our method. We also identified a number of other systems with the potential to exhibit high-T_c superconductivity. The most promising systems overall are studied further in Sec 7.6.

7.6 Focused searches and final results

From the T_c results obtained during training, the most exciting candidate systems could be identified and studied in more depth. More focused structure searches were performed for hydrides of Na, Ca, La, Ac, and K at 100 GPa, hydrides of La, Ac, S, Mg, and Na at 200 GPa, hydrides of Li, Sr, K, Mg, Na, and Sc at 300 GPa, and hydrides of Li, Sr, Mg, Na, Yb, Y, and Ca at 500 GPa. For these calculations, the earlier symmetry constraints were relaxed, but all other parameters remained the same as in the initial searches. No particularly high T_c results were found at pressures as low as 10 GPa at the training stage (a large proportion of the structures were not even metallic), so no additional searches were performed at this pressure.

On completion of the focused searches, we again employed our modified version of ELK to calculate η_H, η_X, $N(E_F)$, and $N_H(E_F)/N(E_F)$ for the stable and metastable structures found. In this phase, stoichiometries with $E_{stoic} \leq 25$ meV/formula unit were selected. For the stoichiometries on the hull ($E_{stoic} = 0$), 2-5 of the most stable structures within were chosen. For the selected off-hull stoichiometries, only the lowest energy structure ($E_{struc} = 0$) was chosen. The inputs were then fed into the final T_c model trained in Section 7.5 and fully converged electron-phonon calcula-
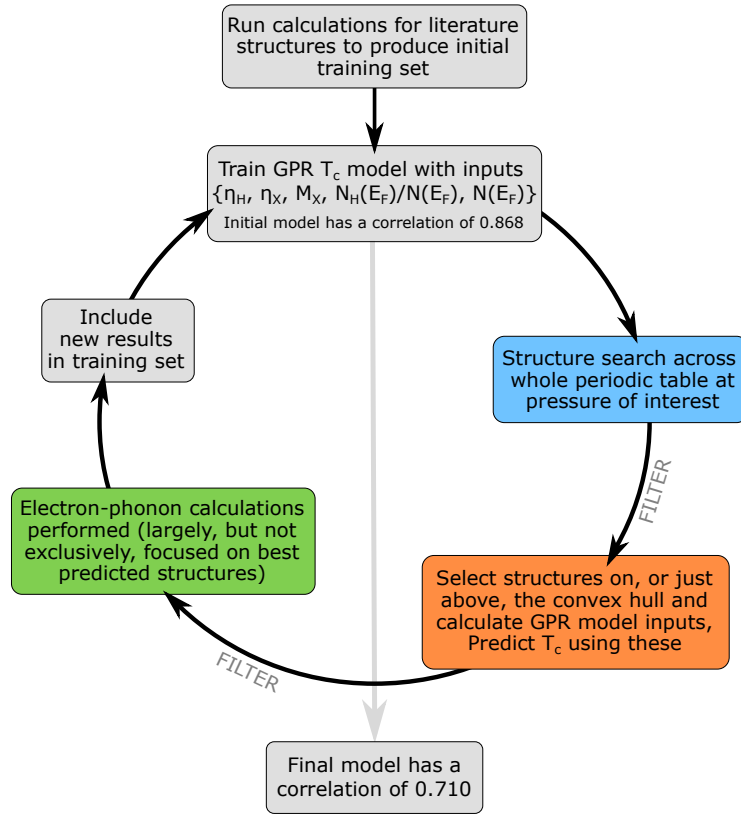
Figure 7.2: A flowchart summarising our methodology.
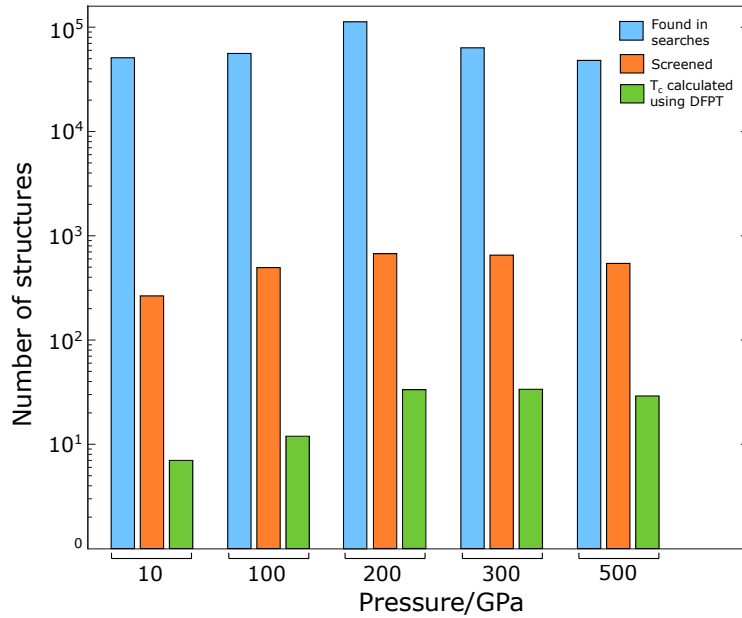
Figure 7.3: A summary of the number of structures studied at each stage of the training process. Note the logarithmic scale.

tions were performed for the structures with the highest predicted T_c values at each pressure (as well as the most promising candidates identified during training). These results are shown in Fig. 7.4. Promising structures that remained dynamically stable after convergence of the **q**-point grids are shown in Table 7.1. We find near room-temperature superconductors at every pressure considered, which we elaborate on in the following sections. Unless otherwise stated, all critical temperatures reported in the following sections are from direct solution of the Eliashberg equations.
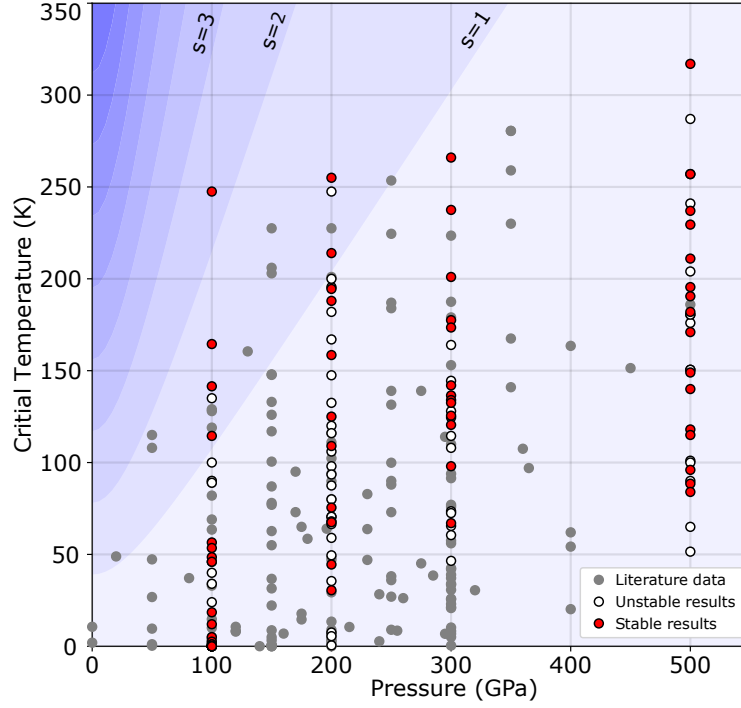
Figure 7.4: Critical temperatures obtained from converged DFPT calculations for the most promising candidates according to our screening process (calculated using the Allen-Dynes equation to facilitate comparison with the literature). The background is shaded according to the figure of merit S, introduced in Ref. [218].
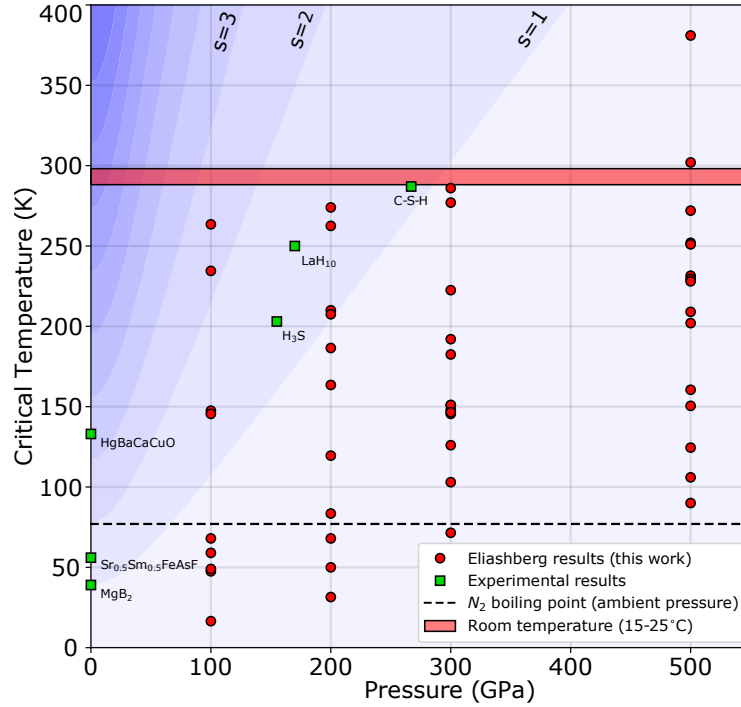
Figure 7.5: As Fig. 7.4, but showing the Eliashberg results from Table 7.1 alongside experimental results for specific superconductors (in order of increasing T_c: [202], [280], [234], [63], [64] and [248]).

Stoic.	Symmetry	P(GPa)	T_c (AD, K)	T_c (EL, K)	λ	E_{stoic}	E_{struc}
NaH_6	$Pm\bar{3}m$	100	228-267	248-279	2.54	28	0
CaH_6	$Im\bar{3}m$	100	150-179	216-253	5.81	9	137
Na_2H_{11}	$Cmmm$	100	127-156	134-161	1.28	0	0
KH_{10}	$C2/m$	100	105-124	134-157	2.45	0	0
AcH_5	$P\bar{1}$	100	48-65	49-69	0.91	0	0
NaH_{16}	$Fmm2$	100	47-60	61-75	1.10	0	0
NaH_{24}	$R\bar{3}$	100	40-57	40-55	0.82	0	0
LaH_5	$P\bar{1}$	100	37-55	40-58	0.83	0	0
AcH_{11}	$C2/m$	100	14-23	13-20	0.71	0	0
NaH_6	$Pm\bar{3}m$	200	235-275	260-288	2.06	39	16
AcH_{12}	$P6_3mc$	200	197-231	245-280	3.92	11	0
MgH_{13}	$Fm\bar{3}m$	200	179-210	196-224	1.98	17	635
SH_3	$Im\bar{3}m$	200	173-203	196-219	1.77	0	0
NaH_8	$I4/mmm$	200	146-171	152-175	1.63	26	124
AcH_6	$Fmmm$	200	110-140	169-204	2.01	0	14
LaH_7	$C2/m$	200	98-120	105-134	1.23	3	0
MgH_4	$I4/mmm$	200	63-88	73-94	0.98	0	101
SH_7	$Fmmm$	200	57-78	58-78	0.91	29	0
AcH_4	$Fm\bar{3}m$	200	35-54	42-58	0.99	19	0
Mg_2H_5	$R\bar{3}m$	200	22-39	24-39	0.74	139	21
MgH_6	$Im\bar{3}m$	300	248-284	271-301	2.28	19	437
YH_9	$F\bar{4}3m$	300	220-255	261-293	2.58	2	0
ScH_8	$Immm$	300	185-217	212-233	2.06	3	0
LiH_2	$P6/mmm$	300	162-193	177-207	1.45	40	75
Na_2H_{14}	$C2/m$	300	157-190	167-198	1.48	3	0
ScH_{12}	$P\bar{1}$	300	127-157	137-165	1.28	0	103
LiH_6	$R\bar{3}m$	300	121-152	130-161	1.30	0	0
ScH_6	$Im\bar{3}m$	300	118-150	135-161	1.26	0	0
NaH_5	$P4/mmm$	300	121-144	138-164	1.92	1	0
LiH_6	$C2/m$	300	109-142	130-163	1.16	0	14
Li_2H_6	$Cmcm$	300	104-137	112-140	1.06	1	0
ScH_{14}	$P\bar{1}$	300	87-109	91-115	1.20	6	0
MgH_4	$I4/mmm$	300	53-81	59-84	0.76	0	0
MgH_{12}	$Pm\bar{3}$	500	294-340	360-402	2.65	0	259
MgH_{13}	$P3m1$	500	239-275	257-287	2.21	12	0
SrH_{10}	$Fm\bar{3}m$	500	239-275	285-319	2.22	8	†
Na_2H_{18}	$P6_3/mmc$	500	218-256	235-269	1.67	0	0
MgH_{10}	$C2/m$	500	209-250	232-270	1.63	9	0
SrH_{24}	$R\bar{3}$	500	195-227	218-245	1.88	9	0
YH_{18}	$P\bar{1}$	500	179-212	213-246	1.99	25	0
YH_{20}	$P\bar{1}$	500	176-205	212-244	2.21	39	0
SrH_{10}	$R\bar{3}m$	500	165-199	190-228	1.31	8	0
CaH_{10}	$R\bar{3}m$	500	155-187	184-220	1.51	3	0
Na_2H_{11}	$Cmmm$	500	132-166	141-180	1.12	0	0
CaH_{15}	$P\bar{6}2m$	500	120-160	134-167	1.01	0	0
SrH_{15}	$P\bar{6}2m$	500	100-136	110-139	0.93	0	0
MgH_8	$C2/m$	500	82-110	91-121	0.96	0	0
Na_2H_{11}	$I4/mmm$	500	72-105	76-104	0.80	0	297

Table 7.1: Allen-Dynes and Eliashberg T_c values for dynamically stable superconductors found in this work, along with calculated λ and stability measures (in meV/formula unit).

7.6.1 100GPa

We start our discussion of the results with the structures at 100 GPa - a relatively low pressure in the field of hydride superconductivity. Of particular note is a $Pm\bar{3}m$ structure of NaH_6, a stoichiometry which we find to be metastable in agreement with Ref. [22] (where a $P\bar{1}$ structure is found to be more favorable until 150 GPa). The structure consists of a cubic lattice of H octahedra with Na at the body-centred positions (see Fig. 7.6). While experimental synthesis of sodium polyhydrides has been demonstrated [253], superconductivity in the system seems under-studied given its promise here with a calculated T_c of 248-279 K. This places the structure at a crucial position in pressure-T_c space, strongly influencing the apparent low-pressure trend of maximum T_c (see Figs. 7.4 and 7.5) and hinting at the exciting possibility of other low-pressure high-T_c structures.

Also of interest at 100 GPa is an $Im\bar{3}m$ structure of CaH_6 (shown in Fig. 7.7). As was the case with NaH_6, this structure is found to be metastable in agreement with previous calculations [275]. However, despite strong electron-phonon coupling (leading to an extremely high λ of 5.81), its critical temperature (216-253 K) is found to be slightly lower than that of $Pm\bar{3}m$-NaH_6 (248-279 K), due to a lower average phonon frequency. This result is comparable to the critical temperature of 220-235 K calculated for this structure at 150 GPa in Ref. [275], also via solution of the Eliashberg equations. The Allen-Dynes equation is found to drastically underestimate the critical temperature of this system as a result of electrons coupling to soft phonon modes (see Fig. 7.9). This is unsurprising as the equation is known to be sensitive to small changes in $\alpha^2F(\omega)$ at low frequencies (in particular, the functional derivative $\delta T_c(\text{AD})/\delta\alpha^2F(\omega)$ diverges as $\omega \to 0$ [29], in this case towards $-\infty$).

Fig. 7.8 also demonstrates that whilst the next best structures, $Cmmm$-Na_2H_{11} and $C2/m$-KH_{10}, have similar average phonon frequencies to $Pm\bar{3}m$-NaH_6, they do not exhibit such high coupling strengths. This leads to lower T_cs of 134-161 K and 134-157 K, respectively. Superconductivity in the KH_{10} stoichiometry has been studied previously [238]; it was found to be on the convex hull at 150 GPa (and off-hull at 50 GPa) and a T_c of 148 K was estimated for an $Immm$ structure based on a regression model. Ref. [116] found KH_{10} to be above the convex hull at 100 GPa and instead found metastable metallic structures of other stoichiometries, but superconductivity was not directly investigated. The calculated critical temperature of KH_{10} at 100 GPa exceeds a previously calculated T_c for KH_6 at higher pressure (73 K at 166 GPa [256]).

(a)   (b)
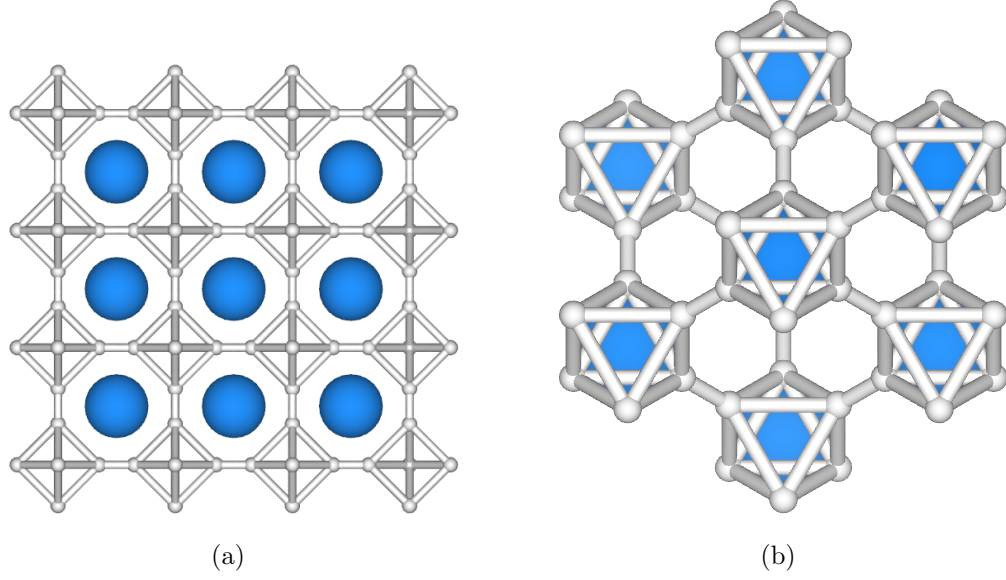
Figure 7.6: The 100 GPa structure of $Pm\bar{3}m$ NaH$_6$ viewed along the (a) [100] and (b) [111] directions of the standardized cell.

(a)   (b)
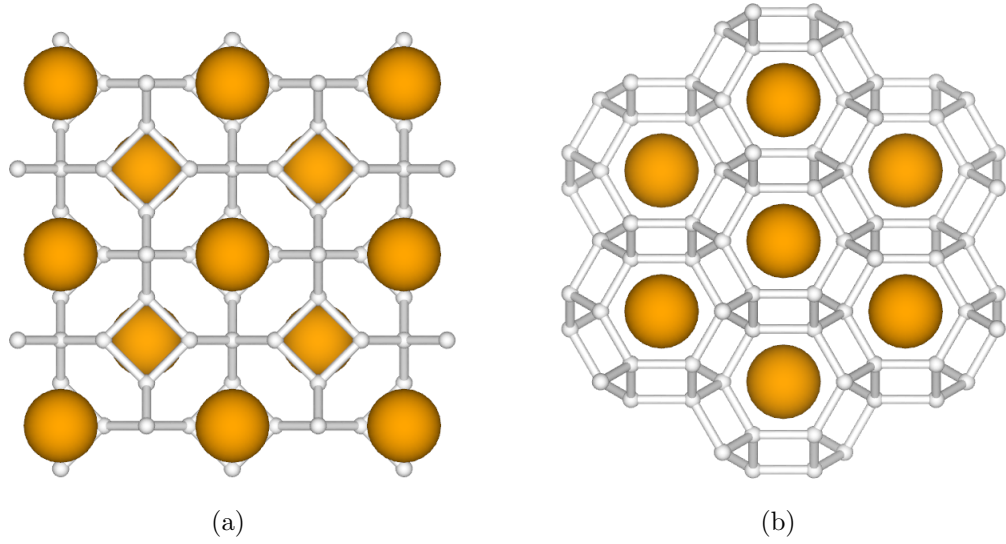
Figure 7.7: The 100 GPa structure of $Im\bar{3}m$ CaH$_6$ viewed along the (a) [100] and (b) [111] directions of the standardised cell. The $Im\bar{3}m$ structure of MgH$_6$ investigated at 300 GPa can be obtained by substituting the Ca atoms with Mg atoms.
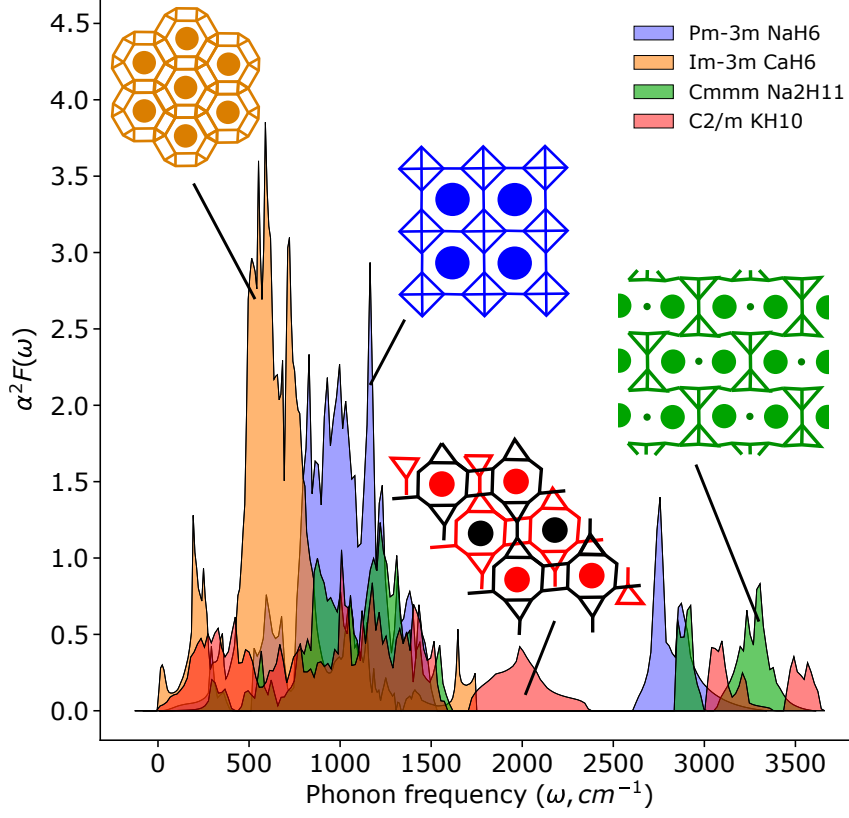
Figure 7.8: The highest-T_c structures and corresponding Eliashberg functions at 100 GPa.
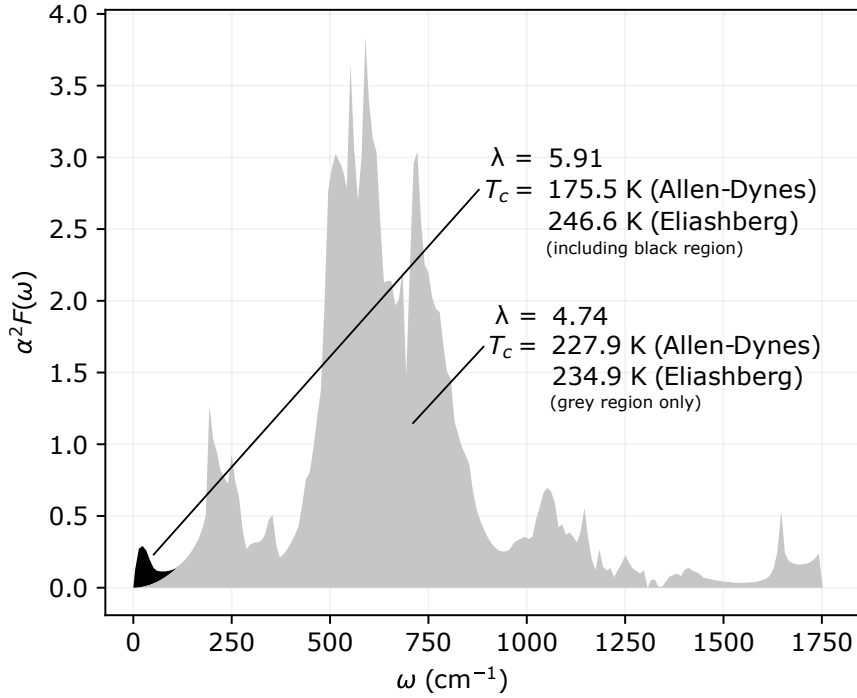
Figure 7.9: The Eliashberg function of $Im\bar{3}m$ CaH$_6$ at 100GPa, demonstrating the reduction of T_c from Allen-Dynes relative to that of solving the Eliashberg equations. This reduction is due to the presence of electrons coupling to soft phonon modes.
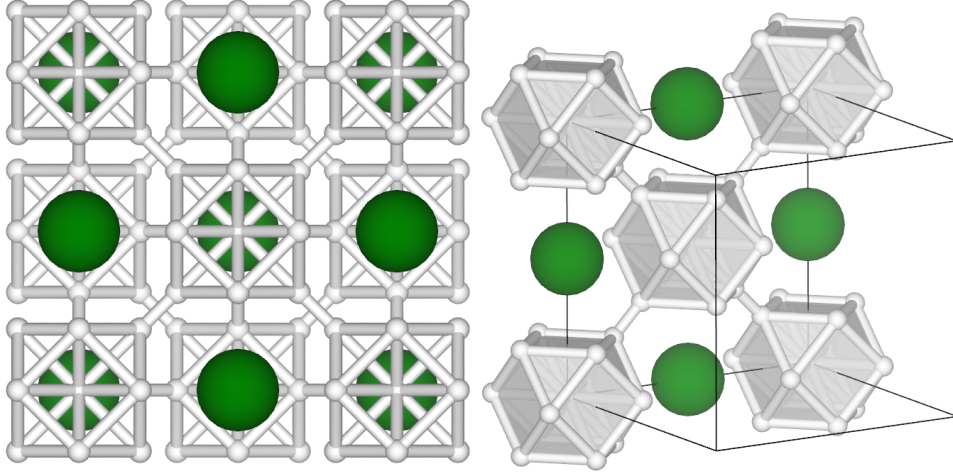
Figure 7.10: The 200 GPa $Fm\bar{3}m$ structure of MgH_{13}. The structure consists of Mg atoms and axis-aligned cuboctahedra of hydrogen in a checkerboard pattern. Each cuboctahedra has an additional hydrogen atom at its centre to make up the necessary 13.

7.6.2   200GPa

At 200 GPa, the $Pm\bar{3}m$ structure of NaH_6 remains the highest T_c structure found, with T_c rising slightly from its 100 GPa value to 260-288 K. However, as pressure increases we find this stoichiometry to be less stable with respect to decomposition. Similarly to Ref. [237], we find several actinium hydrides to be high-temperature superconductors at this pressure, most notably a metastable $P6_3mc$ structure of AcH_{12} (see Fig. 7.11) with a critical temperature of 245-280 K.

200 GPa also marks the appearance of Mg hydrides, which become increasingly prevalent with pressure. Of particular note is a cubic structure of MgH_{13} with the space group $Fm\bar{3}m$ (see Fig. 7.10), which possesses a slightly higher calculated T_c than that of the experimentally-verified $Im\bar{3}m$-H_3S [63] at the same pressure. In agreement with Ref. [179], we find MgH_{13} to lie off of the convex hull at 200 GPa. They report significantly lower critical temperatures for on-hull structures.

A notable absence from the 200 GPa results is LaH_{10}. Several structures of LaH_{10} (including the experimentally-verified $Fm\bar{3}m$ structure [64]) were found in our structure searches and flagged as good candidates by our T_c model, but were dynamically unstable at the harmonic level. This has been noted previously [99, 166, 243]. We elaborate on the consequences of dynamic instabilities on our results and potential ways to deal with them in section 7.6.5.
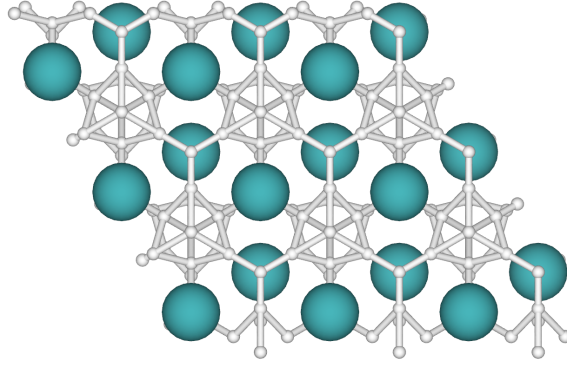
Figure 7.11: The 200 GPa $P6_3mc$ structure of AcH_{12} viewed along the [001] direction of the standardized cell.

7.6.3   300GPa

At 300 GPa an $Im\bar{3}m$ structure of MgH_6 exhibits the highest critical temperature calculated with a T_c of 271-301 K (in agreement with previous calculations [85]). This structure can be obtained by substituting Mg for Ca in the CaH_6 structure investigated at 100 GPa (see Fig. 7.7). By instead substituting only half of the Ca atoms at 200 GPa, it has been reported that one obtains an even higher critical temperature ternary superconductor [254]. Hybridizing compatible binary crystal structures in this way could provide an efficient method to design future high-temperature ternary superconductors.

With a critical temperature of 261-293 K, an $F\bar{4}3m$ structure of YH_9 (shown in Fig. 7.12(a)) is the next highest temperature superconductor found at 300 GPa. This stoichiometry (although with $P6_3/mmc$ symmetry) has been synthesised experimentally and was found to exhibit a critical temperature of 243 K at 201 GPa [146]. Y-H systems have been extensively studied theoretically [111, 214, 243] with critical temperatures in excess of 200 K calculated over large pressure ranges.

We find that an $Immm$ structure of ScH_8 with a similar motif to $F\bar{4}3m$-YH_9 (see Fig. 7.12(b)) is also a high-temperature superconductor at this pressure with a critical temperature of 212-233 K. However, this is significantly higher than the value of $\sim$ 115 K (Allen-Dynes) obtained previously in Ref. [226] using a $16 \times 16 \times 16$ k-point grid and norm-conserving pseudopotentials with 3 valence electrons for Sc. In contrast, we use a $36 \times 36 \times 36$ k-point grid and ultrasoft pseudopotentials with 11 valence electrons for Sc; a more substantial investigation into pseudopotentials is needed to fully solve this discrepancy. Whilst the authors did not calculate T_c for ScH_8, critical temperatures of 213 K and 233 K were obtained at 300 GPa for ScH_7 and ScH_9, respectively, in Ref. [284] (remarkably close to our range for ScH_8).

The next structure of note is a metastable $P6/mmm$ structure of LiH_2, which is interesting both because of its relatively low hydrogen content and because its
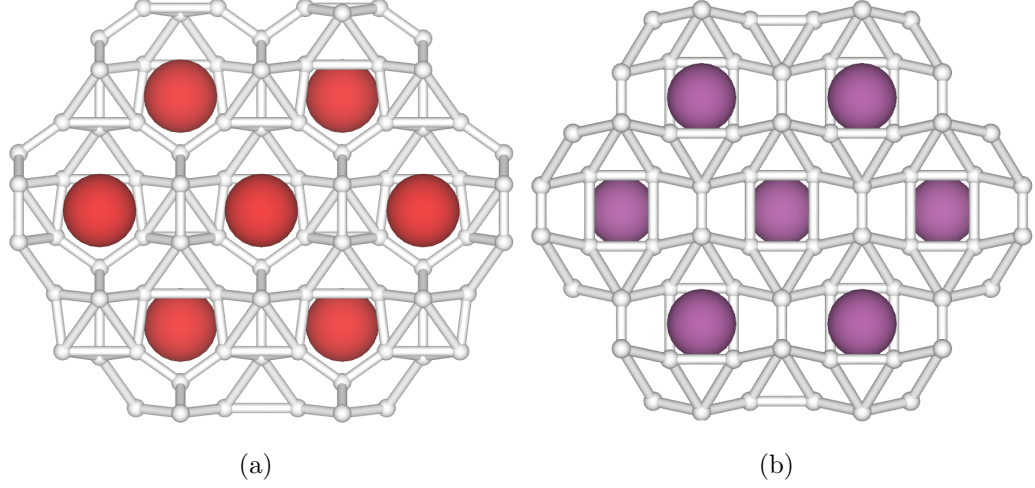
Figure 7.12: (a) The 300 GPa $F\bar{4}3m$ structure of YH$_9$ viewed along the [110] direction of the standardized cell. (b) The 300 GPa $Immm$ structure of ScH$_8$ viewed along the [100] direction of the standardized cell.
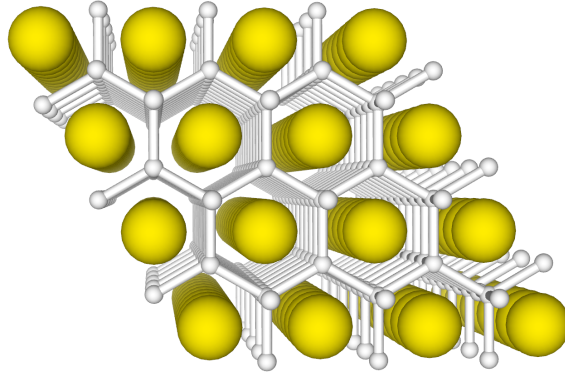
Figure 7.13: The 300 GPa $P6/mmm$ structure of LiH$_2$ viewed with perspective along the [001] direction of the standardized cell.

structure is analogous to the well-known ambient-pressure superconductor, MgB$_2$ [202] (see Fig. 7.13). Superconductivity in Li-H systems has been investigated previously at lower pressures [282], where it was found that the LiH$_2$ stoichiometry did not exhibit superconductivity at 150 GPa. Whilst the LiH$_2$ stoichiometry lies on the convex hull at 300 GPa, in agreement with the previous calculations of Ref. [306], the $P6/mmm$ structure is not the most stable LiH$_2$ structure and is instead found to be metastable at static-lattice level. It appears that, in some systems, structures appearing slightly above the convex hull may have enhanced superconducting properties.
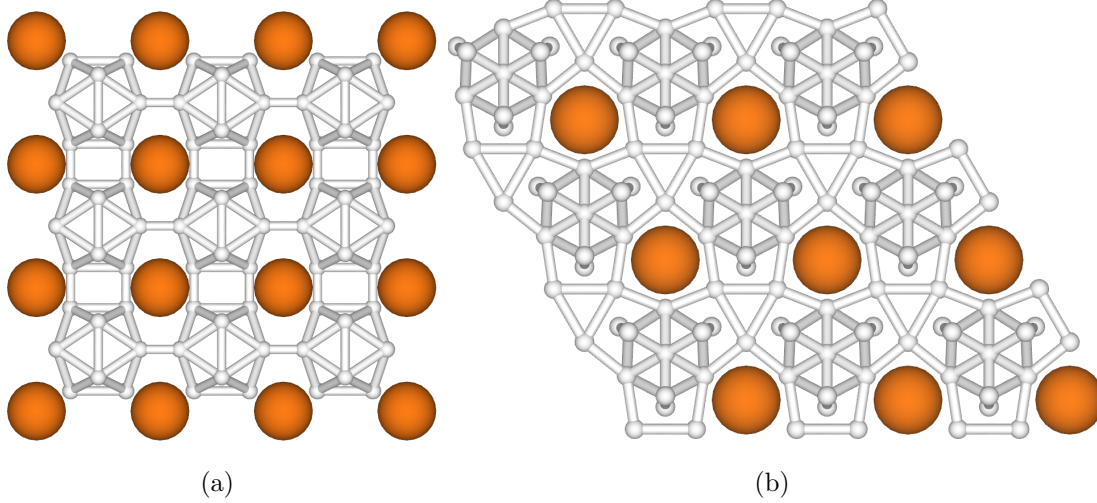
Figure 7.14: (a) The 500 GPa $Pm\bar{3}$ structure of MgH_{12} viewed along the [010] direction of the standardized cell. (b) The 500 GPa $Pm31$ structure of MgH_{13} viewed along the [001] direction of the standardized cell.

7.6.4 500GPa

As we increase pressure further, hydrides with higher hydrogen content can be metallised. In particular, here, we see the appearance of several MgH_n structures with $n > 10$. The highest critical temperature of these belongs to MgH_{12}, which we find to be on the convex hull at this pressure, where an $Pm\bar{3}$ structure (see Fig. 7.14(a)) exhibits hot superconductivity with a critical temperature of 360-402 K (87-129°C). $Pm31$-MgH_{13} (see Fig. 7.14(b)) exhibits a T_c of 257-287 K. Electron-phonon coupling is boosted substantially in this system with increasing pressure; the MgH_{12} stoichiometry has been previously investigated at lower pressures [179], where it was also found to lie on the convex hull, but has a much reduced T_c of 47-60 K at 140 GPa.

7.6.5 Comments - dynamically unstable structures

During a DFT geometry optimization, two properties of the crystal typically remain fixed; the crystal symmetry and the number of atoms in the simulation cell. As a result, phonon modes which either break crystal symmetry, or exist on length scales larger than the simulation cell are implicitly prohibited, even in the case where they might lower the energy of the system. In this case, the resulting crystal structure is artificially stabilized and lies on a saddle point of the potential energy surface, rather than a minimum. Locating such saddle-point structures is the goal of sp-AIRSS [198] (we discuss some related methods in chapter 8), as they could be stabilized by anharmonic or thermal effects, as is the case in $Fm\bar{3}m$-LaH10 at 200 GPa [79,

243, 64], and calculations on the saddle-point structures are cheaper to perform thanks to increased symmetry (or reduced cell size). However, the unstable phonon modes complicate electron-phonon calculations. In principle, one could allow the appropriate breaking of crystal symmetry and/or increase in simulation cell size, and follow the unstable phonon mode to a locally-stable structure (where it would be possible to apply a harmonic theory of electron-phonon coupling). However, it is not clear that this is the sensible thing to do given the (often drastic) increase in the computational cost and neglect of anharmonic effects, which could be important both for the stability and superconductivity of the unperturbed system. Indeed, the competition (or, in more specific cases, co-operation [273, 128]) between structural instability and strong electron-phonon coupling has been the subject of extensive discussion [13, 222, 260]. To decide the correct treatment, one could explicitly map the potential energy as a function of amplitude for the relevant modes. If the resulting potential permits stabilization of the high-symmetry structure (for example a shallow double-well potential) one can extract meaningful contributions to electron-phonon coupling due to exited states within this potential [125, 193], or by constructing a renormalized harmonic theory [81, 77]. Conversely, if the prospects for stabilization are poor, then mode-following to a more stable structure seems sensible. All of these approaches are currently too involved to permit application in a high-throughput setting and so we do not pursue them in the current work. However, for many hydrides, high-frequency phonon modes contribute most to T_c. Therefore, one can roughly establish the promise of a saddle-point superconductor by simply neglecting unstable modes in the calculation of the Eliashberg function [273, 243]. The results of this procedure are plotted in Fig. 7.4 as empty circles, where it can be seen that including the unstable results does not qualitatively change the overall picture.

7.6.6 Comments - Allen-Dynes equation

Having a large number of superconductors for which the Eliashberg equations have been solved directly (table 7.1) provides a unique opportunity to test the Allen-Dynes equation. This comparison is made in Fig. 7.15 (a), where it is clear to see that, whilst the Allen-Dynes result correlates well with the Eliashberg result, it systematically underestimates its value (at least for the binary hydride superconductors studied in this chapter). We fit a modified version of the Allen-Dynes equation of the form

$$T_c = T_c^{(AD)}(a + b\lambda) \tag{7.1}$$

to the data given in table 7.1, which gives $a = 1.0083$ and $b = 0.0654$. As we can see in Fig. 7.15 (b) This removes the systematic underestimation, and slightly reduces

the variance of the T_c predictions. However, given access to $\alpha^2 F(\omega)$, we recommend that the Eliashberg equations be solved directly, to remove the need for approximate T_c predictors entirely.

7.7   Comments - Structural motifs

In this work, the number of promising structures identified was a direct result of the ability to screen an even larger number of candidates (see Fig. 7.2). In order to do this, we required robust and quickly-calculable screening criteria. In the present work, we used the hydrogen/non-hydrogen contributions to the electronic density of states, proxy electron-phonon coupling constants from Gaspari-Gyorffy theory and measures of structural stability (see Sec. 7.5). None of this information is explicitly derived from the crystal structure - it all requires treatment of the structure with some form of quantum-chemistry method. A calculable, and preferably interpretable, quantity derived purely from the structure would be invaluable for screening, especially as it would likely be extremely cheap to calculate given the structure.

Many of the most attractive hydride superconductors in the literature adopt a clathrate "cage-like" structure. As a result, we investigated the construction of geometrical quantities to describe how cage-like a given structure was, in the hope that this would overcome issues from chapter 6, where it seemed that the machine learning algorithm was implicitly assuming a cage-like structure for every stoichiometry. However, the regression model in the present work paid little attention to these inputs. By looking at the structures responsible for the most promising results from Tab. 7.1, as shown in Figs. 7.16 and 7.17, it is clear to see why. Whilst cage-like structures are present, many are not cage-like. Indeed, some of the most promising materials in our results (e.g $Pm\bar{3}m$-NaH_6 at 100 GPa) are better described as containing supermollecular hydrogen clusters. For example, the 100 GPa NaH_6 structure consists of octohedral clusters of hydrogen at the body-centred sites of a cubic Na lattice. The tight packing of hydrogen into these clusters provides both the high-frequency phonons and nearby high-density of electrons necessary for strong electron-phonon coupling and resulting high-temperature superconductivity (as can be seen in the Eliashberg function for NaH6 in Fig. 7.8).

As well as exhibiting strong electron-phonon coupling (ideally to high-frequency phonons), electrons must also be sufficiently delocalised in order to allow bulk conductivity. These concepts form the basis of recent attempts to understand hydride superconductivity via the topology of the electron localization function [28]. It is possible to achieve favorable electronic topology from a wide variety of structures,
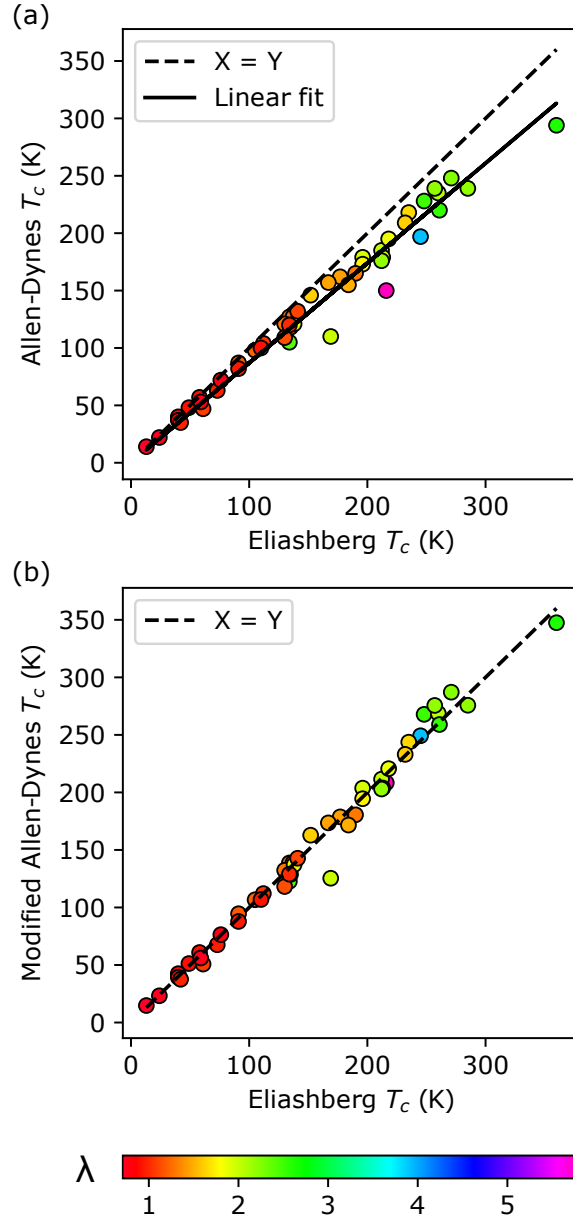
Figure 7.15: (a) Allen-Dynes critical temperatures, plotted against critical temperatures from solution of the Eliashberg equations (data from table 7.1). (b) The same as (a), but using the modified Allen-Dynes equation (Eq. 7.1).

as can be seen in Figs. 7.16 and 7.17, making the construction of geometrical measures of suitability difficult (explaining why our initial attempts failed). Therefore, it seems that the most promising approach for studies in the near future is to try to correlate topological, or geometrical, properties of the *electronic* structure (rather than the crystal structure) with the results of a full Eliashberg treatment. Such considerations may lead to a more broadly-applicable model of hydride superconductors, capable of explaining why each of the wide variety of structures found in this work are suitable. Such a model, when combined with existing physically-motivated screening factors (such as those used in the present work), would allow further extrapolation beyond the available literature data when evaluating the potential of new structures - potentially providing a route to push these materials towards ambient conditions.

7.8 Conclusions

In this chapter, we demonstrate a high-throughput method to efficiently discover high-T_c binary hydrides. We construct a T_c model based on physically-motivated descriptors, trained initially on superconductivity data from the literature and iteratively updated using the results of our own DFPT calculations. Following extensive structure searching, a two-step screening process (based on stability criteria and the predictions of our model) allows us to identify energetically-competitive high-T_c candidates from the large volume of search data. The best candidates include hydrides of sodium, calcium, actinium, lanthanum, magnesium, yttrium, scandium, lithium and strontium.

We have performed a total of 240 T_c calculations using DFPT, split roughly equally between the training phase and final results. This was made possible by optimizing the QUANTUM ESPRESSO electron-phonon code. In the final results, we identify 36 dynamically stable superconductors with $T_c > 100$ K of which 18 have $T_c > 200$ K (see Table 7.1). To the best of our knowledge, superconductivity has not been investigated previously in 27 of the 36 materials. These findings add considerably to the known pressure-T_c behaviour of the binary hydrides. Of particular note, we find a $Pm\bar{3}m$ structure of NaH_6 to have a T_c of 248-279 K at 100 GPa, suggesting the exciting possibility of other low-pressure high-T_c hydride superconductors. We also identify $Pm\bar{3}$-MgH_{12} and $Fm\bar{3}m$-SrH_{10} as above-room-temperature superconductors at 500 GPa, as well as several near-room-temperature superconductors at lower pressures.

Throughout the work carried out for this chapter, our aim has been to consider as wide a range of binary compositions as possible; since our focus is on breadth,
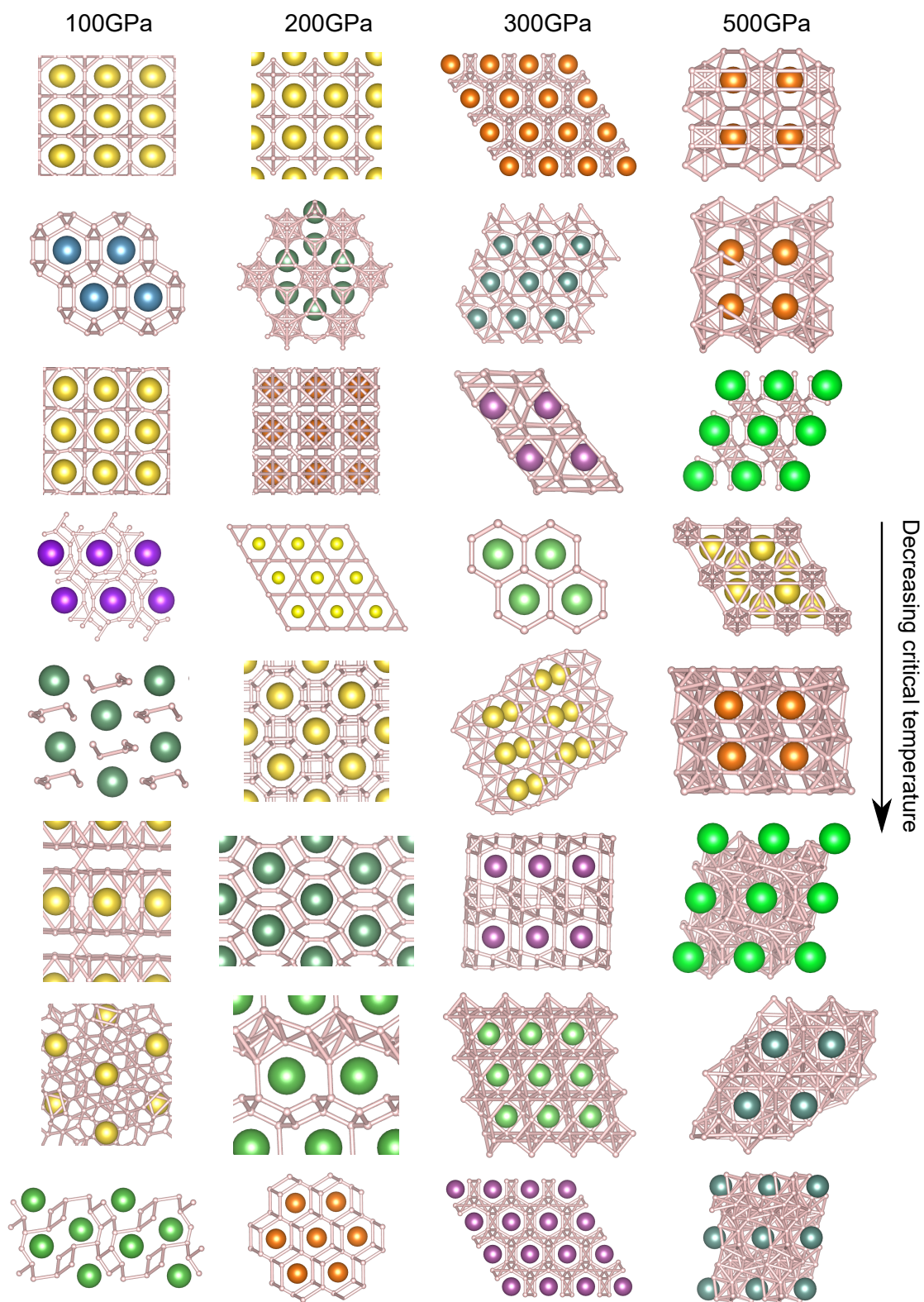
Figure 7.16: Structures from table 7.1, grouped in columns by pressure.
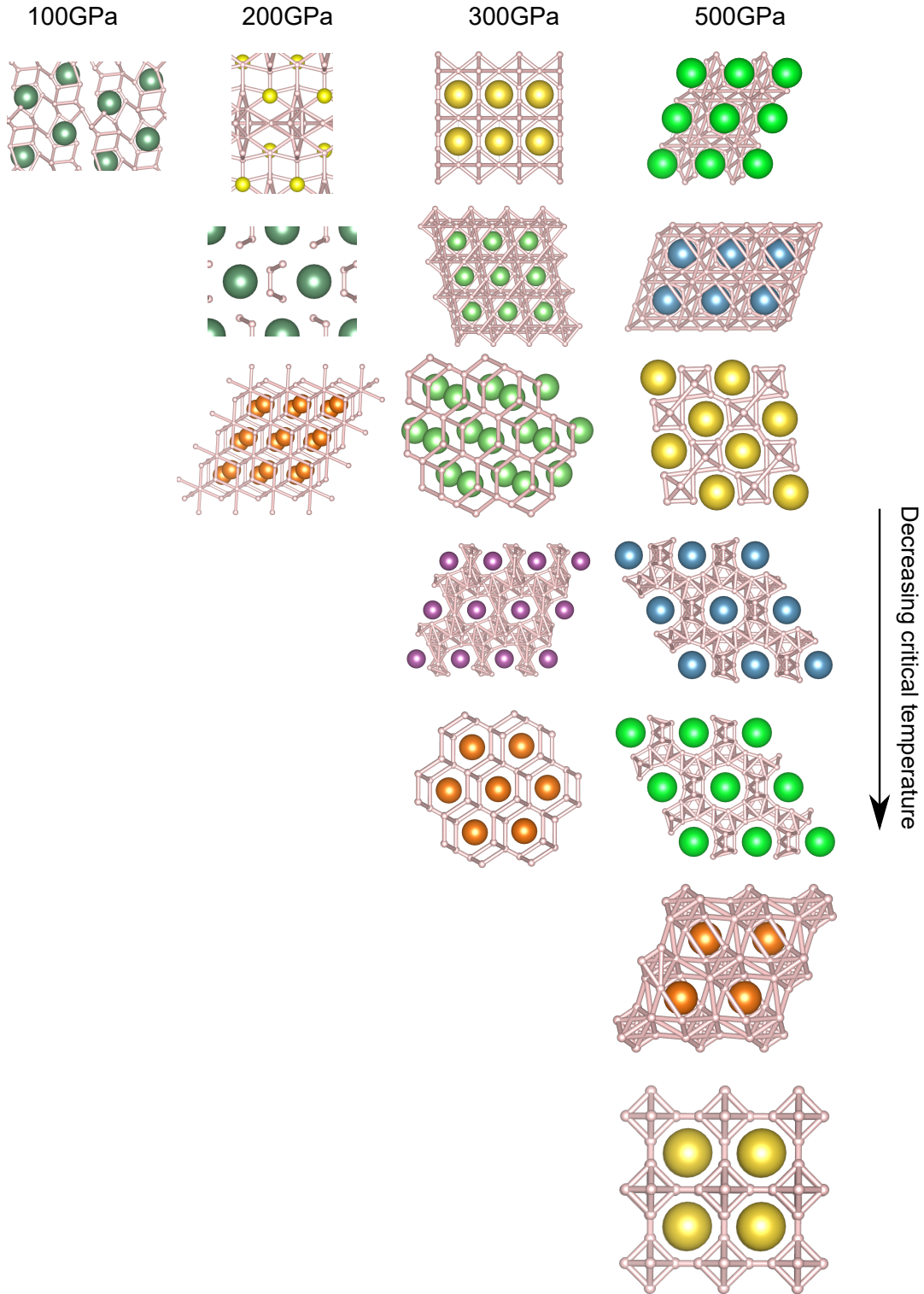
Figure 7.17: Structures from table 7.1, grouped in columns by pressure (vertical continuation of Fig. 7.16).

we make no claim that this study is exhaustive. Despite this, we identify a large number of high-T_c candidates, suggesting the binaries have more to offer in future, more focused, studies. We also note that the highest critical temperature results at each pressure arise from metastable structures or off-hull stoichiometries. Many of these were introduced into the study during the training phase by focusing on high-symmetry structures using sp-AIRSS. This suggests that the additional freedom afforded by allowing some degree of metastability can reveal higher critical temperature superconductors. For example, in the case of SrH_{10} at 500 GPa we see that the ground state $R\bar{3}m$ structure has a critical temperature nearly 100 K lower than a metastable $Fm\bar{3}m$ structure. Exploring avenues such as metastability will be important in future work in order to push the boundaries of high-temperature superconductivity.

Future work on superconducting hydrides is likely to focus on ternary hydrides and higher order systems. Given the increased complexity of these systems, high-throughput screening approaches, such as the one presented here, are likely to become increasingly important. Our high-throughput methodology could be extended to ternary hydrides, although it may be desirable to redefine these systems as effective binaries in order make use of the extensive binary hydride literature data during model training.

Chapter 8

Saddle point searching

8.1 Introduction

The task of geometry optimization (finding minima in the Born-Oppenheimer surface) is a well understood problem. Descent methods such as BFGS [205] are well suited to this problem and are fast enough to allow high-throughput methods such as AIRSS. However, there is another class of stationary points from which we may learn valuable information: saddle points. Near a stationary point (where first derivative terms vanish) the energy surface may be expanded as

$$E(x_{\min} + \delta) = E(x_{\min}) + \frac{1}{2}\delta^T H \delta + O(\delta^3) \tag{8.1}$$

Where H is the *Hessian*, defined by

$$H_{ij} = \frac{\partial^2 E}{\partial x_i \partial x_j} \tag{8.2}$$

Writing the displacement δ in terms of the normalized eigenvectors, v_i, of H (with corresponding eigenvalues λ_i) as $\delta = \sum_i \alpha_i v_i$, we have

$$E(x_{\min} + \delta) = E(x_{\min}) + \frac{\lambda_i \alpha_i^2}{2} + O(\delta^3) \tag{8.3}$$

We see that the nature of the stationary point is determined entirely by the eigenvalues of H. Eigenvectors with +ve (-ve) eigenvalue correspond to directions along which the stationary point is a minimum (maximum). In particular we are interested in *first-order* saddle points, with one negative eigenvalue and the rest positive. These points are of interest for two reasons. Firstly, the lowest activation energy transition path between two minima necessarily passes through a first order saddle point. This point is known as the *transition state* and is the highest-energy point along the path, corresponding to the activation energy. Secondly, first order saddle points are the primary candidates for metastable states which are stabilised by thermal effects. This can be seen by considering the quantum mechanics of a crystal in
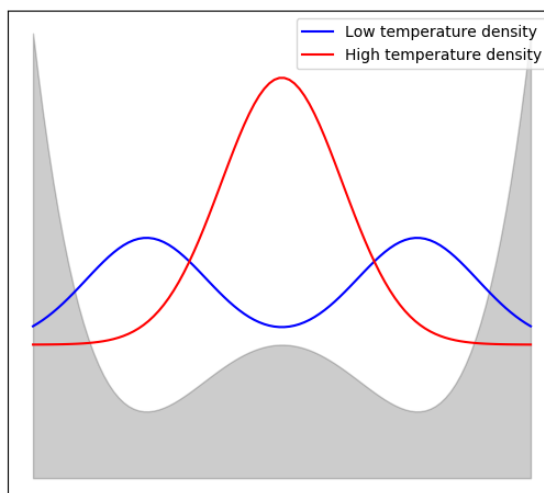
Figure 8.1: The potential $V(\alpha_i) = \alpha_i^4 - \alpha_i^2$ and the resulting particle densities at different temperatures.

a double well potential $V(\alpha_i) = \alpha_i^4 - \alpha_i^2$ as shown in figure 8.1 (we assume quadratic stability in all other directions). At low temperature, the particle density is centred around the nearby minima, and observations of the crystal structure will usually place it in, or near to, one of these minima. However, as the temperature increases and we begin to occupy exited states, we see that the most probable state is actually the higher-symmetry saddle-point phase (a phase that minima searching methods, such as AIRSS, would miss). Locating these saddle-point structures could lead to a better understanding of high temperature phases.

Algorithms to find saddle points exist, such as the nudged elastic band (NEB) method [112]. However, these are often focused on finding transition states given knowledge of both the initial and final states. The approach we pursue to find saddle-points is to start at a minima and use a hill climbing algorithm to arrive at a first-order saddle point, without knowledge of the final state. This choice was made for a few reasons. Firstly, it provides us with a simple statement of the problem and known starting conditions. Secondly, this allows us to associate saddle point with a particular initial minima. We can then continue 'over the hill', performing a subsequent geometry optimization, to find the final state for which this saddle point is the transition state. Thirdly, by the same method, this allows application to the problem of finding other nearby minima. Typically this problem is currently solved via a 'relax and shake' algorithm, whereby random perturbations are made to a stable structure, followed by geometry optimizations, in the hope that nearby minima will result. Going via a saddle point, rather than a random displacement, may increase the efficiency of this process as well as providing us with transition state information.

8.2 Hill climbing algorithms

In order to efficiently find saddle points, knowledge of the Hessian is extremely useful. Hill climbing algorithms are typically split into two groups [210]; *minimum mode* methods, which only require information about the eigenvector of the Hessian with lowest eigenvalue and *all mode* methods which require information about the complete Hessian. Because calculation of the exact Hessian for a crystal structure is computationally expensive, we focus on minimum mode methods and approximate-Hessian all mode methods. To best determine which information from the Hessian is most useful, we start by analysing the failure modes of an algorithm that does not use information from the Hessian at all.

8.2.1 Activation-relaxation

The *activation-relaxation* technique [185] employs only knowledge of the gradient vector (the negative of the force) to approach saddle points. The steps are shown in Algorithm 3. Essentially, we pick a direction (line 2) and attempt to minimize the energy in the perpendicular direction whilst increasing the energy in the parallel direction by inverting the parallel component of the force (line 8). We follow the resulting force, with the only additional constraint being that the step size does not exceed $\Delta_{\max}$ (line 9). The paths taken by this algorithm on a simple 2-D eggbox potential are shown in figure 8.2(a). The algorithm performs well on this potential, especially given the fact that it works without knowledge of the Hessian. However, we can quickly find situations where it works less well. If the normal that the force is projected along does not have a large component pointing up along the valley leading to the saddle point, then the algorithm fails to reach the saddle point. This situation appears when it is necessary to go 'around a corner' as shown in figure 8.2(b). Because the normal is directed away from the starting location at all times, the component of the force that is inverted does not necessarily point in the rough direction of the saddle-point. This causes the algorithm to fail to converge and, in some cases, get lost entirely. Clearly, this is because of a lack of local information about which directions are important. After all we only have two directions to work with; the force and the displacement from starting point.

8.2.2 Choice of direction

Intuitively, the choice of direction that we wish to move along is simple; we wish to walk 'up the valley'. At a given point in the valley x, this direction will be a local minimum of the energy change $\Delta E(\delta) = E(x + \delta) - E(x)$ for a fixed length vector δ. This is shown in figure 8.3. In order to determine this direction we follow the

---

Algorithm 3 Activation-relaxation

1: $s = \text{minima}$

2: $x = s + \text{random step}$

3: **repeat**

4:      $f = \text{force}(x)$

5:      $n = \frac{x-s}{|x-s|}$

6:      $f_{\parallel} = n(n \cdot f)$

7:      $f_{\perp} = f - f_{\parallel}$

8:      $\Delta = f_{\perp} - f_{\parallel}$

9:      **if** $|\Delta| > \Delta_{\max}$ **then**

10:          $\Delta = \Delta \Delta_{\max}/|\Delta|$

11:      **end if**

12:      $x = x + \Delta$

13: **until** $|f| < \text{tol}$
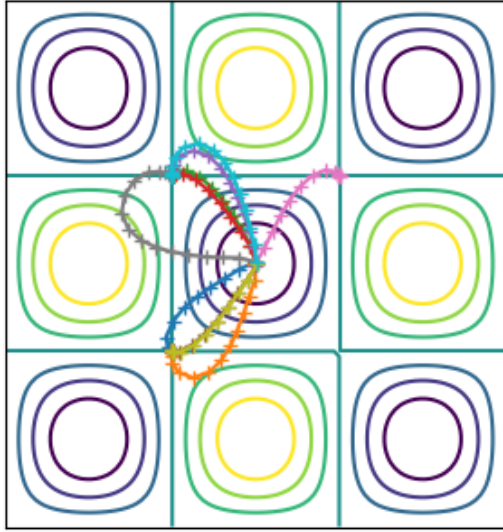
(a) Paths taken by the activation-relaxation algorithm for a simple eggbox potential.
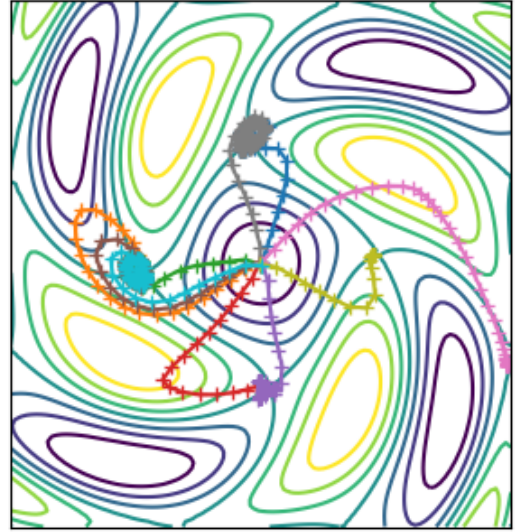
(b) Paths taken by the activation-relaxation algorithm for an eggbox potential that has been twisted.

Figure 8.2: Paths taken by the activation-relaxation algorithm. Low potential regions are shown with blue contours and high potential regions as yellow. We can see that the activation-relaxation algorithm performs well if the lowest-energy path to the saddle point is always directed away from the origin (as in (a)), but poorly if it is curved (as in (b)).
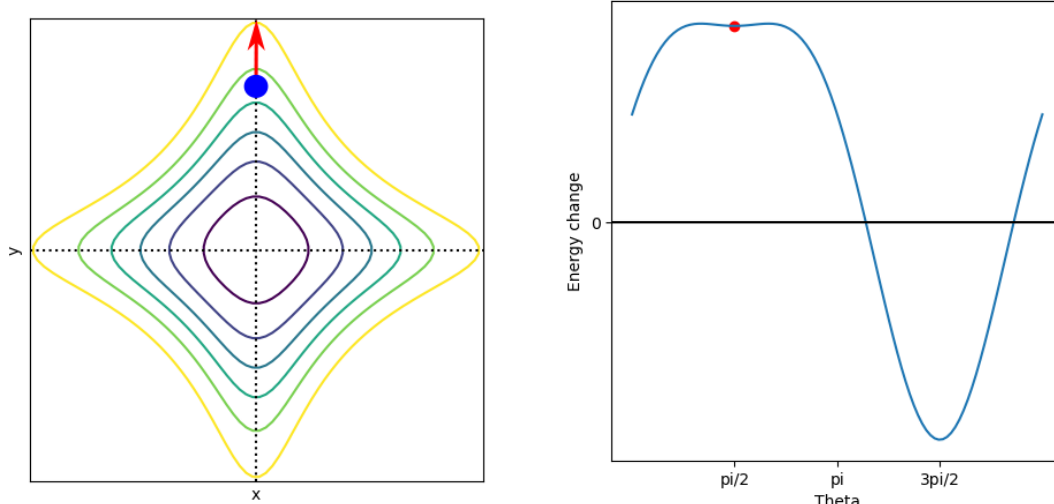
(a) A potential with 'valleys' shown with dotted black lines. At the point indicated by the blue dot, the direction to move 'up the valley' is shown by a red arrow.

(b) The energy change as a function of θ for a small step away from the blue dot in figure 8.3(a) in the direction $dx = cos(\theta), dy = sin(\theta)$. We see that the direction corresponding to a step 'up the valley' (shown by a red dot) is a local minimum.

Figure 8.3: Illustration of the optimal choice of direction for climbing up a valley in a potential energy surface.

approach in Ref. [43] [1]. First we Taylor expand the energy around x

$$\Delta E(\delta) = \nabla E \cdot \delta + \frac{1}{2}\delta^T H \delta + O(\delta^2) \tag{8.4}$$

Note, we are no longer at a minimum in E, so we have a non-zero gradient term. The direction we seek is an extremum of ΔE subject to the constraint that δ is of fixed length S. This can be found by employing the lagrange multiplier λ and instead extremizing

$$\mathcal{L} = \Delta E(\delta) + \lambda(\delta^T \delta - S^2) \tag{8.5}$$

The resulting extremization conditions are

$$\frac{\partial \mathcal{L}}{\partial \delta} = \nabla E + (H - \lambda)\delta \overset{!}{=} 0 \tag{8.6}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = S^2 - \delta^T \delta \overset{!}{=} 0 \tag{8.7}$$

This gives us two equations

$$\delta = (\lambda - H)^{-1}\nabla E \tag{8.8}$$

$$S^2 = [(\lambda - H)^{-1}\nabla E]^T (\lambda - H)^{-1}\nabla E \tag{8.9}$$

---

[1]We note a small mistake in Ref. [43]. Equation 2.13a should read (in their notation) $\Delta^2 = (\lambda - K)^{-1}D \cdot (\lambda - K)^{-1}D$

The first of these gives us an expression for the optimal step in terms of λ and the second constitutes an algebraic expression for λ. Because this method is based on a Taylor expansion of the energy surface it is known as a *rational function optimization* (RFO) method. We choose the smallest (i.e safest) step size S for which there are still solutions to Eq. 8.9 that increase the energy, leading to algorithm 4. Employing this algorithm with the exact Hessian for the same example we had previously gives the behaviour shown in figures 8.4(a) and 8.4(b). We see that this performs much better than activation-relaxation, as we would expect. The task now is to try to keep the performance of activation-relaxation, but to incorporate RFO-like information to improve the convergence properties and success rate.

Algorithm 4 RFO

1: $s = \text{minima}$

2: $x = s + \text{random step}$

3: **repeat**

4:     $g = -\text{force}(x)$

5:     $H = \text{Hessian}(x)$

6:     $S, \lambda = \text{solution of Eq. 8.9 with minimal } S$

7:     $\Delta = (\lambda - H)^{-1} g$

8:     **if** $|\Delta| > \Delta_{\max}$ **then**

9:         $\Delta = \Delta \Delta_{\max} / |\Delta|$

10:    **end if**

11:    $x = x + \Delta$

12: **until** $|f| < \text{tol}$

8.2.3 Improvements to activation-relaxation

Despite its flaws, the activation-relaxation algorithm is attractive because it only uses gradient information which, in DFT, we can get relatively cheaply from the Hellmann-Feynman theorem. Because of this we hope to find robust improvements to the algorithm that also only use gradient information. In particular we hope to improve the poor convergence shown in figure 8.2(b). Prompted by the ease at which the RFO algorithm finds the best direction to search given the exact Hessian, we look at constructing an approximate Hessian with the framework of the activation-relaxation approach. We consider two nearby points in configuration space, x and $x + \delta$. First, we take the Taylor expansion around these points,

$$E(x + \delta) = E(x) + \delta^T \nabla E|_x + \frac{1}{2} \delta^T H|_x \delta + O(\delta^3), \qquad (8.10)$$

$$E(x) = E(x + \delta) - \delta^T \nabla E|_{x+\delta} + \frac{1}{2} \delta^T H|_{x+\delta} \delta + O(\delta^3). \qquad (8.11)$$
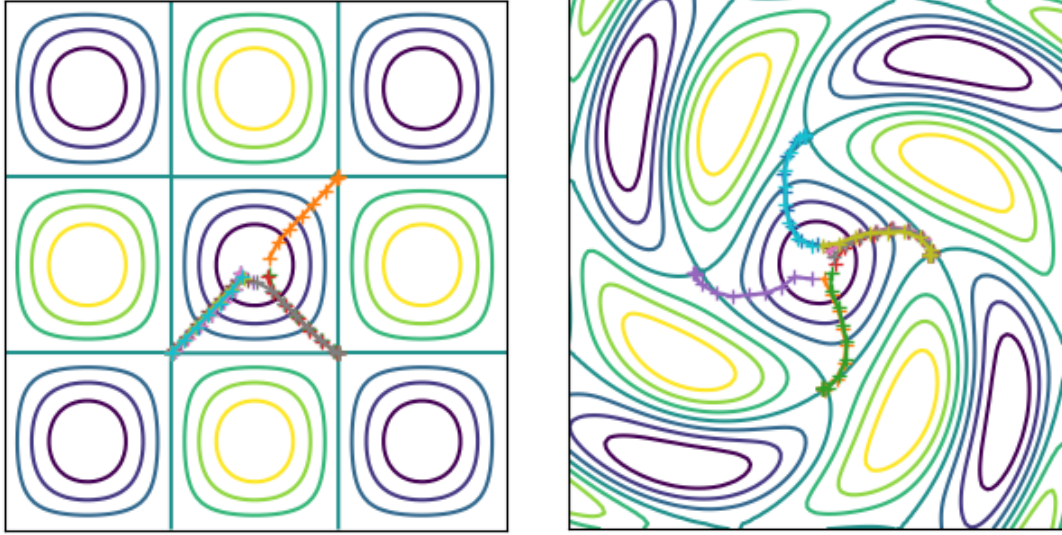
(a) Paths taken by the RFO algorithm using the exact Hessian for a simple eggbox potential.

(b) Paths taken by the RFO algorithm using the exact Hessian for an eggbox potential that has been twisted.

Figure 8.4: Paths taken by the RFO algorithm with the exact Hessian.

Summing these two equations we obtain

$$\delta^T \left[\frac{H|_x + H|_{x+\delta}}{2} \right] \delta = \delta^T (\nabla E|_{x+\delta} - \nabla E|_x). \tag{8.12}$$

Essentially this is a statement about the change in gradient induced by the Hessian. If we take the points x and $x + \delta$ to be adjacent steps in our activation-relaxation algorithm we have access to all of the terms in Eq. 8.12 except the term in square brackets; the average of the Hessian evaluated at x and $x + \delta$. Eq. 8.12 does not have a unique solution for the term in square brackets, but solutions do exist; the simplest of which is

$$\frac{H|_x + H|_{x+\delta}}{2} = \delta^T (\nabla E|_{x+\delta} - \nabla E|_x) \frac{\delta \delta^T}{|\delta|^4} \equiv H_{\mathrm{approx}}|_{x+\delta}. \tag{8.13}$$

Mixing H_{approx} across multiple iterations allows us to build up an approximation to the current Hessian. Once we have an approximate Hessian, we wish to apply the activation-relaxation algorithm to follow the minimum-mode up to the saddle point. Our implementation is shown in algorithm 5 and its performance is shown in figures 8.5(a) and 8.5(b). We see that the convergence of the algorithm has improved drastically over the activation-relaxation method for the twisted eggbox potential (compare figures 8.2(b) and 8.5(b)). For this potential the normal activation-relaxation algorithm failed to find a saddle point after 100 steps roughly half of the time. In contrast, the improved activation-relaxation employing our approximate Hessian never failed to reach a saddle point within 100 steps.

Algorithm 5 Approximate Hessian activation-relaxation

1: $s = \text{minima}$

2: $x = s + \text{random step}$

3: $x_l = s$

4: $f_l = 0$

5: $M = 0.1$ (Hessian mixing parameter)

6: $H = 0$

7: **repeat**

8:     $f = \text{force}(x)$

9:     **if** not first iteration **then**

10:         $\delta = x - x_l$

11:         $\delta_f = f - f_l$

12:         $H_{\text{approx}} = \delta^T \delta_f \delta \delta^T / |\delta|^4$

13:         $H = (1 - M)H + M H_{\text{approx}}$

14:     **end if**

15:     **if** H is diagonalizable **then**

16:         $n = \text{eigenvector of } H \text{ with minimum eigenvalue}$

17:     **else**

18:         $n = \frac{x-s}{|x-s|}$

19:     **end if**

20:     $f_\| = n(n \cdot f)$

21:     $f_\perp = f - f_\|$

22:     $\Delta = f_\perp - f_\|$

23:     **if** $|\Delta| > \Delta_{\text{max}}$ **then**

24:         $\Delta = \Delta \Delta_{\text{max}} / |\Delta|$

25:     **end if**

26:     $x = x + \Delta$

27: **until** $|f| < \text{tol}$
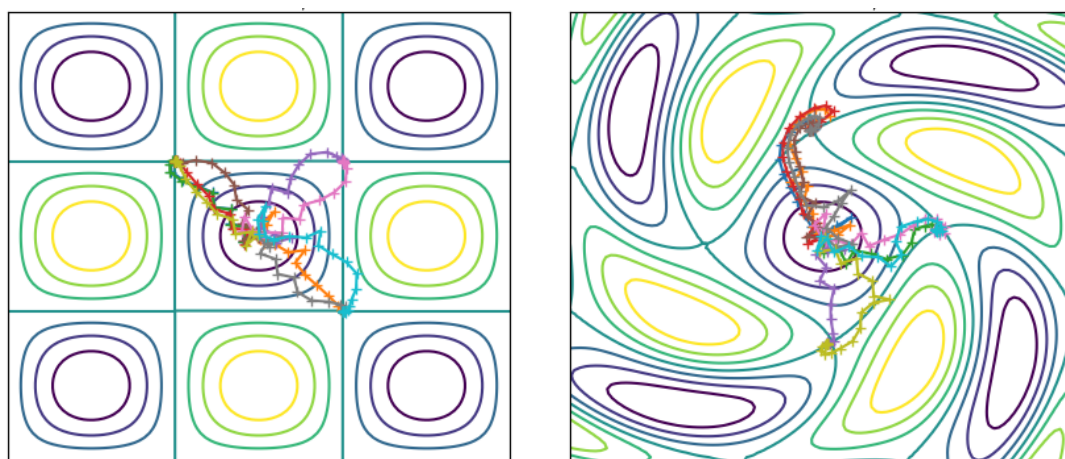
(a) Paths taken by the approximate Hessian algorithm for a simple eggbox potential.

(b) Paths taken by the approximate Hessian algorithm for an eggbox potential that has been twisted.

Figure 8.5: Paths taken by the approximate Hessian algorithm.

8.3 Limitations and future direction

The method developed in this work represents a straightforward improvement to the activation-relaxation algorithm for finding saddle points of functions. In particular, it improves the success rate of the algorithm. In cases where both activation-relaxation and the method described in this work succeed, the method described in this work also required fewer function evaluations than activation-relaxation. However, the number of function evaluations was still of roughly the same order-of-magnitude for both methods, limiting the application of the present method to problems of roughly the same size as could be attacked by activation-relaxation. For example, in order to find transition states in a chemical reaction, the function to be evaluated is the quantum-mechanical energy of the system. Depending on the quantum-chemistry method used to evaluate these energies, and the size of the system under investigation, the number of function evaluations required may be prohibitively expensive. However, preliminary testing indicates that small molecules (10 atoms) can be treated with this method using density functional theory in reasonable timescales on a desktop computer. The use of more accurate quantum chemistry methods such as coupled-cluster or configuration-interaction would require larger computers. On the other hand, treatment of large systems (potentially hundreds of atoms) with classical potentials should be possible - but at these system sizes the interpretation of particular saddle points becomes unclear, unless the dynamics are restricted to a particular sub-system reaction pathway (e.g adsorption of a molecule onto a small region of the surface of a large substrate). It is important to note that, if both ends of the activation pathway are known (i.e reactants

and products) then an algorithm such as nudged elastic band will be much more appropriate. The application for the methods considered in this chapter is strictly to finding transition states away from an arbitrary minimum, without knowledge of the end state.

Besides chemical reactions, this method could be used to search for saddle point crystal structures (of the kind that may be stabilized by anharmonic effects). However, in this case, the method would require some modification to properly deal with periodicity. In particular, one would need to be able to consider perturbations to the crystal structure that change the size of the unit cell (an additional degree of freedom to the atom coordinates). One naive way to achieve this (at the cost of increased expense) is to simply use a supercell of the system that is sufficiently large to describe final-state crystal structures of interest. Alternatively, one could occasionally increase the unit cell size by a random amount during the propagation of the algorithm to see if the extra degrees of freedom were useful in moving towards a saddle point - the unit cell resizing could then be accepted or rejected on this criteria.

Alternative methods for saddle point searching typically require additional information beyond the function and it's gradient. If the function evaluator being employed efficiently produces information about the Hessian (e.g if a second-order version of the Hellmann-Feynman theorem exists, or the same information is easily accessible using perturbation theory), then the RFO approach should be employed with the exact Hessian. In the intermediate regime, where evaluation of the exact Hessian is expensive, but not prohibitively so, one could imagine a modification of the method used in this work, whereby the approximate Hessian was updated to the exact Hessian periodically. This would correspond to evaluating the exact Hessian at some regular interval along the configuration-space trajectory as an additional guide the evolution of the trajectory (and to double-check how successfully the approximate Hessian is being constructed). Information from this exact Hessian would naturally remain for a few subsequent steps of the trajectory, thanks to the Hessian-mixing process employed by the algorithm.

8.4   Conclusion

By comparing the activation-relaxation and rational-function-optimization (RFO) saddle-point-searching algorithms, an improved version of activation-relaxation is developed that requires no additional information. By utilizing information from previous steps of the algorithm, a mixing scheme is developed that provides access to an approximate Hessian. This Hessian is used to decide the direction along

which to perform the "activation" in an improved activation-relaxation algorithm. The resulting algorithm is shown both to be more likely to converge and quicker to converge than activation-relaxation for model potentials, in particular when the path to a saddle point is non-trivial.

Chapter 9

Conclusions

The work carried out in this thesis spans the full depth of *in silico* materials discovery - from the methods used to tackle the fundamental quantum-mechanical problems presented by materials, to the construction of workflows targeting specific properties. It has shown that there are improvements to be made at every stage of the process, particularly when looking at novel classes of materials. These improvements will help theorists to lift more and more of the weight of discovery off of the shoulders of experimentalists, allowing the latter to focus precious resources only on the most promising candidates.

At several points throughout this thesis, novel computational methods are developed. The thesis begins with the development of a new quantum Monte Carlo method that demonstrates that it is possible to treat fermionic systems without the need for a trial wavefunction. This new method, along with several well-known limiting cases, is shown to arise straightforwardly from the treatment of the variational problem as a *constrained* optimization problem. During the thesis, optimizations are made to existing density functional theory (DFT) methods and screening models are developed from newly-available data. The thesis ends with the development of a straightforward improvement to one of the cheapest available saddle-point searching algorithms. These developments demonstrate progress is possible throughout the computational toolset supporting the materials discovery process.

Given suitable computational methods, this thesis demonstrates that an atomistic model of materials can be leveraged to predict a wide variety material properties accurately. It goes on to show that, when combined with crystal structure prediction (CSP) methods, these techniques allow for the discovery of new materials, with desirable properties, entirely *in silico*. In particular, it is shown that by combining CSP methods with density-functional perturbation theory (DFPT) calculations of superconducting properties, it is possible to predict novel structures of superconducting hydrides. An investigation of LaH_{10} and YH_{10} demonstrates similar

structural behaviour and provides an explanation for hexagonal impurities found in high-pressure experiments. The practical limitations resulting from such high pressures is emphasised and a machine-learning based screening approach is used to identify low-pressure superconducting hydrides of Rb and Cs. This work demonstrates that properly designed screening techniques can vastly reduce the amount of computation required to identify materials with specific properties. Careful profiling and optimization of the underlying DFPT calculations leads to a further order-of-magnitude speed up, which is leveraged to carry out an entire-periodic-table search for new superconductors on an unprecedented scale, elucidating general trends in hydride superconductors and identifying numerous high-T_c candidates. In particular, 200 K+ superconductors are identified at pressures as low as 100 GPa, indicating that, in the low-pressure limit, critical temperatures in the hydrides might be significantly higher than the highest conventional superconductors to date.

9.0.1 Directions for future research

Throughout this thesis directions for further research, or follow-up work, have been identified. Chapter 3 presented a mathematical framework for the construction of fermionic QMC methods and demonstrated a particular implementation. However, there are an infinite number of possible methods that satisfy the requirements for obtaining a fermionic state. In many ways the implementation presented in chapter 3 represents the simplest possible way to include the additional walker propagation channels arising from the antisymmetry constraint. Whilst the resulting *exchange moves* cannot modify the exchange symmetry of the overall state, the diffusive moves can. If symmetry is not preserved on a per-timestep basis, the bosonic component introduced at a particular timestep, however small, has the potential to grow into the dominant component - correct results were obtained in chapter 3 only when the rate of introduction of bosonic components was exceeded by the rate of their removal by a cancellation process. It seems likely that any method that relies on a cancellation process to overcome the appearance of bosonic components will eventually succumb to the sign problem (unless the rate of cancellation can be made to scale exponentially with system size). Therefore, it would be interesting to investigate if the exchange and diffusive moves could be combined in some way so that exchange symmetry was preserved on a per-timestep, rather than on-average, basis. In such future methods, the "second-quantized" walkers of Umrigar [270] might serve as the basic entities, to be combined with a symmetry-preserving propagation scheme. The re-introduction of a trial wavefunction for this purpose, or even just as a guide for importance sampling, should also not be overlooked in future work.

The study of the structure of Lithium in chapter 4 demonstrated that AIRSS

was an invaluable tool in identifying the constituents of a low-temperature mixed phase. The resulting mixed close-packed phase model accurately reproduced the experimental neutron pattern when the contribution from BCC remnants was included manually. However, the more straightforward method of simply thermodynamically weighting the structures and patterns from the AIRSS search reproduced many of the features of the experimental pattern, despite the relatively small scale of the search. This hints at the possibility of using more extensive AIRSS searches to construct mixed-phase structure models automatically - something that crystal structure prediction is already capable of for single-phase materials. This certainly holds promise for explaining diffraction patterns (as can be seen for Lithium in chapter 4), but one can imagine that this might be extended to generate grain structure and defect patterns with a more careful treatment of phase tessellation within the crystal. The extension of the quasi-harmonic treatment of vibrations to include anharmonic vibrations was also explored in chapter 4. The effect was not found to be of central importance to the phase behaviour, despite the small mass of Lithium atoms. However, high-temperature hydride superconductors have been the subject of increased interest in recent years and contain a large amount of hydrogen - the lightest element of all. Given that anharmonicity has already been shown to be of crucial importance for the pressure-dependence of the structure of LaH_{10} [80], and that the goal is to obtain superconductors that operate at room temperature, the quasi-anharmonic treatment of finite-temperatures demonstrated in chapter 4 may well find use in future work on hydride superconductors. Indeed, as more hydrides are synthesised in experiment, and neutron (or partial x-ray) patterns become available, the general approach presented in chapter 4 will undoubtedly inform future structural studies.

As more (both theoretical and, crucially, experimental) data on hydride superconductors becomes available, the screening methods developed in chapters 6 and 7 will become more powerful. They already proved invaluable for screening the compositional space of binary materials but will likely become an absolute necessity if we are to explore the space of ternary materials with any degree of completeness. Given the recent discovery of room temperature superconductivity in the C-S-H ternary system [248], broad searches of ternary space will likely appear soon. Such studies will also passively benefit from the optimizations carried out to open-source DFT software during this thesis. As well as looking ahead to ternary systems, the work presented in chapter 7 will directly serve as a guide in future experimental studies of binary hydride systems. Several compositions are suggested as synthesizable with the promise of high-temperature and/or low-pressure superconductivity; in particular synthesis of NaH_6 at 100 GPa and RbH_{12} at 50 GPa would provide important data points to probe the low-pressure limit of obtainable critical temperatures, and

therefore, the potential for superconductivity under ambient conditions.

Appendix A

Appendix

A.1   The Green's function of diffusion Monte Carlo

In this section, we derive the small-timestep ($\delta\tau \ll 1$) approximation of the imaginary-time Green's function

$$G(x, x', \delta\tau) = \langle x'|\exp(-\delta\tau(H - E_T))|x\rangle = \langle x'|\exp(-\delta\tau H)|x\rangle \exp(\delta\tau E_T) \quad \text{(A.1)}$$

With the partitioning of H given by Eq. 2.23 ($H = T + V$), we can use the Suzuki-Trotter [264] expansion of the operator $\exp(-\delta\tau H)$ for small $\delta\tau$:

$$\exp(-\delta\tau H) = \exp(-\delta\tau(T + V)) \approx \exp(-\delta\tau V/2)\exp(-\delta\tau T)\exp(-\delta\tau V/2). \quad \text{(A.2)}$$

Then

$$\begin{aligned}
G(x, x', \delta\tau) &\approx \langle x'|\exp(-\delta\tau V/2)\exp(-\delta\tau T)\exp(-\delta\tau V/2)|x\rangle \exp(\delta\tau E_T) \\
&= \exp\left(-\delta\tau[V(x') + V(x)]/2\right)\langle x'|\exp(-\delta\tau T)|x\rangle \exp(\delta\tau E_T)
\end{aligned} \quad \text{(A.3)}$$

where we have used the fact that the operator V is Hermitian in order to operate with it to the left to obtain $V(x')$. We denote the factor containing the potential related terms as G_V and that containing the kinetic terms as G_D:

$$\begin{aligned}
G_V(x, x', \delta\tau) &= \exp\left(-\delta\tau[V(x') + V(x)]/2\right) \\
G_D(x, x', \delta\tau) &= \langle x'|\exp(-\delta\tau T)|x\rangle.
\end{aligned} \quad \text{(A.4)}$$

The kinetic term is best evaluated in momentum-space;

$$\begin{aligned}
G_D(x, x', \delta\tau) &= \langle x'|\exp(-\delta\tau T)|x\rangle = \int \langle x'|\exp(-\delta\tau T)|p\rangle \langle p|x\rangle d^{3N}p \\
&= \int \langle x'|p\rangle \exp(-\delta\tau|M^{-1}p|^2/2)\exp(ip \cdot x)d^{3N}p \\
&= \int \exp(-ip \cdot x')\exp(-\delta\tau|M^{-1}p|^2/2)\exp(ip \cdot x)d^{3N}p \\
&= \exp\left(\frac{-|M(x - x')|^2}{2\delta\tau}\right)\int \exp(-z^2)d^{3N}(\sqrt{2/\delta\tau}Mz)
\end{aligned} \quad \text{(A.5)}$$

where the second to last line comes from completing the square using the variable substitution $z = \sqrt{\delta\tau} M^{-1} p / \sqrt{2} - iM(x - x')/\sqrt{2\delta\tau}$ and we have used the fact that momentum states $\langle p | x \rangle = \exp(ip \cdot x)$ are eigenstates of the kinetic energy with eigenvalue $|M^{-1}p|^2/2$ (see Eq. 2.23). The remaining integral is simply a constant; all dependence on x and x' is contained within the prefactor:

$$G_D(x, x', \delta\tau) \propto \exp\left(\frac{-|M(x - x')|^2}{2\delta\tau}\right) \tag{A.6}$$

A.2 Phonon coordinates

Here we show that the Hamiltonian in the regime of harmonic vibrations,

$$H_{\text{nuc}}^{(2)} = \sum_{R_i} -\frac{1}{2m_i}\nabla_{R_i}^2 + \frac{1}{2}\sum_{i,j}\delta_i\delta_j\partial_i\partial_j V(R)|_{R^{(0)}} \tag{A.7}$$

can be written in terms of phonon coordinates,

$$p_{q,\sigma} = \sum_i \sqrt{m_i}\exp(-iq \cdot R_i^{(0)})\delta_i P_i(q,\sigma) \tag{A.8}$$

as

$$H_{\text{vib}}^{(2)} = \sum_{q,\sigma} -\frac{1}{2}\frac{\partial^2}{\partial p_{q,\sigma}} + \frac{1}{2}p_{q,\sigma}^2\omega_{q,\sigma}^2 \tag{A.9}$$

We begin by noting that $H_{\text{vib}}^{(2)}$ can be written as

$$\begin{aligned}
H_{\text{vib}}^{(2)} = \sum_{R_i} &-\frac{1}{2m_i}\nabla_{R_i}^2 \\
&+ \frac{1}{2}\sum_{i,j}\sqrt{m_i m_j}\delta_i\delta_j D_{i,j}(q)\exp\left[-iq \cdot (R_j^{(0)} - R_i^{(0)})\right]
\end{aligned} \tag{A.10}$$

which, if we define $\Delta_i = \sqrt{m_i}\exp\left[-iq \cdot R_i^{(0)}\right]\delta_i$, can be written as

$$H_{\text{vib}}^{(2)} = \sum_{R_i} -\frac{1}{2m_i}\nabla_{R_i}^2 + \frac{1}{2}\Delta^\dagger D\Delta \tag{A.11}$$

The dynamical matrix, D, is diagonalized by our eigenvectors $P_i(q,\sigma)$ via $P^\dagger D P = \Omega$, where $P_{\sigma,i} = P_i(q,\sigma)$ and $\Omega_{\sigma,\sigma} = \omega_{q,\sigma}^2$. This allows us to write $\Delta^\dagger D\Delta = \Delta^\dagger P\Omega P^\dagger\Delta$. We then see that

$$(P^\dagger\Delta)_\sigma = \sum_i P_{\sigma,i}^\dagger\Delta_i = \sum_i P_i(q,\sigma)\sqrt{m_i}\exp\left[-iq \cdot R_i^{(0)}\right]\delta_i = p_{q,\sigma} \tag{A.12}$$

These are our phonon coordinates (compare to Eq. A.8). Our Hamiltonian can then be written as

$$\begin{aligned}
H_{\text{nuc}}^{(2)} &= \sum_{R_i} -\frac{1}{2m_i}\nabla_{R_i}^2 + \frac{1}{2}\sum_{q,\sigma}p_{q,\sigma}^2\omega_{q,\sigma}^2 \\
&= \sum_i -\frac{1}{2m_i}\partial_i^2 + \frac{1}{2}\sum_{q,\sigma}p_{q,\sigma}^2\omega_{q,\sigma}^2
\end{aligned} \tag{A.13}$$

We now turn our attention to the kinetic energy term. Using the chain rule we have

$$\partial_i = \frac{\partial}{\partial \delta_i} = \sum_{q,\sigma} \frac{\partial p_{q,\sigma}}{\partial \delta_i} \frac{\partial}{\partial p_{q,\sigma}} \tag{A.14}$$

Noting that

$$\frac{\partial p_{q,\sigma}}{\partial \delta_i} = P_i(q,\sigma) \exp\left[-iq \cdot R_i^{(0)}\right] \sqrt{m_i} \tag{A.15}$$

We then have

$$\sum_i \frac{1}{m_i} \partial_i^2 = \sum_{\substack{i,q,\sigma \\ q',\sigma'}} P_i(q,\sigma) P_i(q',\sigma') \exp\left[-i(q-q') \cdot R_i^{(0)}\right] \frac{\partial^2}{\partial p_{q,\sigma}^2} \tag{A.16}$$

Using the fact that $\sum_i \exp\left[-i(q-q') \cdot R_i^{(0)}\right] = \delta_{q,q'}$ and that our eiegnvectors are normalized, we arrive at

$$\sum_i \frac{1}{m_i} \partial_i^2 = \sum_{q,\sigma} \frac{\partial^2}{\partial p_{q,\sigma}^2} \tag{A.17}$$

which allows us to write, finally,

$$H_{\text{nuc}}^{(2)} = \sum_{q,\sigma} -\frac{1}{2} \frac{\partial^2}{\partial p_{q,\sigma}} + \frac{1}{2} p_{q,\sigma}^2 \omega_{q,\sigma}^2 \tag{A.18}$$

A.3   Generalized linear regression

The goal of generalized linear regression is to make a linear model of the relationship

$$\underbrace{O_n}_{\text{Observable}} = O_n(\underbrace{f_{1,n}, f_{2,n}, \ldots, f_{N,n}}_{\text{Features}}) \tag{A.19}$$

The linear model looks like

$$\underbrace{P_n}_{\text{Prediction}} = \underbrace{\sum_i c_i f_{i,n}}_{\text{Linear model}} + \underbrace{\epsilon_n}_{\text{Residual}} \tag{A.20}$$

Where c_i are the linear coefficients of the features. In matrix notation this looks like $P = fc + \epsilon$, where we call f the *feature matrix*. We pick the coefficient vector c by minimizing the modulus of the residual vector; $c = \text{argmin}_c |\epsilon|^2$. In order to simplify the model, we may bias the coefficient vector using Tikhonov regularization. This involves adding a cost function, $|Tc|^2$, which is large when the coefficient vector has many significant entries. The matrix T is known as the Tikhonov matrix. This results in the minimization

$$\begin{aligned} c &= \text{argmin}_c \ |\epsilon|^2 + |Tc|^2 \\ &= \text{argmin}_c \ |fc - P|^2 + |Tc|^2 \\ &\equiv \text{argmin}_c \ \mathcal{L}(c) \end{aligned} \tag{A.21}$$

Using implied summation our objective function can be written as

$$\mathcal{L}(c) = (f_{ij}c_j - P_i)(f_{ik}c_k - P_i) + T_{il}c_l T_{im}c_m \tag{A.22}$$

Minimizing with respect to c_n we require

$$\frac{\partial \mathcal{L}}{\partial c_n} = 2[f_{in}(f_{ik}c_k - P_i) + T_{in}T_{ij}c_j] \overset{!}{=} 0 \tag{A.23}$$

In matrix notation this reads

$$f^T(fc - p) + T^T T c = 0 \implies c = (f^T f + T^T T)^{-1} f^T p \tag{A.24}$$

which gives us the optimal feature coefficients c for the model in Eq. A.20.

A.4 Neutron patterns

The neutron patterns in the main text were derived by fitting

$$I(2\theta) = aI_{\mathrm{BCC}}(2\theta) + b\sum_i I_i(2\theta) \frac{w_i}{\exp\left(\frac{E_i - \mu}{k_B T}\right) - 1} \tag{A.25}$$

to the experimental data, where $I_{\mathrm{BCC}}(2\theta)$ is the remnant BCC pattern and $\{I_i(2\theta)\}$ are the close packed patterns. The variables a, b, μ and T are treated as fitting parameters. Different, equivalent, close packed structures are included via a multiplicity ω_i, described in the main text. In order to derive this form we consider the system as a mixed phase, consisting of m_i atoms in the ith phase. We then write the energy as

$$E = \sum_i m_i E_i \tag{A.26}$$

where E_i is the per atom energy of the ith phase. In order to write this we neglect the energetic contributions of defects. The resulting partition function in the grand canonical ensemble can be written as;

$$\mathcal{Z} = \sum_{\{m_i\}} \exp\left(-\beta \sum_i m_i(E_i - \mu)\right) \tag{A.27}$$

Where μ is a lagrange multiplier fixing the number (or density) of atoms and the sum is over possible sets of phase occupation numbers m_i. Assuming an infinite

system, we can manipulate this into a nicer form:

$$\mathcal{Z} = \sum_{\{m_i\}} \exp\left(-\beta \sum_i m_i(E_i - \mu)\right) \tag{A.28}$$

$$= \sum_{\{m_i\}} \prod_i \exp\left(-\beta m_i(E_i - \mu)\right) \tag{A.29}$$

$$= \prod_i \sum_{m_i=0}^{\infty} \exp\left(-\beta m_i(E_i - \mu)\right) \tag{A.30}$$

$$= \prod_i \frac{1}{1 - \exp\left(-\beta m_i(E_i - \mu)\right)} \tag{A.31}$$

From Eq. A.27 it is clear to see that

$$\langle m_i \rangle = -\frac{1}{\beta} \frac{\partial \ln(\mathcal{Z})}{\partial E_i} = \frac{1}{\exp\left(\beta(E_i - \mu)\right) - 1} \tag{A.32}$$

which we use to weight our i^{th} pattern.

A.5   Variation of the Lagrangian $\mathcal{L}$

Given a Hamiltonian H, we define the effective Hamiltonian

$$H_X = H - E_T + \sum_i \mu_i(x)(P_i + 1) \tag{A.33}$$

and look for extrema of

$$\mathcal{L}[\psi] = E_T + \langle\psi| H_X |\psi\rangle \tag{A.34}$$

with respect to variation of ψ and ψ^*. Variations in ψ^* are straightforward

$$\mathcal{L}[\psi^* + \delta\psi^*] = E_T + \int (\psi^* + \delta\psi^*) H_X \psi dx$$

$$= \mathcal{L}[\psi] + \int \delta\psi^* H_X \psi dx \tag{A.35}$$

$$\underbrace{\overset{!}{=} \mathcal{L}[\psi] \ \forall \ \delta\psi^*}_{\text{Extremization}} \implies H_X \psi = 0.$$

Variations in ψ are more involved

$$\mathcal{L}[\psi + \delta\psi] = E_T + \int \psi^* H_X(\psi + \delta\psi) dx$$

$$= \mathcal{L}\mathcal{L}[\psi] + [\psi] + \int \psi^* H_X \delta\psi dx \tag{A.36}$$

$$= \mathcal{L}[\psi] + \int \psi^* \left[T + V + \sum_i \mu_i(1 + P_i)\right] \delta\psi.$$

We can shift the kinetic term to instead operate on ψ^* by using integration by parts twice:

$$
\int \psi^* \frac{\partial^2 \delta\psi}{\partial x_i^2} dx
$$
$$
= \left[\psi^* \frac{\partial \delta\psi}{\partial x_i} \right]_\infty^0 - \int \frac{\partial \psi^*}{\partial x_i} \frac{\partial \delta\psi}{\partial x_i} dx \tag{A.37}
$$
$$
= -\left[\frac{\partial \psi^*}{\partial x_i} \delta\psi \right]_\infty^0 + \int \frac{\partial^2 \psi^*}{\partial x_i^2} \delta\psi dx
$$

where we have assumed that $\psi \to 0$ as $|x| \to \infty$ to cancel the boundary terms. We can also act with permutation operators to the left within the integral because

$$
\int f(x) P_i g(x) dx
$$
$$
= \int f(x) g(P_i x) dx
$$
$$
\text{let } z = P_i x \to
$$
$$
= \int f(P_i z) g(z) dz \tag{A.38}
$$
$$
\text{relabel } z \to x
$$
$$
= \int f(P_i x) g(x) dx.
$$

Putting this together we can write

$$
\mathcal{L}[\psi + \delta\psi] = \mathcal{L}[\psi] + \int \delta\psi \left[T + V + \sum_i (1 + P_i) \mu_i \right] \psi^* dx \tag{A.39}
$$

where the permutation operators act to the right. Note that μ_i now appears after the permutation operators. If we assume μ_i is symmetric with respect to permutations (in the main text it is a constant because we take the exchanges to be equiprobable) then we can pull it back through the permeation operators and write

$$
\mathcal{L}[\psi + \delta\psi] = \mathcal{L}[\psi] + \int \delta\psi H_X \psi^* dx
$$
$$
\underbrace{\stackrel{!}{=} \mathcal{L}[\psi] \ \forall \ \delta\psi}_{\text{Extremization}} \implies H_X \psi^* = 0. \tag{A.40}
$$

A.6 The Green's function of H_X

For small timesteps, we derive the form of the Green's function

$$
G(x, x', \delta\tau) = \langle x | \exp\left(-\delta\tau H_X\right) | x' \rangle. \tag{A.41}
$$

Writing $H = T + V$ where T is the kinetic energy operator and V is a local potential we can apply the Suzuki-Trotter [264, 90] expansion of the Green's function to obtain

$$G(x, x', \delta\tau) \approx G_V(x, x', \delta\tau) \times$$
$$\langle x| \exp(-\delta\tau[T + \sum_i \mu_i(x)(P_i + 1)]) |x'\rangle. \tag{A.42}$$

For small timesteps, the exponential can be factorized, allowing us to write

$$\langle x| \exp(-\delta\tau T)[1 - \sum_i \delta\tau\mu_i(x)(P_i + 1)] |x'\rangle$$
$$\approx \left[1 - \sum_i \delta\tau\mu_i(x')\right] G_D(x, x', \delta\tau) \tag{A.43}$$
$$- \sum_i \delta\tau\mu_i(P_i x') G_D(x, P_i x', \delta\tau)$$

where we have introduced the diffusive Green's function

$$G_D(x, x', \delta\tau) = \langle x| \exp(-\delta\tau T) |x'\rangle, \tag{A.44}$$

whos form is derived in appendix A.1. Noting that

$$|x - P_i x'| = |P_i(x - P_i x')| = |P_i x - x'|$$
$$\implies G_D(x, P_i x', \delta\tau) = G_D(P_i x, x', \delta\tau) \tag{A.45}$$

and that, because P_i corresponds to exchanging identical particles,

$$V(P_i x) = V(x) \implies G_V(P_i x, x', \delta\tau) = G_V(x, x', \delta\tau), \tag{A.46}$$

we can finally write the Green's function as

$$G(x, x', \delta\tau) = \left[\left(1 - \sum_i \delta\tau\mu_i(x')\right) - \sum_i \delta\tau\mu_i(P_i x')P_i \right]$$
$$\times G_V(x, x', \delta\tau) G_D(x, x', \delta\tau) \tag{A.47}$$

where P_i now acts on the unprimed (pre-propagation) coordinates.

A.7 Maximum-separation propagation

In order to encourage the formation of nodal pockets, we seek the form of $P_\pm(x)$ that maximizes the expected separation of +ve and -ve walkers, given by

$$\langle |x_+ - x_-| \rangle = \int P_+(x_+) P_-(x_-) |x_+ - x_-| dx_+ dx_-. \tag{A.48}$$

This is equivalent to extremizing

$$\mathcal{S} = \int S_+^2(x_+) S_-^2(x_-) |x_+ - x_-| dx_+ dx_-$$
$$+ \int \lambda(x)[S_+^2(x) - S_-^2(x) - \psi_D(x)]dx \tag{A.49}$$

with respect to $S_{\pm}^2(x) = P_{\pm}(x)$ (introduced to ensure $P_{\pm}(x) \geq 0$) and the Lagrange multiplier $\lambda(x)$ which enforces the constraint $\psi_D(x) = P_+(x) - P_-(x)$. Extremization of $\mathcal{S}$ leads to

$$\frac{\delta \mathcal{S}}{\delta S_+(y)} = \int 2S_+(y)S_-^2(z)|z - y|dz + 2S_+(y)\lambda(y) \stackrel{!}{=} 0, \tag{A.50}$$

$$\frac{\delta \mathcal{S}}{\delta S_-(y)} = \int 2S_-(y)S_+^2(z)|z - y|dz - 2S_-(y)\lambda(y) \stackrel{!}{=} 0. \tag{A.51}$$

Now, if we assume that both $S_+(y) \neq 0$ and $S_-(y) \neq 0$, Eqs. A.50 and A.51 read

$$\frac{1}{2S_+(y)}\frac{\delta \mathcal{S}}{\delta S_+(y)} = \int S_-^2(z)|z - y|dz + \lambda(y) = 0, \tag{A.52}$$

$$\frac{1}{2S_-(y)}\frac{\delta \mathcal{S}}{\delta S_-(y)} = \int S_+^2(z)|z - y|dz - \lambda(y) = 0. \tag{A.53}$$

Adding these equations gives

$$\int [S_+^2(z) + S_-^2(z)]|z - y|dz = 0 \implies S_+^2(z) + S_-^2(z) = 0, \tag{A.54}$$

a contradiction. This means that at most one of $S_+^2(y) = P_+(x)$ and $S_-^2(y) = P_-(x)$ is non-zero (i.e the distributions of +ve walkers and -ve walkers are mutually exclusive). Combined with the condition $\psi_D(x) = P_+(x) - P_-(x)$, we must have

$$P_{\pm}(x) = \begin{cases} |\psi_D(x)| & \text{if sign}(\psi_D(x)) = \pm 1, \\ 0 & \text{otherwise.} \end{cases} \tag{A.55}$$

Note that this derivation does not depend on the form of $\psi_D(x)$. It also results in the same distributions $P_{\pm}(x)$ for any measure of separation that is symmetric in x_+ and x_-, not just $|x_+ - x_-|$.

A.8 Form of cancellation functions

Here, we show that

$$P_{\pm}(x) = \begin{cases} |\psi_D(x)| & \text{if sign}(\psi_D(x)) = \pm 1, \\ 0 & \text{otherwise.} \end{cases} \tag{A.56}$$

can be written as

$$P_{\pm}(x) = \psi_{\pm}(x)f_{\pm}(x) \tag{A.57}$$

where

$$\psi_+(x) = \sum_{w_i > 0} w_i G_D(x, x_i(\tau), \delta\tau) \geq 0,$$
$$\psi_-(x) = \sum_{w_i < 0} |w_i| G_D(x, x_i(\tau), \delta\tau) \geq 0 \tag{A.58}$$

and

$$f_\pm(x) = \max\left(1 - \psi_\mp(x)/\psi_\pm(x),\ 0\right) \in [0, 1]. \tag{A.59}$$

Examining the form of $\psi_D(x)$, we have

$$\psi_D(x) = \psi_+(x) - \psi_-(x) \tag{A.60}$$

$$= \psi_+(x) \underbrace{\left[1 - \frac{\psi_-(x)}{\psi_+(x)}\right]}_{=\ f_+(x)\ \text{if}\ \psi_D(x) > 0} \tag{A.61}$$

$$= -\psi_-(x) \underbrace{\left[1 - \frac{\psi_+(x)}{\psi_-(x)}\right]}_{=\ f_-(x)\ \text{if}\ \psi_D(x) < 0} \tag{A.62}$$

and

$$\begin{aligned} 0 \leq f_+(x) \leq 1 \ \ \text{if} \ \ \psi_D(x) > 0, \\ 0 \leq f_-(x) \leq 1 \ \ \text{if} \ \ \psi_D(x) < 0. \end{aligned} \tag{A.63}$$

Using Eq. A.61 when $\psi_D(x) > 0$ and Eq. A.62 when $\psi_D(x) < 0$ allows us to combine both into the compact form of Eq. A.57. The form of Eq. A.57 then allows us to interpret $f_\pm(x)$ as a weight cancellation function. In certain limits, this function leads to cancellation-based schemes proposed in the past [195, 14]. The prefactor of $f_\pm(x)$ in Eq. A.57 is simply the diffused wavefunction for the corresponding sign, $\psi_\pm(x)$. This means we can diffuse a walker with weight w from $x \to y$ normally according to $G_D(y, x, \delta\tau)$ so long as we then apply the weight update

$$w \to \begin{cases} f_+(y)w & \text{if}\ \psi_+(y) > \psi_-(y)\ \&\ w > 0, \\ f_-(y)w & \text{if}\ \psi_+(y) < \psi_-(y)\ \&\ w < 0, \\ 0 & \text{otherwise.} \end{cases} \tag{A.64}$$

Where we evaluate $\psi_\pm(x, \tau + \delta\tau)$ directly via Eq. A.58.

A.9 Gaspari-Gyorffy theory

McMillan [192] showed that for strong-coupled superconductors the electron-phonon coupling constant, λ, can be expressed as

$$\lambda = 2 \int \frac{d\omega \alpha^2(\omega) F(\omega)}{\omega} = \frac{N(E_F)\langle I^2 \rangle}{M\langle \omega^2 \rangle} \tag{A.65}$$

λ can also be reformatted as

$$\lambda = \frac{\eta}{M\langle \omega^2 \rangle}$$

where η is the so-called Hopfield parameter. Hopfield was one of the first to stress the importance of the local environment in determining λ [119]. In situations where

we have nearly perfect separation of vibrational modes into those of different atomic character (such as we may see in hydrides) we can write

$$\lambda = \sum_j \lambda_j = \sum_j \frac{\eta_j}{M_j \langle \omega_j^2 \rangle} \tag{A.66}$$

where j is the atom type.

The quantity $\langle I^2 \rangle$ appearing in Eq. A.65 can be approximated using Gaspari-Gyorffy (GG) theory [97]. Recent work has emerged using this theory for metal hydrides under high pressure [212, 44] despite it originally being designed for elemental transition metals. The theory, based on the rigid muffin-tin approximation (RMTA), relies on several approximations [212] and allows us to reformulate the electron-phonon interaction in terms of phase shifts for a scattering potential. A self-consistent DOS calculation is thus all that is required to calculate $\langle I^2 \rangle$ for each atom type and hence obtain η_j. The GG equation is

$$\langle I^2 \rangle = \frac{E_F}{\pi^2 N^2(E_F)}$$
$$\sum_l \frac{2(l+1)\sin^2(\delta_{l+1} - \delta_l)N_l(E_F)N_{l+1}(E_F)}{N_l^{(1)}N_{l+1}^{(1)}} \tag{A.67}$$

where $N_l^{(1)}$ is the free-scatterer DOS given by

$$N_l^{(1)} = \frac{\sqrt{E_F}}{\pi}(2l+1)\int_0^{R_{MT}} R_l^2(r, E_F)r^2 dr \tag{A.68}$$

and the δ_l are the scattering phase shifts. Here R_{MT} is the muffin-tin radius associated with atom type j and R_l is the scattering solution of the Schrödinger equation. The phase shifts, which characterise the long-distance behaviour of the wavefunction, can be written in terms of the logarithmic derivative of the radial wavefunction

$$\tan(\delta_l(R_{MT}, E_F)) =$$
$$\frac{j_l'(kR_{MT}) - j_l(kR_{MT})L_l(R_{MT}, E_F)}{n_l'(kR_{MT}) - n_l(kR_{MT})L_l(R_{MT}, E_F)} \tag{A.69}$$

where $k = \sqrt{E_F}$, $L_l = R_l'/R_l$ is the logarithmic derivative, j_l are spherical Bessel functions and n_l are Neumann functions. We can therefore directly calculate the logarithmic derivative and use Eq. A.69 to obtain the phase shifts [229].

Since $M_j \langle \omega_j^2 \rangle$ is often considerably smaller for hydrogen than for the other components, it is clear from Eq. A.66 that the hydrogen atoms can provide a considerable fraction of λ even if the Hopfield parameter of the other atom type is similar in magnitude. Calculating η_H can therefore, in some cases, provide a cheap screening

method for identifying potential high-T_c hydrides. In particular, the average phonon frequencies for different structures are often similar when considering the same stoichiometry at the same pressure. If the average phonon frequencies are assumed to be exactly equivalent in such cases, we then arrive at a potential way of estimating T_c ordering between structures, simply by considering η_H. It is in this context that we assess the utility of GG theory in this work.

Bibliography

[1] http://elk.sourceforge.net/. The ELK FP-LAPW code.

[2] The XDMC code: https://github.com/miicck/xdmc.

[3] Input files for structures found in the machine learning chapter https://doi.org/10.17863/CAM.48347.

[4] https://uk.mathworks.com/help/stats/fitrgp.html. MATLAB GPR fitting documentation.

[5] Kazutaka Abe. "High-pressure properties of dense metallic zirconium hydrides studied by ab initio calculations". In: *Physical Review B* 98.13 (2018), p. 134103.

[6] Kazutaka Abe. "Hydrogen-rich scandium compounds at high pressures". In: *Physical Review B* 96.14 (2017), p. 144108.

[7] Kazutaka Abe and NW Ashcroft. "Crystalline diborane at high pressures". In: *Physical Review B* 84.10 (2011), p. 104118.

[8] G. J. Ackland and A. P. Jones. "Applications of local crystal structure measures in experiment and simulation". In: *Phys. Rev. B* 73 (5 Feb. 2006), p. 054104. DOI: 10.1103/PhysRevB.73.054104. URL: https://link.aps.org/doi/10.1103/PhysRevB.73.054104.

[9] Graeme J. Ackland et al. "Quantum and isotope effects in lithium metal". In: *Science* 356.6344 (2017), pp. 1254–1259. ISSN: 0036-8075. DOI: 10.1126/science.aal4886. URL: http://science.sciencemag.org/content/356/6344/1254.

[10] S. A. Alexander and R. L. Coldwell. "Atomic wave function forms". In: *International Journal of Quantum Chemistry* 63.5 (1997), pp. 1001–1022. DOI: 10.1002/(SICI)1097-461X(1997)63:5<1001::AID-QUA9>3.0.CO;2-1.

[11] P. B. Allen and R. C. Dynes. "Transition temperature of strong-coupled superconductors reanalyzed". In: *Physical Review B* 12.3 (1975), p. 905.

[12] P. B. Allen and R. C. Dynes. "Transition temperature of strong-coupled superconductors reanalyzed". In: *Phys. Rev. B* 12 (3 Aug. 1975), pp. 905–922. DOI: 10.1103/PhysRevB.12.905. URL: https://link.aps.org/doi/10.1103/PhysRevB.12.905.

[13] Philip B. Allen and Božidar Mitrović. "Theory of Superconducting Tc". In: ed. by Henry Ehrenreich, Frederick Seitz, and David Turnbull. Vol. 37. Solid State Physics. Academic Press, 1983, pp. 1–92. DOI: https://doi.org/10.1016/S0081-1947(08)60665-7. URL: http://www.sciencedirect.com/science/article/pii/S0081194708606657.

[14] James B. Anderson, Carol A. Traynor, and Bruce M. Boghosian. "Quantum chemistry by random walk: Exact treatment of many-electron systems". In: *The Journal of Chemical Physics* 95.10 (1991), pp. 7418–7425. DOI: 10.1063/1.461368.

[15] N. W. Ashcroft. "Hydrogen Dominant Metallic Alloys: High Temperature Superconductors?" In: *Phys. Rev. Lett.* 92 (18 May 2004), p. 187002. DOI: 10.1103/PhysRevLett.92.187002. URL: https://link.aps.org/doi/10.1103/PhysRevLett.92.187002.

[16] N. W. Ashcroft. "Hydrogen Dominant Metallic Alloys: High Temperature Superconductors?" In: *Phys. Rev. Lett.* 92.18 (2004), p. 187002.

[17] N. W. Ashcroft. "Metallic Hydrogen: A High-Temperature Superconductor?" In: *Phys. Rev. Lett.* 21 (26 Dec. 1968), pp. 1748–1749. DOI: 10.1103/PhysRevLett.21.1748. URL: https://link.aps.org/doi/10.1103/PhysRevLett.21.1748.

[18] N. W. Ashcroft. "Metallic Hydrogen: A High-Temperature Superconductor?" In: *Phys. Rev. Lett.* 21.26 (1968), p. 1748.

[19] N. W. Ashcroft. "Quantum-solid behavior and the electronic structure of the light alkali metals". In: *Phys. Rev. B* 39 (15 May 1989), pp. 10552–10559. DOI: 10.1103/PhysRevB.39.10552. URL: https://link.aps.org/doi/10.1103/PhysRevB.39.10552.

[20] Roland Assaraf, Michel Caffarel, and Anatole Khelif. In: *Journal of Physics A: Mathematical and Theoretical* 40.6 (Jan. 2007), pp. 1181–1214. DOI: 10.1088/1751-8113/40/6/001.

[21] Brian M. Austin, Dmitry Yu Zubarev, and William A. Lester. "Quantum Monte Carlo and Related Approaches". In: *Chemical Reviews* 112.1 (Jan. 2012), pp. 263–288. ISSN: 0009-2665. DOI: 10.1021/cr2001564.

[22] Pio Baettig and Eva Zurek. "Pressure-Stabilized Sodium Polyhydrides: NaH_n". In: *Phys. Rev. Lett.* 106 (23 June 2011), p. 237002. DOI: 10.1103/PhysRevLett.106.237002. URL: https://link.aps.org/doi/10.1103/PhysRevLett.106.237002.

[23] C. B. Barber, D. P. Dobkin, and H. T. Huhdanpaa. "The quickhull Algorithm for Convex Hulls". In: *ACM Trans. on Mathematical Software* 22.4 (1996), pp. 469–483.

[24] J. Bardeen, L. N. Cooper, and J. R. Schrieffer. "Theory of Superconductivity". In: *Phys. Rev.* 108 (5 Dec. 1957), pp. 1175–1204. DOI: 10.1103/PhysRev.108.1175. URL: https://link.aps.org/doi/10.1103/PhysRev.108.1175.

[25] Stefano Baroni et al. "Phonons and related crystal properties from density-functional perturbation theory". In: *Rev. Mod. Phys.* 73 (2 July 2001), pp. 515–562. DOI: 10.1103/RevModPhys.73.515. URL: https://link.aps.org/doi/10.1103/RevModPhys.73.515.

[26] C. S. Barrett. "X-ray study of the alkali metals at low temperatures". In: *Acta Crystallographica* 9.8 (1956), pp. 671–677. ISSN: 0365-110X. DOI: 10.1107/S0365110X56001790.

[27] C. S. Barrett and O. R. Trautz. "Transactions". In: *AIME* 175 (1948), p. 579.

[28] Francesco Belli et al. "Strong correlation between electronic bonding network and critical temperature in hydrogen-based superconductors". In: *Nature Communications* 12.1 (Sept. 2021), p. 5381. ISSN: 2041-1723. DOI: 10.1038/s41467-021-25687-0. URL: https://doi.org/10.1038/s41467-021-25687-0.

[29] G. Bergmann and D. Rainer. "The sensitivity of the transition temperature to changes in $\alpha 2F(\omega)$". In: *Zeitschrift für Physik* 263.1 (Aug. 1973), pp. 59–68. ISSN: 0044-3328. DOI: 10.1007/BF02351862. URL: https://doi.org/10.1007/BF02351862.

[30] R. Berliner and S. A. Werner. "Effect of stacking faults on diffraction: The structure of lithium metal". In: *Phys. Rev. B* 34 (6 Sept. 1986), pp. 3586–3603. DOI: 10.1103/PhysRevB.34.3586. URL: https://link.aps.org/doi/10.1103/PhysRevB.34.3586.

[31] T. Bi et al. "The Search for Superconductivity in High Pressure Hydrides". In: Elsevier, 2019.

[32] T. Bi et al. "The search for superconductivity in high pressure hydrides". In: Elsevier, 2019.

[33] Francis Birch. "Finite Elastic Strain of Cubic Crystals". In: *Phys. Rev.* 71 (11 June 1947), pp. 809–824. DOI: 10.1103/PhysRev.71.809. URL: https://link.aps.org/doi/10.1103/PhysRev.71.809.

[34] Peter Blaha et al. http://www.wien2k.at/reg_user/textbooks/usersguide.pdf/. Wien2k user guide. 2001.

[35] L. Boeri and G. B. Bachelet. "Viewpoint: the road to room-temperature conventional superconductivity". In: *J. Phys: Condens. Matt.* 31.23 (2019), p. 234002.

[36] G. E. P. Box and Mervin E. Muller. "A Note on the Generation of Random Normal Deviates". In: *The Annals of Mathematical Statistics* 29.2 (1958), pp. 610–611. DOI: 10.1214/aoms/1177706645. URL: https://doi.org/10.1214/aoms/1177706645.

[37] Sergiy Bubin and Ludwik Adamowicz. "Correlated-Gaussian calculations of the ground and low-lying excited states of the boron atom". In: *Phys. Rev. A* 83 (2 Feb. 2011), p. 022505. DOI: 10.1103/PhysRevA.83.022505.

[38] Kyle J. Caspersen and Emily A. Carter. "Finding transition states for crystalline solid–solid phase transformations". In: *Proceedings of the National Academy of Sciences* 102.19 (2005), pp. 6738–6743. ISSN: 0027-8424. DOI: 10.1073/pnas.0408127102.

[39] D. M. Ceperley. "Fermion nodes". In: *Journal of Statistical Physics* 63.5 (June 1991), pp. 1237–1267. ISSN: 1572-9613. DOI: 10.1007/BF01030009.

[40] D. M. Ceperley and B. J. Alder. "Ground State of the Electron Gas by a Stochastic Method". In: *Phys. Rev. Lett.* 45 (7 Aug. 1980), pp. 566–569. DOI: 10.1103/PhysRevLett.45.566.

[41] D. M. Ceperley and B. J. Alder. "Quantum Monte Carlo for molecules: Green's function and nodal release". In: *The Journal of Chemical Physics* 81.12 (1984), pp. 5833–5844. DOI: 10.1063/1.447637. URL: https://doi.org/10.1063/1.447637.

[42] D. Ceperley, G. V. Chester, and M. H. Kalos. "Monte Carlo simulation of a many-fermion study". In: *Phys. Rev. B* 16 (7 Oct. 1977), pp. 3081–3099. DOI: 10.1103/PhysRevB.16.3081. URL: https://link.aps.org/doi/10.1103/PhysRevB.16.3081.

[43] Charles J. Cerjan and William H. Miller. "On finding transition states". In: *The Journal of Chemical Physics* 75.6 (1981), pp. 2800–2806. DOI: 10.1063/1.442352.

[44] P.-H. Chang et al. "Pressure-Induced High Temperature Superconductivity in H_3X (X= As, Se, Br, Sb, Te and I)". In: *arXiv preprint arXiv:1903.03255* (2019), p. 109315.

[45] Changbo Chen et al. "Pressure induced phase transition in MH2 (M = V, Nb)". In: *The Journal of Chemical Physics* 140.11 (2014), p. 114703. DOI: `10.1063/1.4866179`. eprint: `https://doi.org/10.1063/1.4866179`. URL: `https://doi.org/10.1063/1.4866179`.

[46] Y. Chen et al. "Calculations for the transverse N-point phonons in bcc Zr, Nb, and Mo". In: *Phys. Rev. B* 31 (10 May 1985), pp. 6775–6778. DOI: `10.1103/PhysRevB.31.6775`.

[47] François Chollet et al. *Keras*. `https://keras.io`. 2015.

[48] S. J. Clark et al. "First principles methods using CASTEP". In: *Zeitschrift für Kristallographie* (2005), pp. 567–570. URL: `http://eprints.whiterose.ac.uk/8521/`.

[49] D. A. Clevert, T. Unterthiner, and S. Hochreiter. *Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs)*. 2015. eprint: `arXiv:1511.07289`.

[50] E. Richard Cohen, David R. Lide, and George L. Trigg. *Physics Desk Reference, 3rd edition*. Springer-Verlag New York, Inc., 2003, p. 826.

[51] Leon N. Cooper. "Bound Electron Pairs in a Degenerate Fermi Gas". In: *Phys. Rev.* 104 (4 Nov. 1956), pp. 1189–1190. DOI: `10.1103/PhysRev.104.1189`. URL: `https://link.aps.org/doi/10.1103/PhysRev.104.1189`.

[52] R. S. Crisp. "Observation of the low-temperature Martensitic transformation in Li and a Li-Mg alloy by soft X-ray emission". In: *Journal of Physics: Condensed Matter* 3.30 (1991), p. 5761. URL: `http://stacks.iop.org/0953-8984/3/i=30/a=009`.

[53] Philip Dalladay-Simpson, Ross T Howie, and Eugene Gregoryanz. "Evidence for a new phase of dense hydrogen above 325 gigapascals". In: *Nature* 529.7584 (2016), pp. 63–67.

[54] Shanti Deemyad and Rong Zhang. "Probing quantum effects in lithium". In: *Physica C: Superconductivity and its Applications* 548 (2018), pp. 68–71. ISSN: 0921-4534. DOI: `10.1016/j.physc.2018.02.007`. URL: `http://www.sciencedirect.com/science/article/pii/S0921453417302320`.

[55]   Ranga P. Dias and Isaac F. Silvera. "Observation of the Wigner-Huntington transition to metallic hydrogen". In: *Science* 355.6326 (2017), pp. 715–718. ISSN: 0036-8075. DOI: 10.1126/science.aal1579. eprint: https://science.sciencemag.org/content/355/6326/715.full.pdf. URL: https://science.sciencemag.org/content/355/6326/715.

[56]   Ranga P. Dias and Isaac F. Silvera. "Observation of the Wigner-Huntington transition to metallic hydrogen". In: *Science* 355.6326 (2017), pp. 715–718.

[57]   O. Dmitrieva et al. "Chemical gradients across phase boundaries between martensite and austenite in steel studied by atom probe tomography and simulation". In: *Acta Materialia* 59.1 (2011), pp. 364–374. ISSN: 1359-6454. DOI: 10.1016/j.actamat.2010.09.042. URL: http://www.sciencedirect.com/science/article/pii/S1359645410006191.

[58]   K. Doll, J. C. Schön, and M. Jansen. "Structure prediction based on ab initio simulated annealing for boron nitride". In: *Physical Review B* 78.144110 (2017).

[59]   A. P. Drozdov et al. "Conventional Superconductivity at 203 K at High Pressures in the Sulfur Hydride System". In: *Nature* 525.7567 (2015), p. 73.

[60]   A. P. Drozdov et al. "Superconductivity at 215 K in Lanthanum Hydride at High Pressures". In: *arXiv preprint arXiv:1808.07039* (2018).

[61]   A. P. Drozdov et al. "Superconductivity at 250 K in lanthanum hydride under high pressures". In: *Nature* 569.7757 (May 2019), pp. 528–531. ISSN: 1476-4687. DOI: 10.1038/s41586-019-1201-8. URL: https://doi.org/10.1038/s41586-019-1201-8.

[62]   A. P. Drozdov et al. "Superconductivity at 250 K in lanthanum hydride under high pressures". In: *Nature* 569.7757 (2019), p. 528.

[63]   AP Drozdov et al. "Conventional superconductivity at 203 kelvin at high pressures in the sulfur hydride system". In: *Nature* 525.7567 (2015), p. 73.

[64]   AP Drozdov et al. "Superconductivity at 250 K in lanthanum hydride under high pressures". In: *Nature* 569.7757 (2019), p. 528.

[65]   D. Duan et al. "Pressure-Induced Metallization of Dense $(H_2S)_2H_2$ with High-T_c Superconductivity". In: *Nature - Scientific Reports* 4 (2014), p. 6968.

[66]   D. Duan et al. "Structure and Superconductivity of Hydrides at High Pressures". In: *Natl. Sci. Rev* 4.1 (2017), pp. 121–135.

[67]   Defang Duan et al. "Enhancement of T c in the atomic phase of iodine-doped hydrogen at high pressures". In: *Physical Chemistry Chemical Physics* 17.48 (2015), pp. 32335–32340.

[68] Defang Duan et al. "Pressure-induced metallization of dense (H 2 S) 2 H 2 with high-T c superconductivity". In: *Scientific reports* 4 (2014), p. 6968.

[69] Defang Duan et al. "Structure and superconductivity of hydrides at high pressures". In: *National Science Review* 4.1 (2017), pp. 121–135.

[70] A. P. Durajski and R. Szczesniak. "Properties of the Pressure-Induced Superconducting State in Trihydrides ScH_3 and LaH_3". In: *Supercond. Sci. and Tech.* 27.11 (2014), p. 115012.

[71] G. M. Eliashberg. "Interactions Between Electrons and Lattice Vibrations in a Superconductor". In: *Sov. Phys. JETP* 11.3 (1960), pp. 696–702.

[72] G.M. Eliashberg. "Interactions between Electrons and Lattice Vibrations in a Superconductor". In: *Soviet Phys. JETP* 11.2 (). URL: `http://www.jetp.ac.ru/cgi-bin/dn/e_011_03_0696.pdf`.

[73] M. I. Eremets et al. "Superconductivity in Hydrogen Dominant Materials: Silane". In: *Science* 319.5869 (2008), pp. 1506–1509. ISSN: 0036-8075. DOI: `10.1126/science.1153282`. URL: `https://science.sciencemag.org/content/319/5869/1506`.

[74] MI Eremets and IA Troyan. "Conductive dense hydrogen". In: *Nature Materials* 10.12 (2011), pp. 927–931.

[75] I. Errea, M. Calandra, and F. Mauri. "First-Principles Theory of Anharmonicity and the Inverse Isotope Effect in Superconducting Palladium-Hydride Compounds". In: *Phys. Rev. Lett.* 111.17 (2013), p. 177002.

[76] I. Errea et al. "High-Pressure Hydrogen Sulfide from First Principles: A Strongly Anharmonic Phonon-Mediated Superconductor". In: *Phys. Rev. Lett.* 114.15 (2015), p. 157004.

[77] Ion Errea, Matteo Calandra, and Francesco Mauri. "Anharmonic free energies and phonon dispersions from the stochastic self-consistent harmonic approximation: Application to platinum and palladium hydrides". In: *Phys. Rev. B* 89 (6 Feb. 2014), p. 064302. DOI: `10.1103/PhysRevB.89.064302`. URL: `https://link.aps.org/doi/10.1103/PhysRevB.89.064302`.

[78] Ion Errea et al. "High-pressure hydrogen sulfide from first principles: A strongly anharmonic phonon-mediated superconductor". In: *Physical review letters* 114.15 (2015), p. 157004.

[79] Ion Errea et al. "Quantum crystal structure in the 250-kelvin superconducting lanthanum hydride". In: *Nature* 578.7793 (2020), pp. 66–69. DOI: `10.1038/s41586-020-1955-z`. URL: `https://doi.org/10.1038/s41586-020-1955-z`.

[80] Ion Errea et al. "Quantum crystal structure in the 250-kelvin superconducting lanthanum hydride". In: *Nature* 578.7793 (Feb. 2020), pp. 66–69. ISSN: 1476-4687. DOI: `10.1038/s41586-020-1955-z`. URL: `https://doi.org/10.1038/s41586-020-1955-z`.

[81] Ion Errea et al. "Quantum hydrogen-bond symmetrization in the superconducting hydrogen sulfide system". In: *Nature* 532.7597 (Apr. 2016), pp. 81–84. ISSN: 1476-4687. DOI: `10.1038/nature17175`. URL: `https://doi.org/10.1038/nature17175`.

[82] M Mahdi Davari Esfahani et al. "Superconductivity and unexpected chemistry of germanium hydrides under pressure". In: *Physical Review B* 95.13 (2017), p. 134506.

[83] M Mahdi Davari Esfahani et al. "Superconductivity of novel tin hydrides (Sn n H m) under pressure". In: *Scientific reports* 6 (2016), p. 22873.

[84] Xiaolei Feng et al. "Compressed sodalite-like MgH6 as a potential high-temperature superconductor". In: *RSC Adv.* 5 (73 2015), pp. 59292–59296. DOI: `10.1039/C5RA11459D`. URL: `http://dx.doi.org/10.1039/C5RA11459D`.

[85] Xiaolei Feng et al. "Compressed sodalite-like MgH6 as a potential high-temperature superconductor". In: *RSC Advances* 5.73 (2015), pp. 59292–59296.

[86] Enrico Fermi. "un metodo statistico per la determinazione di alcune proprietà dell'atomo". In: *rend. accad. naz. lincei.* 6 (1927), pp. 602–607.

[87] José A. Flores-Livas et al. "A perspective on conventional high-temperature superconductors at high pressure: Methods and materials". In: *Physics Reports* (2020). ISSN: 0370-1573. DOI: `https://doi.org/10.1016/j.physrep.2020.02.003`. URL: `http://www.sciencedirect.com/science/article/pii/S0370157320300363`.

[88] José A. Flores-Livas et al. "A perspective on conventional high-temperature superconductors at high pressure: Methods and materials". In: *Physics Reports* (2020). ISSN: 0370-1573. DOI: `https://doi.org/10.1016/j.physrep.2020.02.003`. URL: `http://www.sciencedirect.com/science/article/pii/S0370157320300363`.

[89] H. Flyvbjerg and H. G. Petersen. "Error estimates on averages of correlated data". In: *The Journal of Chemical Physics* 91.1 (1989), pp. 461–466. DOI: `10.1063/1.457480`.

[90] W. M. C. Foulkes et al. "Quantum Monte Carlo simulations of solids". In: *Rev. Mod. Phys.* 73 (1 Jan. 2001), pp. 33–83. DOI: `10.1103/RevModPhys.73.33`.

[91] Y. Fu et al. "High-Pressure Phase Stability and Superconductivity of Pnictogen Hydrides and Chemical Trends for Compressed Hydrides". In: *Chemistry of Materials* 28.6 (2016), pp. 1746–1755.

[92] Yuhao Fu et al. "High-pressure phase stability and superconductivity of pnictogen hydrides and chemical trends for compressed hydrides". In: *Chemistry of Materials* 28.6 (2016), pp. 1746–1755.

[93] Brent Fultz. "Vibrational thermodynamics of materials". In: *Progress in Materials Science* 55.4 (2010), pp. 247–352. ISSN: 0079-6425. DOI: `10.1016/j.pmatsci.2009.05.002`. URL: `http://www.sciencedirect.com/science/article/pii/S0079642509000577`.

[94] Guoying Gao et al. "Superconducting high pressure phase of germane". In: *Physical review letters* 101.10 (2008), p. 107002.

[95] Guoying Gao et al. "Theoretical study of the ground-state structures and properties of niobium hydrides under pressure". In: *Phys. Rev. B* 88 (18 Nov. 2013), p. 184104. DOI: `10.1103/PhysRevB.88.184104`. URL: `https://link.aps.org/doi/10.1103/PhysRevB.88.184104`.

[96] Guoying Gao et al. "Theoretical study of the ground-state structures and properties of niobium hydrides under pressure". In: *Physical Review B* 88.18 (2013), p. 184104.

[97] G. D. Gaspari and B. L. Gyorffy. "Electron-phonon interactions, d resonances, and superconductivity in transition metals". In: *Phys. Rev. Lett.* 28.13 (1972), p. 801.

[98] Z. M. Geballe et al. "Synthesis and Stability of Lanthanum Superhydrides". In: *Angewandte Chemie International Edition* 57.3 (2018), pp. 688–692.

[99] Zachary M. Geballe et al. "Synthesis and Stability of Lanthanum Superhydrides". In: *Angewandte Chemie International Edition* 57.3 (2018), pp. 688–692. DOI: `https://doi.org/10.1002/anie.201709970`. eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1002/anie.201709970`. URL: `https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.201709970`.

[100] Paolo Giannozzi et al. "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials". In: *Journal of Physics: Condensed Matter* 21.39 (2009), 395502 (19pp). URL: `http://www.quantum-espresso.org`.

[101] P Giannozzi et al. "Advanced capabilities for materials modelling with Quantum ESPRESSO". In: *Journal of Physics: Condensed Matter* 29.46 (Oct. 2017), p. 465901. DOI: `10.1088/1361-648x/aa8f79`. URL: `https://doi.org/10.1088%2F1361-648x%2Faa8f79`.

[102] Feliciano Giustino. "Electron-phonon interactions from first principles". In: *Rev. Mod. Phys.* 89 (1 Feb. 2017), p. 015003. DOI: `10.1103/RevModPhys.89.015003`. URL: `https://link.aps.org/doi/10.1103/RevModPhys.89.015003`.

[103] R. J. Gooding and J. A. Krumhansl. "Theory of the BCC-to-9R structural phase transformation of Li". In: *Phys. Rev. B* 38 (3 July 1988), pp. 1695–1704. DOI: `10.1103/PhysRevB.38.1695`. URL: `https://link.aps.org/doi/10.1103/PhysRevB.38.1695`.

[104] F. A. Gorelli et al. "Lattice Dynamics of Dense Lithium". In: *Phys. Rev. Lett.* 108 (5 Jan. 2012), p. 055501. DOI: `10.1103/PhysRevLett.108.055501`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.108.055501`.

[105] Jeffrey C. Grossman. "Benchmark quantum Monte Carlo calculations". In: *The Journal of Chemical Physics* 117.4 (2002), pp. 1434–1440. DOI: `10.1063/1.1487829`. eprint: `https://doi.org/10.1063/1.1487829`.

[106] Qinyan Gu et al. "High-temperature superconducting phase of HBr under pressure predicted by first-principles calculations". In: *Physical Review B* 96.6 (2017), p. 064517.

[107] C. Heil, G. B. Bachelet, and L. Boeri. "Absence of Superconductivity in Iron Polyhydrides at High Pressures". In: *Phys. Rev. B* 97.21 (2018), p. 214510.

[108] C. Heil and L. Boeri. "Influence of Bonding on Superconductivity in High-Pressure Hydrides". In: *Phys. Rev. B* 92.6 (2015), p. 060508.

[109] Christoph Heil, Giovanni B Bachelet, and Lilia Boeri. "Absence of superconductivity in iron polyhydrides at high pressures". In: *Physical Review B* 97.21 (2018), p. 214510.

[110] Christoph Heil et al. "Superconductivity in sodalite-like yttrium hydride clathrates". In: *Phys. Rev. B* 99.22 (2019), 220502(R).

[111] Christoph Heil et al. "Superconductivity in sodalite-like yttrium hydride clathrates". In: *Physical Review B* 99.22 (2019), p. 220502.

[112] Graeme Henkelman, Blas P. Uberuaga, and Hannes Jónsson. "A climbing image nudged elastic band method for finding saddle points and minimum energy paths". In: *The Journal of Chemical Physics* 113.22 (2000), pp. 9901–9904. DOI: `10.1063/1.1329672`.

[113]   Detlef W.M. Hofmanna and Joannis Apostolakis. "Crystal structure prediction by data mining". In: *Molecular Structure* 647 (2002), pp. 17–39.

[114]   P. Hohenberg and W. Kohn. "Inhomogeneous Electron Gas". In: *Phys. Rev.* 136 (3B Nov. 1964), B864–B871. DOI: `10.1103/PhysRev.136.B864`. URL: `https://link.aps.org/doi/10.1103/PhysRev.136.B864`.

[115]   P. Hohenberg and W. Kohn. "Inhomogeneous Electron Gas". In: *Phys. Rev* 136.3B (1964), B864.

[116]   James Hooper and Eva Zurek. "High Pressure Potassium Polyhydrides: A Chemical Perspective". In: *The Journal of Physical Chemistry C* 116.24 (2012), pp. 13322–13328. DOI: `10.1021/jp303024h`. eprint: `https://doi.org/10.1021/jp303024h`. URL: `https://doi.org/10.1021/jp303024h`.

[117]   James Hooper and Eva Zurek. "Rubidium polyhydrides under pressure: emergence of the linear H3- species". In: *Chemistry–A European Journal* 18.16 (2012), pp. 5013–5021.

[118]   James Hooper et al. "Polyhydrides of the alkaline earth metals: a look at the extremes under pressure". In: *The Journal of Physical Chemistry C* 117.6 (2013), pp. 2982–2992.

[119]   J. J. Hopfield. "Angular momentum and transition-metal superconductivity". In: *Physical Review* 186.2 (1969), p. 443.

[120]   Pugeng Hou et al. "High pressure structures and superconductivity of AlH 3 (H 2) predicted by first principles". In: *RSC Advances* 5.7 (2015), pp. 5096–5101.

[121]   Pugeng Hou et al. "High pressure structures and superconductivity of AlH3(H2) predicted by first principles". In: *RSC Adv.* 5 (7 2015), pp. 5096–5101. DOI: `10.1039/C4RA14990D`. URL: `http://dx.doi.org/10.1039/C4RA14990D`.

[122]   Ross T Howie et al. "High-pressure synthesis of lithium hydride". In: *Physical Review B* 86.6 (2012), p. 064108.

[123]   Chao-Hao Hu et al. "Pressure-Induced Stabilization and Insulator-Superconductor Transition of BH". In: *Phys. Rev. Lett.* 110 (16 Apr. 2013), p. 165504. DOI: `10.1103/PhysRevLett.110.165504`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.110.165504`.

[124]   Chao-Hao Hu et al. "Pressure-induced stabilization and insulator-superconductor transition of BH". In: *Physical review letters* 110.16 (2013), p. 165504.

[125] J C K Hui and P B Allen. "Effect of lattice anharmonicity on superconductivity". In: *Journal of Physics F: Metal Physics* 4.3 (Mar. 1974), pp. L42–L45. DOI: `10.1088/0305-4608/4/3/003`. URL: `https://doi.org/10.1088/0305-4608/4/3/003`.

[126] Michael Hutcheon. *Quantum Espresso Tools (QET) - a library for carrying out high-throughput calculations using Quantum Espresso.* URL: `https://github.com/miicck/qet`.

[127] Michael J. Hutcheon, Alice M. Shipley, and Richard J. Needs. "Predicting novel superconducting hydrides using machine learning approaches". In: *Phys. Rev. B* 101 (14 Apr. 2020), p. 144505. DOI: `10.1103/PhysRevB.101.144505`. URL: `https://link.aps.org/doi/10.1103/PhysRevB.101.144505`.

[128] A. Yu. Ignatov, S. Y. Savrasov, and T. A. Tyson. "Superconductivity near the vibrational-mode instability in $MgCNi_3$". In: *Phys. Rev. B* 68 (22 Dec. 2003), p. 220504. DOI: `10.1103/PhysRevB.68.220504`. URL: `https://link.aps.org/doi/10.1103/PhysRevB.68.220504`.

[129] R. D. James and K. F. Hane. "Martensitic transformations and shape-memory materials". In: *Acta Materialia* 48.1 (2000), pp. 197–222. ISSN: 1359-6454. DOI: `10.1016/S1359-6454(99)00295-5`. URL: `http://www.sciencedirect.com/science/article/pii/S1359645499002955`.

[130] Robert Jastrow. "Many-Body Problem with Strong Forces". In: *Phys. Rev.* 98 (5 June 1955), pp. 1479–1484. DOI: `10.1103/PhysRev.98.1479`. URL: `https://link.aps.org/doi/10.1103/PhysRev.98.1479`.

[131] Xilian Jin et al. "Superconducting high-pressure phases of disilane". In: *Proceedings of the National Academy of Sciences* 107.22 (2010), pp. 9969–9973. ISSN: 0027-8424. DOI: `10.1073/pnas.1005242107`. eprint: `https://www.pnas.org/content/107/22/9969.full.pdf`. URL: `https://www.pnas.org/content/107/22/9969`.

[132] Xilian Jin et al. "Superconducting high-pressure phases of disilane". In: *Proceedings of the National Academy of Sciences* 107.22 (2010), pp. 9969–9973.

[133] M. H. Kalos. "Monte Carlo Calculations of the Ground State of Three- and Four-Body Nuclei". In: *Phys. Rev.* 128 (4 Nov. 1962), pp. 1791–1795. DOI: `10.1103/PhysRev.128.1791`. URL: `https://link.aps.org/doi/10.1103/PhysRev.128.1791`.

[134] M. H. Kalos and Francesco Pederiva. "Exact Monte Carlo Method for Continuum Fermion Systems". In: *Phys. Rev. Lett.* 85 (17 Oct. 2000), pp. 3547–3551. DOI: `10.1103/PhysRevLett.85.3547`.

[135] M. I. Katsnelson, I. I. Naumov, and A. V. Trefilov. "Singularities of the electronic structure and pre-martensitic anomalies of lattice properties in β-phases of metals and alloys". In: *Phase Transitions* 49.1-3 (1994), pp. 143–191. DOI: 10.1080/01411599408201172.

[136] Duck Young Kim, Ralph H Scheicher, and Rajeev Ahuja. "Predicted high-temperature superconducting state in the hydrogen-dense transition-metal hydride YH_3 at 40 K and 17.7 GPa". In: *Phys. Rev. Lett.* 103.7 (2009), p. 077002.

[137] D. P. Kingma and J. Ba. *Adam: A Method for Stochastic Optimization.* 2014. eprint: arXiv:1412.6980.

[138] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. "Optimization by Simulated Annealing". In: *Science* 220.4598 (1983).

[139] Charles Kittel. *Introduction to Solid State Physics, 8th edition.* Hoboken, NJ: John Wiley & Sons, Inc, 2005, p. 50.

[140] G. Klambauer et al. "Self-Normalizing Neural Networks". In: *Advances in Neural Information Processing Systems (NIPS)* 30 (2017). eprint: arXiv:1706.02515.

[141] Won-Seok Ko and Jong Bae Jeon. "Interatomic potential that describes martensitic phase transformations in pure lithium". In: *Computational Materials Science* 129 (2017), pp. 202–210. ISSN: 0927-0256. DOI: 10.1016/j.commatsci.2016.12.018. URL: http://www.sciencedirect.com/science/article/pii/S0927025616306413.

[142] W. Kohn and L. J. Sham. "Self-Consistent Equations Including Exchange and Correlation Effects". In: *Phys. Rev* 140.4A (1965), A1133.

[143] J. S. Kole and H. De Raedt. "Quantum Monte Carlo method for attractive Coulomb potentials". In: *Phys. Rev. E* 64 (1 June 2001), p. 016704. DOI: 10.1103/PhysRevE.64.016704.

[144] A. N. Kolmogorov et al. "New Superconducting and Semiconducting Fe-B Compounds Predicted with an Ab Initio Evolutionary Search". In: *Physical Review letters* 105.217003 (2010).

[145] P. P Kong et al. "Superconductivity up to 243 K in yttrium hydrides under high pressure". In: *arXiv preprint arXiv:1909.10482* (2019).

[146] PP Kong et al. "Superconductivity up to 243 K in yttrium hydrides under high pressure". In: *arXiv preprint arXiv:1909.10482* (2019).

[147] Vladimir Z Kresin. "High-Tc hydrides: interplay of optical and acoustic modes and comments regarding the upper limit of Tc". In: *Journal of Superconductivity and Novel Magnetism* 31.11 (2018), pp. 3391–3395.

[148] Ivan A Kruglov et al. "Superconductivity of LaH 10 and LaH 16 polyhydrides". In: *Phys. Rev. B* 101.2 (2020), p. 024508.

[149] Ivan A Kruglov et al. "Uranium polyhydrides at moderate pressures: Prediction, synthesis, and expected superconductivity". In: *Science advances* 4.10 (2018), eaat9776.

[150] Alexander A. Kunitsa and So Hirata. "Grid-based diffusion Monte Carlo for fermions without the fixed-node approximation". In: *Physical Review E* 101.1 (Jan. 2020). DOI: `10.1103/PhysRevE.101.013311`.

[151] A. G. Kvashnin et al. "High-Temperature Superconductivity in a Th-H System Under Pressure Conditions". In: *ACS Appl. Mat. & Interfaces* 10.50 (2018), pp. 43809–43816.

[152] A. G. Kvashnin et al. "Iron Superhydrides FeH_5 and FeH_6: Stability, Electronic Properties, and Superconductivity". In: *J. Phys. Chem. C* 122.8 (2018), pp. 4731–4736.

[153] Alexander G Kvashnin et al. "High-temperature superconductivity in a Th–H system under pressure conditions". In: *ACS applied materials & interfaces* 10.50 (2018), pp. 43809–43816.

[154] Yongkyung Kwon, D. M. Ceperley, and Richard M. Martin. "Effects of three-body and backflow correlations in the two-dimensional electron gas". In: *Phys. Rev. B* 48 (16 Oct. 1993), pp. 12037–12046. DOI: `10.1103/PhysRevB.48.12037`. URL: `https://link.aps.org/doi/10.1103/PhysRevB.48.12037`.

[155] Mel Levy. "Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem". In: *Proceedings of the National Academy of Sciences* 76.12 (1979), pp. 6062–6065. ISSN: 0027-8424. DOI: `10.1073/pnas.76.12.6062`. eprint: `https://www.pnas.org/content/76/12/6062.full.pdf`. URL: `https://www.pnas.org/content/76/12/6062`.

[156] Fei Li et al. "Structural evolution of FeH4 under high pressure". In: *RSC Adv.* 7 (21 2017), pp. 12570–12575. DOI: `10.1039/C6RA25591D`. URL: `http://dx.doi.org/10.1039/C6RA25591D`.

[157] Xiao-Feng Li, Zi-Yu Hu, and Bing Huang. "Phase diagram and superconductivity of compressed zirconium hydrides". In: *Phys. Chem. Chem. Phys.* 19 (5 2017), pp. 3538–3543. DOI: 10.1039/C6CP08036G. URL: http://dx.doi.org/10.1039/C6CP08036G.

[158] Xiao-Feng Li, Zi-Yu Hu, and Bing Huang. "Phase diagram and superconductivity of compressed zirconium hydrides". In: *Physical Chemistry Chemical Physics* 19.5 (2017), pp. 3538–3543.

[159] Xiaofeng Li, Hanyu Liu, and Feng Peng. "Crystal structures and superconductivity of technetium hydrides under pressure". In: *Physical Chemistry Chemical Physics* 18.41 (2016), pp. 28791–28796.

[160] Xiaofeng Li and Feng Peng. "Superconductivity of pressure-stabilized vanadium hydrides". In: *Inorganic chemistry* 56.22 (2017), pp. 13759–13765.

[161] Xin Li et al. "Polyhydride CeH_9 with an atomic-like hydrogen clathrate structure". In: *Nature Communications* 10.1 (2019), p. 3461.

[162] Y. Li et al. "Pressure-Stabilized Superconductive Yttrium Hydrides". In: *Scientific Reports* 5 (2015), p. 9948.

[163] Yinwei Li et al. "Superconductivity at ~100 K in dense SiH4(H2)2 predicted by first principles". In: *Proceedings of the National Academy of Sciences* 107.36 (2010), pp. 15708–15711. ISSN: 0027-8424. DOI: 10.1073/pnas.1007354107. eprint: https://www.pnas.org/content/107/36/15708.full.pdf. URL: https://www.pnas.org/content/107/36/15708.

[164] H. Liu et al. "Dynamics and Superconductivity in Compressed Lanthanum Superhydride". In: *Phys. Rev. B* 98.10 (2018), 100102(R).

[165] H. Liu et al. "Potential High-T_c Superconducting Lanthanum and Yttrium Hydrides at High Pressure". In: *Proc. Nat. Acad. Sci.* 114.27 (2017), pp. 6990–6995.

[166] Hanyu Liu et al. "Dynamics and superconductivity in compressed lanthanum superhydride". In: *Physical Review B* 98.10 (2018), p. 100102.

[167] Hanyu Liu et al. "Potential high-Tc superconducting lanthanum and yttrium hydrides at high pressure". In: *Proceedings of the National Academy of Sciences* 114.27 (2017), pp. 6990–6995.

[168] Lu-Lu Liu et al. "High-pressure structures of yttrium hydrides". In: *Journal of Physics: Condensed Matter* 29.32 (July 2017), p. 325401. DOI: 10.1088/1361-648x/aa787d. URL: https://doi.org/10.1088%5C%2F1361-648x%5C%2Faa787d.

[169] Lu-Lu Liu et al. "High-pressure structures of yttrium hydrides". In: *Journal of Physics: Condensed Matter* 29.32 (2017), p. 325401.

[170] Yunxian Liu et al. "Prediction of stoichiometric PoH n compounds: crystal structures and properties". In: *RSC Advances* 5.125 (2015), pp. 103445–103450.

[171] Yunxian Liu et al. "Pressure-Induced Structures and Properties in Indium Hydrides". In: *Inorganic Chemistry* 54.20 (Oct. 2015), pp. 9924–9928. ISSN: 0020-1669. DOI: `10.1021/acs.inorgchem.5b01684`. URL: `https://doi.org/10.1021/acs.inorgchem.5b01684`.

[172] Yunxian Liu et al. "Stability and properties of the Ru–H system at high pressure". In: *Physical Chemistry Chemical Physics* 18.3 (2016), pp. 1516–1520.

[173] Yunxian Liu et al. "Stability and properties of the Ru–H system at high pressure". In: *Phys. Chem. Chem. Phys.* 18 (3 2016), pp. 1516–1520. DOI: `10.1039/C5CP06617D`. URL: `http://dx.doi.org/10.1039/C5CP06617D`.

[174] Yunxian Liu et al. "Structures and properties of osmium hydrides under pressure from first principle calculation". In: *The Journal of Physical Chemistry C* 119.28 (2015), pp. 15905–15911.

[175] J. H. Lloyd-Williams, R. J. Needs, and G. J. Conduit. "Pseudopotential for the electron-electron interaction". In: *Phys. Rev. B* 92 (7 Aug. 2015), p. 075106. DOI: `10.1103/PhysRevB.92.075106`.

[176] Jonathan H. Lloyd-Williams and Bartomeu Monserrat. "Lattice dynamics and electron-phonon coupling using nondiagonal supercells". In: *Phys. Rev. B* 92 (2015), p. 184301. DOI: `10.1103/PhysRevB.92.184301`.

[177] Christian H. Loach and Graeme J. Ackland. "Stacking Characteristics of Close Packed Materials". In: *Phys. Rev. Lett.* 119 (20 Nov. 2017), p. 205701. DOI: `10.1103/PhysRevLett.119.205701`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.119.205701`.

[178] David C. Lonie et al. "Metallization of magnesium polyhydrides under pressure". In: *Phys. Rev. B* 87 (5 Feb. 2013), p. 054107. DOI: `10.1103/PhysRevB.87.054107`. URL: `https://link.aps.org/doi/10.1103/PhysRevB.87.054107`.

[179] David C Lonie et al. "Metallization of magnesium polyhydrides under pressure". In: *Physical Review B* 87.5 (2013), p. 054107.

[180]   David C. Loniea and Eva Zureka. "XtalOpt: An Open–Source Evolutionary Algorithm for Crystal Structure Prediction". In: *Computer physics communications* 182 (2011), pp. 372–387.

[181]   Paul Loubeyre, Florent Occelli, and Paul Dumas. "Synchrotron infrared spectroscopic evidence of the probable transition to metal hydrogen". In: *Nature* 577.7792 (2020), pp. 631–635.

[182]   Yanbin Ma et al. "High-pressure structures and superconductivity of bismuth hydrides". In: *arXiv preprint arXiv:1511.05291* (2015).

[183]   Yanbin Ma et al. "The unexpected binding and superconductivity in SbH4 at high pressure". In: *arXiv preprint arXiv:1506.03889* (2015).

[184]   A. Majumdar et al. "Superconductivity in FeH_5". In: *Phys. Rev. B* 96.20 (2017), p. 201107.

[185]   Rachid Malek and Normand Mousseau. "Dynamics of Lennard-Jones clusters: A characterization of the activation-relaxation technique". In: *Phys. Rev. E* 62 (6 Dec. 2000), pp. 7723–7728. DOI: `10.1103/PhysRevE.62.7723`. URL: `https://link.aps.org/doi/10.1103/PhysRevE.62.7723`.

[186]   Sebastian Manten and Arne Luchow. "On the accuracy of the fixed-node diffusion quantum Monte Carlo method". In: *The Journal of Chemical Physics* 115.12 (2001), pp. 5362–5366. DOI: `10.1063/1.1394757`.

[187]   Ho-Kwang Mao and Russell J Hemley. "Optical studies of hydrogen above 200 Gigapascals: Evidence for metallization by band overlap". In: *Science* 244.4911 (1989), pp. 1462–1465.

[188]   F. Marsiglio and J. P. Carbotte. *Electron - Phonon Superconductivity*. 2001. eprint: `arXiv:cond-mat/0106143`.

[189]   Martin Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015. URL: `https://www.tensorflow.org/`.

[190]   C. M. McCarthy, C. W. Tompson, and S. A. Werner. "Anharmonicity and the low-temperature phase in lithium metal". In: *Phys. Rev. B* 22 (2 July 1980), pp. 574–580. DOI: `10.1103/PhysRevB.22.574`. URL: `https://link.aps.org/doi/10.1103/PhysRevB.22.574`.

[191]   Jeffrey M McMahon et al. "The properties of hydrogen and helium under extreme conditions". In: *Rev. Mod. Phys* 84.4 (2012), p. 1607.

[192]   W. L. McMillan. "Transition temperature of strong-coupled superconductors". In: *Physical Review* 167.2 (1968), p. 331.

[193] V. Meregalli and S. Y. Savrasov. "Electron-phonon coupling and properties of doped BaBiO$_3$". In: *Phys. Rev. B* 57 (22 June 1998), pp. 14453–14469. DOI: 10.1103/PhysRevB.57.14453. URL: https://link.aps.org/doi/10.1103/PhysRevB.57.14453.

[194] Nicholas Metropolis and S. Ulam. "The Monte Carlo Method". In: *Journal of the American Statistical Association* 44.247 (1949). PMID: 18139350, pp. 335–341. DOI: 10.1080/01621459.1949.10483310. eprint: https://www.tandfonline.com/doi/pdf/10.1080/01621459.1949.10483310. URL: https://www.tandfonline.com/doi/abs/10.1080/01621459.1949.10483310.

[195] Yuriy Mishchenko. "Remedy for the fermion sign problem in the diffusion Monte Carlo method for few fermions with antisymmetric diffusion process". In: *Phys. Rev. E* 73 (2 Feb. 2006), p. 026706. DOI: 10.1103/PhysRevE.73.026706.

[196] Lubos Mitas. "Structure of Fermion Nodes and Nodal Cells". In: *Phys. Rev. Lett.* 96 (24 June 2006), p. 240402. DOI: 10.1103/PhysRevLett.96.240402.

[197] Hendrik J. Monkhorst and James D. Pack. "Special points for Brillouin-zone integrations". In: *Phys. Rev. B* 13 (12 June 1976), pp. 5188–5192. DOI: 10.1103/PhysRevB.13.5188. URL: https://link.aps.org/doi/10.1103/PhysRevB.13.5188.

[198] B. Monserrat et al. "Structure and metallicity of phase V of hydrogen". In: *Phys. Rev. Lett* 120.25 (2018), p. 255701.

[199] Bartomeu Monserrat, N. D. Drummond, and R. J. Needs. "Anharmonic vibrational properties in periodic systems: energy, electron-phonon coupling, and stress". In: *Phys. Rev. B* 87 (14 Apr. 2013), p. 144302. DOI: 10.1103/PhysRevB.87.144302. URL: https://link.aps.org/doi/10.1103/PhysRevB.87.144302.

[200] P. Morel and P. W. Anderson. "Calculation of the Superconducting State Parameters with Retarded Electron-Phonon Interaction". In: *Phys. Rev.* 125 (4 Feb. 1962), pp. 1263–1271. DOI: 10.1103/PhysRev.125.1263. URL: https://link.aps.org/doi/10.1103/PhysRev.125.1263.

[201] F. D. Murnaghan. "The Compressibility of Media under Extreme Pressures". In: *Proceedings of the National Academy of Sciences* 30.9 (1944), pp. 244–247. ISSN: 0027-8424. DOI: 10.1073/pnas.30.9.244. eprint: http://www.pnas.org/content/30/9/244.full.pdf. URL: http://www.pnas.org/content/30/9/244.

[202] Jun Nagamatsu et al. "Superconductivity at 39 K in magnesium diboride". In: *Nature* 410.6824 (Mar. 2001), pp. 63–64. ISSN: 1476-4687. DOI: `10.1038/35065039`. URL: `https://doi.org/10.1038/35065039`.

[203] R. J. Needs and C. J. Pickard. "Perspective: Role of Structure Prediction in Materials Discovery and Design". In: *APL Materials* 4.5 (2016), p. 053210.

[204] Richard J. Needs and Chris J. Pickard. "Perspective: Role of structure prediction in materials discovery and design". In: *APL Materials* 4.5 (2016), p. 053210. DOI: `10.1063/1.4949361`.

[205] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. Springer, 2006, p. 136.

[206] A. R. Oganov et al. "Structure Prediction Drives Materials Discovery". In: *Nature Reviews Materials* 4 (2019), p. 331.

[207] Artem R. Oganov and Colin W. Glass. "Crystal structure prediction using ab initio evolutionary techniques: Principles and applications". In: *Chemical Physics* 124.244704 (2006).

[208] D. Ohlendorf and E. Wicke. "Heat capacities between 1.5 and 16 K and superconductivity of V/H and Nb/H alloys". In: *Journal of Physics and Chemistry of Solids* 40.10 (1979), pp. 721–728. ISSN: 0022-3697. DOI: `https://doi.org/10.1016/0022-3697(79)90154-9`. URL: `http://www.sciencedirect.com/science/article/pii/0022369779901549`.

[209] Jeppe Olsen, Poul Jørgensen, and Jack Simons. "Passing the one-billion limit in full configuration-interaction (FCI) calculations". In: *Chemical Physics Letters* 169.6 (1990), pp. 463–472. ISSN: 0009-2614. DOI: `https://doi.org/10.1016/0009-2614(90)85633-N`. URL: `https://www.sciencedirect.com/science/article/pii/000926149085633N`.

[210] R. A. Olsen et al. "Comparison of methods for finding saddle points without knowledge of the final states". In: *The Journal of Chemical Physics* 121.20 (2004), pp. 9776–9792. DOI: `10.1063/1.1809574`.

[211] A. W. Overhauser. "Crystal Structure of Lithium at 4.2 K". In: *Phys. Rev. Lett.* 53 (1 July 1984), pp. 64–65. DOI: `10.1103/PhysRevLett.53.64`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.53.64`.

[212] D. A. Papaconstantopoulos et al. "Cubic H_3S around 200GPa: An Atomic Hydrogen Superconductor Stabilized by Sulfur". In: *Phys. Rev. B* 91.18 (2015), p. 184511.

[213] F. Peng et al. "Hydrogen Clathrate Structures in Rare Earth Hydrides at High Pressures: Possible Route to Room-Temperature Superconductivity". In: *Phys. Rev. Lett.* 119.10 (2017), p. 107001.

[214] Feng Peng et al. "Hydrogen clathrate structures in rare earth hydrides at high pressures: possible route to room-temperature superconductivity". In: *Physical review letters* 119.10 (2017), p. 107001.

[215] J. P. Perdew, K. Burke, and M. Ernzerhof. "Generalized Gradient Approximation Made Simple". In: *Phys. Rev. Lett.* 77.18 (1996), p. 3865.

[216] W. Pichl et al. "The martensite phase of high-purity lithium". In: *J. Phys. IV France* 112 (2003), pp. 1095–1098. DOI: 10.1051/jp4:20031073.

[217] C J Pickard and R J Needs. "High-pressure phases of silane". In: *Phys. Rev. Lett.* 97.4 (2006), p. 045504.

[218] C. J. Pickard, I. Errea, and M. I. Eremets. "Superconducting Hydrides Under Pressure". In: *Annual Review of Condensed Matter Physics* 11 (2020), p. 57.

[219] C. J Pickard and R. J. Needs. "Ab Initio Random Structure Searching". In: *J. Phys: Condens. Matt.* 23.5 (2011), p. 053201.

[220] Chris J. Pickard and R. J. Needs. "Ab initio Random Structure Searching". In: *J. Phys.: Condens. Matter* 23 (2011), p. 053201. DOI: 10.1088/0953-8984/23/5/053201.

[221] Chris J. Pickard and R. J. Needs. "High-Pressure Phases of Silane". In: *Phys. Rev. Lett.* 97 (4 July 2006), p. 045504. DOI: 10.1103/PhysRevLett.97.045504. URL: https://link.aps.org/doi/10.1103/PhysRevLett.97.045504.

[222] W.E. Pickett. "The next breakthrough in phonon-mediated superconductivity". In: *Physica C: Superconductivity* 468.2 (2008). Room Temperature Superconductivity, pp. 126–135. ISSN: 0921-4534. DOI: https://doi.org/10.1016/j.physc.2007.08.018. URL: http://www.sciencedirect.com/science/article/pii/S0921453407013238.

[223] U. Pinsook and G. J. Ackland. "Calculation of anomalous phonons and the hcp-bcc phase transition in zirconium". In: *Phys. Rev. B* 59 (21 June 1999), pp. 13642–13649. DOI: 10.1103/PhysRevB.59.13642.

[224] J.-P. Poirier and A. Tarantola. "A logarithmic equation of state". In: *Physics of the Earth and Planetary Interiors* 109 (Nov. 1998), pp. 1–8. DOI: 10.1016/S0031-9201(98)00112-5.

[225] Mariusz Puchalski and Krzysztof Pachucki. "Ground-state wave function and energy of the lithium atom". In: *Phys. Rev. A* 73 (2 Feb. 2006), p. 022503. DOI: `10.1103/PhysRevA.73.022503`.

[226] Shifeng Qian et al. "Theoretical study of stability and superconductivity of ScH n (n= 4–8) at high pressure". In: *Physical Review B* 96.9 (2017), p. 094513.

[227] Kevin M. Rasch, Shuming Hu, and Lubos Mitas. "Communication: Fixed-node errors in quantum Monte Carlo: Interplay of electron density and node nonlinearities". In: *The Journal of Chemical Physics* 140.4 (2014), p. 041102. DOI: `10.1063/1.4862496`. URL: `https://doi.org/10.1063/1.4862496`.

[228] M.J. Rutter. "C2x: A tool for visualisation and input preparation for CASTEP and other electronic structure codes". In: *Computer Physics Communications* 225 (2018), pp. 174–179. DOI: `https://doi.org/10.1016/j.cpc.2017.12.008`. URL: `http://www.sciencedirect.com/science/article/pii/S0010465517304137`.

[229] J. J. Sakurai and J. Napolitano. "Scattering Theory". In: *Modern Quantum Mechanics*. 2nd. Cambridge University Press, 2017, pp. 386–445. DOI: `10.1017/9781108499996.010`.

[230] Nilesh P Salke et al. "Synthesis of clathrate cerium superhydride CeH 9 at 80-100 GPa with atomic hydrogen sublattice". In: *Nature communications* 10.1 (2019), pp. 1–10.

[231] Nilesh P Salke et al. "Synthesis of clathrate cerium superhydride CeH_9 at 80-100 GPa with atomic hydrogen sublattice". In: *Nature Communications* 10.1 (2019), pp. 1–10.

[232] Antonio Sanna. *Eliashberg's Theory of Superconductivity with ELK*. URL: `https://elk.sourceforge.io/CECAM/Sanna-Eliashberg.pdf`.

[233] Anne Marie Schaeffer et al. "Boundaries for martensitic transition of ^{7}Li under pressure". In: *Nature Communications* 6 (Aug. 2015), p. 8030. DOI: `10.1038/ncomms9030`.

[234] A. Schilling et al. "Superconductivity above 130 K in the Hg–Ba–Ca–Cu–O system". In: *Nature* 363.6424 (May 1993), pp. 56–58. ISSN: 1476-4687. DOI: `10.1038/363056a0`. URL: `https://doi.org/10.1038/363056a0`.

[235] W. Schwarz and O. Blaschko. "Polytype structures of lithium at low temperatures". In: *Phys. Rev. Lett.* 65 (25 Dec. 1990), pp. 3144–3147. DOI: `10.1103/PhysRevLett.65.3144`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.65.3144`.

[236] D. V. Semenok et al. "Actinium Hydrides AcH_{10}, AcH_{12}, and AcH_{16} as High-Temperature Conventional Superconductors". In: *J. Phys. Chem. Lett.* 9.8 (2018), pp. 1920–1926.

[237] Dmitrii V Semenok et al. "Actinium hydrides AcH10, AcH12, and AcH16 as high-temperature conventional superconductors". In: *The journal of physical chemistry letters* 9.8 (2018), pp. 1920–1926.

[238] Dmitrii V Semenok et al. "On distribution of superconductivity in metal hydrides". In: *arXiv preprint arXiv:1806.00865* (2018).

[239] Dmitry V Semenok et al. "Superconductivity at 161 K in thorium hydride ThH10: Synthesis and properties". In: *Materials Today* (2019).

[240] Andrew Shamp, James Hooper, and Eva Zurek. "Compressed cesium polyhydrides: Cs+ sublattices and H3–three-connected nets". In: *Inorganic chemistry* 51.17 (2012), pp. 9333–9342.

[241] Andrew Shamp and Eva Zurek. "Superconducting high-pressure phases composed of hydrogen and iodine". In: *The journal of physical chemistry letters* 6.20 (2015), pp. 4067–4072.

[242] K. V. Shanavas, L. Lindsay, and D. S. Parker. "Electronic structure and electron-phonon coupling in TiH2". In: *Scientific Reports* 6.1 (2016), p. 28102. ISSN: 2045-2322. DOI: 10.1038/srep28102. URL: https://doi.org/10.1038/srep28102.

[243] Alice M. Shipley et al. "Stability and superconductivity of lanthanum and yttrium decahydrides". In: *arXiv:2001.05305* (2020).

[244] James S. Sims and Stanley Hagstrom. "Combined Configuration-Interaction—Hylleraas-Type Wave-Function Study of the Ground State of the Beryllium Atom". In: *Phys. Rev. A* 4 (3 Sept. 1971), pp. 908–916. DOI: 10.1103/PhysRevA.4.908.

[245] T Skoskiewicz et al. "Isotope effect in the superconducting palladium-hydrogen-deuterium system". In: *Journal of Physics C: Solid State Physics* 7.15 (Aug. 1974), pp. 2670–2676. DOI: 10.1088/0022-3719/7/15/015. URL: https://doi.org/10.1088%5C%2F0022-3719%5C%2F7%5C%2F15%5C%2F015.

[246] H. G. Smith. "Martensitic phase transformation of single-crystal lithium from bcc to a 9R-related structure". In: *Phys. Rev. Lett.* 58 (12 Mar. 1987), pp. 1228–1231. DOI: 10.1103/PhysRevLett.58.1228. URL: https://link.aps.org/doi/10.1103/PhysRevLett.58.1228.

[247] H. G. Smith et al. "Pressure effects on the martensitic transformation in metallic lithium". In: *Phys. Rev. B* 41 (2 Jan. 1990), pp. 1231–1234. DOI: 10.1103/PhysRevB.41.1231. URL: https://link.aps.org/doi/10.1103/PhysRevB.41.1231.

[248] Elliot Snider et al. "Room-temperature superconductivity in a carbonaceous sulfur hydride". In: *Nature* 586.7829 (Oct. 2020), pp. 373–377. ISSN: 1476-4687. DOI: 10.1038/s41586-020-2801-z. URL: https://doi.org/10.1038/s41586-020-2801-z.

[249] M. Somayazulu et al. "Evidence for Superconductivity Above 260 K in Lanthanum Superhydride at Megabar Pressures". In: *Phys. Rev. Lett.* 122.2 (2019), p. 027001.

[250] Maddury Somayazulu et al. "Evidence for superconductivity above 260 K in lanthanum superhydride at megabar pressures". In: *Physical review letters* 122.2 (2019), p. 027001.

[251] V. I. Spitsyn et al. "Structure and superconductive properties of the high-pressure phase in the technetium-hydrogen system". In: *Doklady Physical Chemistry* 260.1 (1982), pp. 795–798. ISSN: 0012-5016. URL: http://inis.iaea.org/search/search.aspx?orig_q=RN:14782582.

[252] P. Staikov, A. Kara, and T. S. Rahman. "First-principles studies of the thermodynamic properties of bulk Li". In: *Journal of Physics: Condensed Matter* 9.10 (1997), p. 2135. URL: http://stacks.iop.org/0953-8984/9/i=10/a=004.

[253] Viktor V. Struzhkin et al. "Synthesis of sodium polyhydrides at high pressures". In: *Nature Communications* 7.1 (July 2016), p. 12267. DOI: 10.1038/ncomms12267. URL: https://doi.org/10.1038/ncomms12267.

[254] Wiwittawin Sukmas et al. "Near-room-temperature superconductivity of Mg/Ca substituted metal hexahydride under pressure". In: *Journal of Alloys and Compounds* 849 (2020), p. 156434. ISSN: 0925-8388. DOI: https://doi.org/10.1016/j.jallcom.2020.156434. URL: http://www.sciencedirect.com/science/article/pii/S0925838820327985.

[255] Ying Sun et al. "Route to a superconducting phase above room temperature in electron-doped hydride compounds under high pressure". In: *Phys. Rev. Lett.* 123.9 (2019), p. 097001.

[256] D. Szczesniak and R. Szczesniak. "Thermodynamics of the hydrogen dominant potassium hydride superconductor at high pressure". In: *Solid State Communications* 212 (2015), pp. 1–4. ISSN: 0038-1098. DOI: `https://doi.org/10.1016/j.ssc.2015.03.015`. URL: `http://www.sciencedirect.com/science/article/pii/S0038109815001088`.

[257] R Szczesniak and A P Durajski. "Superconducting state above the boiling point of liquid nitrogen in the GaH3 compound". In: *Superconductor Science and Technology* 27.1 (Nov. 2013), p. 015003. DOI: `10.1088/0953-2048/27/1/015003`. URL: `https://doi.org/10.1088%5C%2F0953-2048%5C%2F27%5C%2F1%5C%2F015003`.

[258] K Tanaka, JS Tse, and H Liu. "Electron-phonon coupling mechanisms for hydrogen-rich metals at high pressure". In: *Physical Review B* 96.10 (2017), p. 100502.

[259] K Tanaka, JS Tse, and H Liu. "Electron-phonon coupling mechanisms for hydrogen-rich metals at high pressure". In: *Physical Review B* 96.10 (2017), p. 100502.

[260] L. R. Testardi. "Structural instability and superconductivity in A-15 compounds". In: *Rev. Mod. Phys.* 47 (3 July 1975), pp. 637–648. DOI: `10.1103/RevModPhys.47.637`. URL: `https://link.aps.org/doi/10.1103/RevModPhys.47.637`.

[261] L. H. Thomas. "The calculation of atomic fields". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 23.5 (1927), pp. 542–548. DOI: `10.1017/S0305004100011683`.

[262] William W. Tipton, Catherine A. Matulis, and Richard G. Hennig. "Ab initio prediction of the Li_5Ge_2 Zintl compound". In: *Computational Materials Science* 93 (2014), pp. 133–136.

[263] Julien Toulouse, Roland Assaraf, and Cyrus J. Umrigar. "Chapter Fifteen - Introduction to the Variational and Diffusion Monte Carlo Methods". In: *Electron Correlation in Molecules - ab initio Beyond Gaussian Quantum Chemistry*. Ed. by Philip E. Hoggan and Telhat Ozdogan. Vol. 73. Advances in Quantum Chemistry. Academic Press, 2016, pp. 285–314. DOI: `10.1016/bs.aiq.2015.07.003`.

[264] H. F Trotter. In: *Proc. Am. Math. Soc.* 10.545 (1959).

[265] Ivan A. Troyan et al. "Anomalous high-temperature superconductivity in YH_6". In: *arXiv preprint arXiv:1908.01534* (2020).

[266] Matthias Troyer and Uwe-Jens Wiese. "Computational Complexity and Fundamental Limitations to Fermionic Quantum Monte Carlo Simulations". In: *Phys. Rev. Lett.* 94 (17 May 2005), p. 170201. DOI: `10.1103/PhysRevLett.94.170201`.

[267] John S. Tse et al. "Structure and electronic properties of BaH2 at high pressure". In: *Solid State Communications* 149.43 (2009), pp. 1944–1946. ISSN: 0038-1098. DOI: `https://doi.org/10.1016/j.ssc.2009.07.044`. URL: `http://www.sciencedirect.com/science/article/pii/S0038109809004669`.

[268] Norm M. Tubman et al. "Prospects for release-node quantum Monte Carlo". In: *The Journal of Chemical Physics* 135.18 (2011), p. 184109. DOI: `10.1063/1.3659143`.

[269] Giovanni A.C. Ummarino. *Eliashberg Theory.* URL: `https://www.cond-mat.de/events/correl13/manuscripts/ummarino.pdf`.

[270] C. J. Umrigar. "Observations on variational and projector Monte Carlo methods". In: *The Journal of Chemical Physics* 143.16 (2015), p. 164105. DOI: `10.1063/1.4933112`.

[271] Pascal Vinet et al. "Temperature effects on the universal equation of state of solids". In: *Phys. Rev. B* 35 (4 Feb. 1987), pp. 1945–1953. DOI: `10.1103/PhysRevB.35.1945`.

[272] David J. Wales and Jonathan P. K. Doye. "Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms". In: *Physical chemistry A* 101 (1997), pp. 5111–5116.

[273] Xiangang Wan et al. "Electron-phonon superconductivity near charge-density-wave instability in $LaO_{0.5}F_{0.5}BiS_2$: Density-functional calculations". In: *Phys. Rev. B* 87 (11 Mar. 2013), p. 115124. DOI: `10.1103/PhysRevB.87.115124`. URL: `https://link.aps.org/doi/10.1103/PhysRevB.87.115124`.

[274] H. Wang et al. "Superconductive Sodalite-like Clathrate Calcium Hydride at High Pressures". In: *Proc. Nat. Acad. Sci.* 109.17 (2012), pp. 6463–6466.

[275] Hui Wang et al. "Superconductive sodalite-like clathrate calcium hydride at high pressures". In: *Proceedings of the National Academy of Sciences* 109.17 (2012), pp. 6463–6466.

[276] Yanchao Wang et al. "Crystal structure prediction via particle-swarm optimization". In: *Physical Review B* 82.094116 (2010).

[277] Yong-Kai Wei et al. "Elastic, superconducting, and thermodynamic properties of the cubic metallic phase of AlH3 via first-principles calculations". In: *Journal of Applied Physics* 114.11 (2013), p. 114905. DOI: 10.1063/1.4821287. eprint: https://doi.org/10.1063/1.4821287. URL: https://doi.org/10.1063/1.4821287.

[278] Yong-Kai Wei et al. "Pressure induced superconductivity and electronic structure properties of scandium hydrides using first principles calculations". In: *RSC Adv.* 6 (85 2016), pp. 81534–81541. DOI: 10.1039/C6RA11862C. URL: http://dx.doi.org/10.1039/C6RA11862C.

[279] Malgorzata Wierzbowska, Stefano de Gironcoli, and Paolo Giannozzi. "Origins of low-and high-pressure discontinuities of T_c in niobium". In: *arXiv:cond-mat/0504077* (2005).

[280] G Wu et al. "Superconductivity at 56 K in samarium-doped SrFeAsF". In: *Journal of Physics: Condensed Matter* 21.14 (Mar. 2009), p. 142203. DOI: 10.1088/0953-8984/21/14/142203. URL: https://doi.org/10.1088/0953-8984/21/14/142203.

[281] Yu Xie et al. "Superconductivity of lithium-doped hydrogen under high pressure". In: *Acta Crystallographica Section C* 70.2 (Feb. 2014), pp. 104–111. DOI: 10.1107/S2053229613028337. URL: https://doi.org/10.1107/S2053229613028337.

[282] Y Xie et al. "Superconductivity of lithium-doped hydrogen under high pressure". In: *Acta Crystallographica Section C: Structural Chemistry* 70.2 (2014), pp. 104–111.

[283] Y Yao et al. "Superconductivity in high-pressure SiH4". In: *Europhysics Letters (EPL)* 78.3 (Apr. 2007), p. 37003. DOI: 10.1209/0295-5075/78/37003. URL: https://doi.org/10.1209%5C%2F0295-5075%5C%2F78%5C%2F37003.

[284] Xiaoqiu Ye et al. "High hydrides of scandium under pressure: potential superconductors". In: *The Journal of Physical Chemistry C* 122.11 (2018), pp. 6298–6309.

[285] Shuyin Yu et al. "Exploration of stable compounds, crystal structures, and superconductivity in the Be-H system". In: *AIP Advances* 4.10 (2014), p. 107118. DOI: 10.1063/1.4898145. eprint: https://doi.org/10.1063/1.4898145. URL: https://doi.org/10.1063/1.4898145.

[286] Shuyin Yu et al. "Exploration of stable compounds, crystal structures, and superconductivity in the Be-H system". In: *AIP Advances* 4.10 (2014), p. 107118.

[287] Shuyin Yu et al. "Pressure-driven formation and stabilization of superconductive chromium hydrides". In: *Scientific Reports* 5.1 (2015), p. 17764. ISSN: 2045-2322. DOI: `10.1038/srep17764`. URL: `https://doi.org/10.1038/srep17764`.

[288] Shuyin Yu et al. "Pressure-driven formation and stabilization of superconductive chromium hydrides". In: *Scientific reports* 5 (2015), p. 17764.

[289] Niloofar Zarifi et al. "Crystal structures and properties of iron hydrides at high pressure". In: *The Journal of Physical Chemistry C* 122.42 (2018), pp. 24262–24269.

[290] Tomasz P Zemła et al. "Characterization of the superconducting phase in tellurium hydride at high pressure". In: *Modern Physics Letters B* (2019), p. 1950169.

[291] Chao Zhang et al. "Chemical Trend of Pressure-Induced Metallization in Alkaline Earth Hydrides". In: *J. Phys. Chem. C* 114.34 (Sept. 2010), pp. 14614–14617. ISSN: 1932-7447. DOI: `10.1021/jp103968c`. URL: `https://doi.org/10.1021/jp103968c`.

[292] Shoutao Zhang et al. "Phase diagram and high-temperature superconductivity of compressed selenium hydrides". In: *Scientific reports* 5 (2015), p. 15433.

[293] Shuang Zheng et al. "Structural and superconducting properties of tungsten hydrides under high pressure". In: *Frontiers in Physics* 6 (2018), p. 101.

[294] Guohua Zhong et al. "Structural, electronic, dynamical, and superconducting properties in dense GeH4 (H2) 2". In: *The Journal of Physical Chemistry C* 116.8 (2012), pp. 5225–5234.

[295] Xin Zhong et al. "Tellurium hydrides at high pressures: High-temperature superconductors". In: *Physical review letters* 116.5 (2016), p. 057002.

[296] Dawei Zhou et al. "Ab initio study revealing a layered structure in hydrogen-rich KH 6 under high pressure". In: *Physical Review B* 86.1 (2012), p. 014118.

[297] Dawei Zhou et al. "Ab initio study revealing a layered structure in hydrogen-rich KH_6 under high pressure". In: *Phys. Rev. B* 86 (1 July 2012), p. 014118. DOI: `10.1103/PhysRevB.86.014118`. URL: `https://link.aps.org/doi/10.1103/PhysRevB.86.014118`.

[298] Di Zhou et al. "High-Pressure Synthesis of Magnetic Neodymium Polyhydrides". In: *Journal of the American Chemical Society* 142.6 (Feb. 2020), pp. 2803–2811. ISSN: 0002-7863. DOI: `10.1021/jacs.9b10439`. URL: `https://doi.org/10.1021/jacs.9b10439`.

[299] Di Zhou et al. "High-Pressure Synthesis of Magnetic Neodymium Polyhydrides". In: *Journal of the American Chemical Society* 142.6 (Feb. 2020), pp. 2803–2811. ISSN: 0002-7863. DOI: 10.1021/jacs.9b10439. URL: https://doi.org/10.1021/jacs.9b10439.

[300] Di Zhou et al. "Superconducting praseodymium superhydrides". In: *Science Advances* 6.9 (2020). DOI: 10.1126/sciadv.aax6849. eprint: https://advances.sciencemag.org/content/6/9/eaax6849.full.pdf. URL: https://advances.sciencemag.org/content/6/9/eaax6849.

[301] Di Zhou et al. "Superconducting praseodymium superhydrides". In: *Science Advances* 6.9 (2020). DOI: 10.1126/sciadv.aax6849. eprint: https://advances.sciencemag.org/content/6/9/eaax6849.full.pdf. URL: https://advances.sciencemag.org/content/6/9/eaax6849.

[302] Quan Zhuang et al. "Pressure-Stabilized Superconductive Ionic Tantalum Hydrides". In: *Inorganic Chemistry* 56.7 (Apr. 2017), pp. 3901–3908. ISSN: 0020-1669. DOI: 10.1021/acs.inorgchem.6b02822. URL: https://doi.org/10.1021/acs.inorgchem.6b02822.

[303] Quan Zhuang et al. "Pressure-stabilized superconductive ionic tantalum hydrides". In: *Inorganic chemistry* 56.7 (2017), pp. 3901–3908.

[304] E. Zurek and T. Bi. "High-temperature superconductivity in alkaline and rare earth polyhydrides at high pressure: A theoretical perspective". In: *J. Chem. Phys.* 150.5 (2019), p. 050901.

[305] Eva Zurek and Tiange Bi. "High-temperature superconductivity in alkaline and rare earth polyhydrides at high pressure: A theoretical perspective". In: *The Journal of chemical physics* 150.5 (2019), p. 050901.

[306] Eva Zurek et al. "A little bit of lithium does a lot for hydrogen". In: *Proceedings of the National Academy of Sciences* 106.42 (2009), pp. 17640–17643. ISSN: 0027-8424. DOI: 10.1073/pnas.0908262106. URL: https://www.pnas.org/content/106/42/17640.