

# Developing a Robust Material Classification Dataset for Application to Machine Learning

George Bird `george.bird@student.manchester.ac.uk`

November 2022

## Abstract

In this project, a large dataset is generated with the intention of training machine learning models to replicate its mappings. The crystal Bismuth-Tellurium-Iodide was selected as it exhibits a wide range of emergent quantum phenomena, including topologically non-trivial states, under specific deformations. Furthermore, the categorisation of these phenomena can be determined analytically. Therefore, this dataset enables systematic comparisons of quantum and classical machine learning techniques to modelling condensed matter systems. Techniques proven using this dataset can then be applied to an expanded set of crystals to streamline research in material development.

To develop this dataset required efficient and robust methods of classifying Bismuth-Tellurium-Iodide states such that many samples can be generated. To achieve this required novel implementation of condensed matter models which can be computed in parallel using GPU-acceleration to generate many samples. Furthermore, we used a novel application of the gradient descent algorithm to determine key features in the material's dispersion relations which aid in this classification process. Several methods of enlarging the dataset are also presented, such that the machine learning approach can better model the system. The condensed matter models and assumptions used to develop the procedures are derived and discussed, alongside the characteristic behaviours of several topologically trivial and non-trivial material states. The resultant classification procedures are shown to be successful and offer insight into the studied materials.

# Contents

<b>1</b>	<b>Introduction: Quantum Mechanical Analysis of Crystals</b>	<b>3</b>
1.1	Many-body Schrödinger Equation . . . . .	3
1.2	Crystal Lattices . . . . .	3
1.3	Bloch Wavefunctions . . . . .	4
1.4	Condensed Matter Models and Assumptions . . . . .	4
1.4.1	Born-Oppenheimer Approximation . . . . .	4
1.4.2	Linear Combination of Atomic Orbitals . . . . .	5
1.4.3	Nth Nearest Neighbour Cutoff Assumption . . . . .	5
1.4.4	Tight-Binding Model . . . . .	6
<b>2</b>	<b>Classification of Materials</b>	<b>6</b>
2.1	Classical Materials . . . . .	7
2.1.1	Conductors . . . . .	7
2.1.2	(Trivial) Insulators . . . . .	7
2.1.3	Semiconductors . . . . .	7
2.1.4	(Weyl & Dirac) Semimetals . . . . .	8
2.2	Overview of Topology and its Connection to Material Classification . . . . .	9
2.3	Topological Insulators . . . . .	9
2.3.1	Rashba Effect . . . . .	10
<b>3</b>	<b>Methodology and Results of Classifying Topological States for Application to Machine Learning</b>	<b>11</b>
3.1	Novel Approach to Highly Parallelised Tight-Binding Equation Solvers . . . . .	12
3.2	Material Classification Procedure . . . . .	14
3.2.1	Determining Minimum Bandgap by Grid Search and Random Sampling . . . . .	14
3.2.2	Gradient descent - A Novel Approach for Determining the Bandgap Minima . . . . .	15
3.3	Dataset Expansion . . . . .	16
3.3.1	Interpolation . . . . .	16
3.3.2	Random Noise . . . . .	16
3.4	Energy Eigenvalue Analysis . . . . .	18
3.5	Eigenvector Analysis and Spin Behaviour . . . . .	19
3.6	Analysis of Spin-Orbit Terms . . . . .	23
<b>4</b>	<b>Conclusion</b>	<b>23</b>
4.1	Outlook of Work . . . . .	24
4.1.1	Benchmarking Supervised Learning Models . . . . .	24
4.1.2	Quantum Machine Learning Approach . . . . .	24
4.1.3	Generalising for Many Crystalline Materials . . . . .	24

# 1 Introduction: Quantum Mechanical Analysis of Crystals

## 1.1 Many-body Schrödinger Equation

Arguably the single most insightful and defining equation of modern physics is the Schrödinger equation, shown in its time-dependent form in *Eqn. 1*. It describes the nature, evolution and interaction of waves defining matter. The equation follows from the Hamiltonian approach to classical mechanics, where classical terms are replaced with quantum operators. There are a multitude of cases for which the Schrödinger equation can be solved analytically, such as atomic orbitals.

$$\hat{H}\Psi(\vec{x}, t) = \frac{-\hbar^2}{2m} \nabla^2 \Psi(\vec{x}, t) + V(\vec{x}, t) \Psi(\vec{x}, t) = -i\hbar \frac{\partial}{\partial t} \Psi(\vec{x}, t) \quad (1)$$

However, analytical solutions of the schrödinger equation become intractable in many-body scenarios, due to the growing number of pair-wise interactions [1]. The many nuclei and electron interactions involved in macroscopic materials make their analytical analysis prohibitive to compute. This is demonstrated in *Eqn. 2* where the pairwise-interactions needing consideration grows quadratically with the number of particles, alongside a linear growth in the dimension of the hilbert space. Thus, several simplifying assumptions have been developed to yield insight. Particularly the tight-binding equation is utilised in this project, as it captures the quantum behaviour of interest whilst having reasonable computational cost. Therefore, this model, its associated assumptions and prerequisite models are outlined in this introduction.

$$\hat{H} = \underbrace{\sum_{i=0}^{N_{\text{nuc.}}} \frac{-\hbar^2}{2m_i} \nabla_i^2}_{\text{Nuclei Kinetic Term}} + \underbrace{\sum_{j=0}^{N_{\text{elec.}}} \frac{-\hbar^2}{2m_j} \nabla_j^2}_{\text{Electron Kinetic Term}} + \underbrace{\sum_{k=0}^{N_{\text{nuc.}}} \sum_{l=0}^{k-1} \frac{Z_k Z_l e^2}{4\pi\epsilon_0 |\vec{x}_k - \vec{x}_l|}}_{\text{Nuclei-Nuclei Interaction}} + \underbrace{\sum_{p=0}^{N_{\text{nuc.}}} \sum_{q=0}^{N_{\text{elec.}}} \frac{-Z_p e^2}{4\pi\epsilon_0 |\vec{x}_p - \vec{x}_q|}}_{\text{Nuclei-Electron Interaction}} + \underbrace{\sum_{m=0}^{N_{\text{elec.}}} \sum_{n=0}^{m-1} \frac{e^2}{4\pi\epsilon_0 |\vec{x}_m - \vec{x}_n|}}_{\text{Electron-Electron Interaction}} \quad (2)$$

These streamlined models have shown quantum mechanics to be an effective descriptor for many of the characteristics of bulk materials [2]. A variety of material classifications can be understood through the resultant quantum dispersion relations, detailed in *Sec. 2.1*. Moreover, some materials classifications are realised uniquely through this approach, including topological insulators which are emergent quantum phenomena. The identification of these emergent materials form the basis of this project. To begin analysing material properties requires a schema to represent the materials of interest, this is discussed in the following section.

## 1.2 Crystal Lattices

The most elementary material is a perfect crystal, without impurities or defects, as any additional disorder adds complexity. These crystalline materials are described through the tessellation of a base object, a unit cell, depicted in *Fig. 1*. Unit-cells are characterised by their constituent atoms and their relative positions. These cells are offset by integer multiples of lattice-vectors described in *Eqn. 3*. This formulation of the unit-cell and its associated lattice-vectors is the minimal, complete description for these materials.

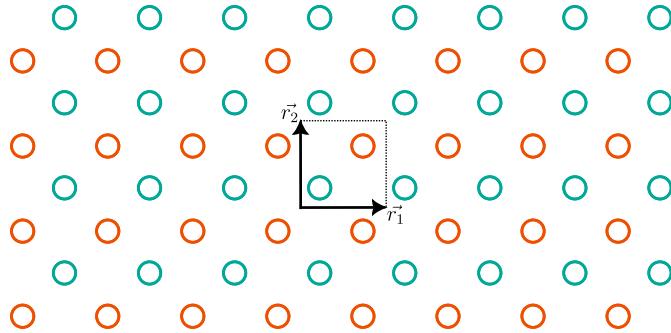


Figure 1: Depicts how a unit cell, outlined with a dashed box and containing two atoms of orange and blue type, is tiled by integer multiples of the lattice vectors  $\vec{a}_1$  and  $\vec{a}_2$  to produce the crystal.

$$\vec{R} = \sum_{i=0}^m n_i \vec{r}_i \quad (3)$$

Where  $n_i \in \mathbb{Z}$  giving each of the integer factors for the offsets, and  $\vec{r}_i \in \mathbb{R}^m$  are each of the lattice vectors. These vectors should be assumed to be mutually independent and span an  $m$  dimensional space. It is further assumed that the crystal is infinite in extent or abides by some periodic boundary condition.

As a consequence of tessellation, perfect crystalline materials feature at least discrete translational symmetry. If electron-electron interactions are also neglected, which would otherwise develop the problem into a many-body dynamic with its associated complexity, then a periodic potential arises that incurs constraints on the wavefunction. Thus, the wavefunction must also exhibit the same discrete translational symmetry which is described by the Bloch-theorem.

### 1.3 Bloch Wavefunctions

The coordinate system's origin can be arbitrarily shifted by the vector in *Eqn. 3*. Consequently, the local environment experienced by the electron is independent of which cell is defined as the centre. Therefore, it is expected that observable properties of the wavefunction vary within the unit cell, but repeat with the same discrete translations, resulting in periodic observable properties of the wavefunction. Complex, global phase of the wavefunction is unobservable, so it can evolve through the crystal without being constrained by the symmetry. This leads to the Bloch wavefunction [3] shown in *Eqn. 4*.

$$\Psi(\vec{x}) = e^{i\vec{k} \cdot \vec{x}} u(\vec{x}) \cap u(\vec{x} + \vec{R}) = u(\vec{x}) \quad (4)$$

The function  $u(\vec{x})$  results from the specific periodic potential involved, modulating the plane-wave solution. A position dependent phase, characterised by the quantum number  $\vec{k}$ , enters through the plane-wave term. This parameter can be associated to the crystal-momentum, in analogy to the standard plane wave [4]. Thus Bloch's theorem indicates that electrons are distributed as modulated waves, with various crystal momenta, within the crystalline material.

### 1.4 Condensed Matter Models and Assumptions

As mentioned, a range of simplifying assumptions to the many-body Schrödinger description are employed to make it computationally feasible to model wavefunctions within a material. The aim is to capture all the dynamics of interest, contributing to the emergent material properties, whilst jettisoning any redundant information. This reduction in the degrees of freedom results in a smaller search-space for solutions and hence a quantum description becomes computationally attainable. A broad range of techniques have been developed to express different emergent features of a material. In this work the tight-binding model is extensively used. A general progression of models up to the tight-binding model, alongside contributing assumptions, are detailed below.

#### 1.4.1 Born-Oppenheimer Approximation

A large relative mass difference between the nuclei and electron makes it practical to model their dynamics separately, since their time-evolution's occur at vastly differing rates, as seen in *Eqn. 1*. This separable approach is known as the Born-Oppenheimer approximation [5] and it has wide application in physics and chemistry. In *Eqn. 5*, the many-body Schrödinger equation is separated into a nuclear and electronic hamiltonian, which act on their respective separated wavefunctions.

$$\hat{H}\Psi(\{\vec{r}_i\}, \{\vec{r}_j\}) \approx (\hat{H}_{\text{nucl}} + \hat{H}_{\text{elec}}) \psi(\{\vec{r}_i\}) \phi(\{\vec{r}_j\}) \approx \hat{H}_{\text{nucl}}\psi(\{\vec{r}_i\}) + \hat{H}_{\text{elec}}\phi(\{\vec{r}_j\}) \quad (5)$$

The semi-independent modelling of both these systems has given insight into material behaviour. For example, the displacement of the nuclei from their equilibrium position describes the mechanism of both heat [6] and sound transmittance [7, 8] throughout a crystalline solid. These quantised vibrational modes give rise to the phonon quasiparticle, a type of emergent boson that can distribute energy throughout the crystal [9]. Due to a large time-scale difference between the two, it is typical to neglect the nuclei wavefunction entirely in an extension to the Born-Oppenheimer approximation. This is

an appropriate measure to uphold the crystal's discrete translational symmetry. In addition, it is often the general time-independent electronic states which are of interest, so only time-averaged nuclear wavefunctions are sufficient to provide solutions.

However, this simplification can lead to deficits in a material's electronic description. Despite the differing time-scales for motion, the separated systems have been shown to have important interplay producing emergent quantum behaviour [10–12]. A further composite boson quasiparticle, the cooper pair [13], can arise from many electron-phonon interactions within a crystal. The condensation of these bound states has been demonstrated to be an effective description for superconductivity [14]. Thus, some emergent phenomena cannot be modelled from analysis of the electronic system alone.

However, with just the extended Born-Oppenheimer approximation, the degrees of freedom representing the electronic system remain in excess of what is needed to explain many of the observed material properties. With just this assumption the modelling remains inviable. The transition away from the position basis to describe wavefunctions, reduces the number of degrees of freedom. In the following section the linear combination of atomic orbitals is used to approximately represent the wavefunction, further simplifying the many-body dynamics.

#### 1.4.2 Linear Combination of Atomic Orbitals

Interactions between electrons and neighbouring nuclei tend to zero as the nuclei are infinitely separated. This results in the electronic wavefunctions tending towards their independent atomic wavefunction solutions. At finite separation these atomic wavefunctions are increasingly perturbed. However, if this process is considered insignificant, the resultant crystal wavefunction can be represented using a linear combination of these atomic orbitals [15]. This assumption offers a greatly simplified, approximate basis for solving the wavefunctions of a crystal lattice. This superposition of atomic orbitals, to describe the electron wavefunction, is shown in *Eqn. 6*.

$$\Psi_n(\vec{x}) = \sum_{\vec{R}} \sum_a C_{n,a,\vec{R}} \psi_a(\vec{x} - \vec{R}) \quad (6)$$

Individual normalised atomic orbitals are given by  $\psi_a$  centred on each of their respective nuclei. These orbitals can also include electron spin to account for the pauli-exclusion principle [16]. Each of these orbitals are then displaced by vector  $\vec{R}$ , producing the full extent of the crystal. The resultant electron wavefunction  $\Psi$  should also be normalised to one. Therefore, *Eqn. 6* can be used to describe the single electron wavefunction in a crystal. The many-body schrödinger equation in the position basis can be represented in the atomic basis, for a single electron, as shown in *Eqn. 7*. Thus, the hamiltonian is represented as a matrix, enabling the matrix-mechanics approach. Where  $i$  and  $j$  go through all atomic orbital permutations from *Eqn. 6*. This equation can interpreted as a stochastic matrix for continuous-time, discrete-space markovian dynamics.

$$\hat{H}_{ij\vec{R}} = \langle \psi_i | \hat{H} | \psi_j \rangle = \int d\vec{x} \psi_i^*(\vec{x}) \hat{H}(\vec{x}) \psi_j(\vec{x} - \vec{R}) \quad (7)$$

The elements can be determine using Hartree-Fock method [17, 18] or density-functional theory [19]. These techniques address the electron-electron interaction terms using an electron mean-field, this decouples the individual electron in the Hamiltonian allowing each nuclei and electron equation to be approximated separately. Finally, the energy eigenvalues and mixing coefficients  $C_{a,\vec{R}} \in \mathbb{C}$  are deduced such that they fulfill the time-independent schrödinger relation  $\hat{H}\Psi_n \approx E_n\Psi_n$ .

It is apparent from this description that when superposing the wavefunctions, the corresponding energy eigenvalues can combine to form a wider range of energies than the original quantised ones from the atomic orbitals. This begins a trend of quasi-continuous energy bands for large crystals. To ascertain better understanding of crystal properties, a dependence between the energy value and crystal-momenta must be found. To achieve this requires the Tight-Binding equation and a truncation to the atomic orbitals being considered.

#### 1.4.3 Nth Nearest Neighbour Cutoff Assumption

For a given infinitely extending crystal, a finite number of  $\vec{R}$  are sampled. These are chosen in order of proximity ( $|\vec{R}|$ ). It is expected that long-range interactions become negligible. Thus, a cutoff number is chosen for how many neighbouring cells

are evaluated. The distance cutoff used depends on the crystal being studied. The term "n<sup>th</sup>-next nearest neighbours" is used to describe an equidistant set of cells from the central cell.

#### 1.4.4 Tight-Binding Model

The Tight-Binding model [20] can be interpreted as the application of Bloch-wavefunctions to the linear combination of atomic orbitals (LCAO), using the Born-Oppenheimer and  $N^{th}$  nearest neighbour cutoff assumptions. The term tight-binding arises from the atomic orbital basis used, as tightly bound electrons will have minimally perturbed atomic orbitals as described in the LCAO technique. The tight-binding technique can be viewed as a further simplification to LCAO method, by utilising the discrete translational symmetry to relate displaced atomic orbitals through the Bloch-theorem.

In *Eqn. 7*, each of the diagonal terms ( $i = j \cap \vec{R} = \vec{0}$ ) represents the energy eigenvalues for the atomic orbitals, which take on their usual discrete values. Meanwhile all other terms represent the likelihood of an electron hopping from one orbital to another. These hopping terms represent the bond-energies between neighbouring orbitals. Now orbitals displaced by  $\vec{R}$  must be related using *Eqn. 4*. This requires a contraction of  $H_{ij\vec{R}}$  with the phase acquired from the Bloch theorem ( $\phi_{\vec{R}} = e^{i\vec{k}\cdot\vec{R}}$ ). The resultant expression is shown in *Eqn 8*.

$$\hat{H}(\vec{k}) = \hat{H}_{ij}(\vec{k}) = \sum_{\forall \vec{R}} e^{i\vec{k}\cdot\vec{R}} \hat{H}_{ij\vec{R}} \quad (8)$$

This hamiltonian is trivially Hermitian if the n-nearest neighbour approach is used for sampling  $\vec{R}$ . Using the time-independent schrödinger equation, shown in *Eqn. 1*, it can be seen that the energy-band structure is related to the crystal-momenta. This relation is demonstrated in *Eqn. 9*.

$$\hat{H}(\vec{k}) \Psi_{n\vec{k}} = E_n(\vec{k}) \Psi_{n\vec{k}} \quad (9)$$

Where  $\Psi_{n\vec{k}}$  represents a vector of mixing coefficients for the various atomic orbitals. These coefficients, similar to those in *Eqn. 6*, and energy eigenvalues can be found by diagonalisation of the hamiltonian  $H_{ij}$ . These represent stationary states for the electrons in the crystal, so called "Bloch-electrons" characterised by the crystal momenta  $\vec{k}$  describing their phase evolution through the crystal. In addition, for each wavevector  $\vec{k}$ , the Bloch-electron can also exist in several excited states, indicated by the several energy eigenvalues, these correspond to various complex hybridisations of the atomic orbitals. Since the energy is a function of the crystal momenta, and multiple excited states exist for each momenta, then continuous energy bands form. A direct consequence of the periodicity in position, a discrete translational symmetry is also exhibited in the wavevectors  $\vec{k}$ , seen in the phase factor of *Eqn. 8*, this results in the "reciprocal lattice". Reciprocal-lattice vectors ( $\vec{G}$ ) characterises this discrete translational symmetry as demonstrated in *Eqn. 10*.

$$e^{2\pi i \vec{G}\cdot\vec{R}} = 1 \quad (10)$$

The individual vectors  $b_j$  are combined to make  $\vec{G} = \sum_{j=0}^m p_j \vec{b}_j$  where  $p_j \in \mathbb{Z}$ , analogous to *Eqn. 3*. These are related by  $\vec{r}_i \cdot \vec{b}_j = 2\pi\delta_{ij}$ . Thus, the full electron dynamics, up to the given assumptions, are summarised in a finite region called the Brillouin zone [21]. This Brillouin zone is tiled by the reciprocal lattice vectors to make the full  $\vec{k}$ -space. As the Tight-binding equation is a bijective map, there is a unique correspondence between the unit-cell for the real space crystal being studied and the Brillouin-zone's energy-momentum relations. Consequently, all arrangements of crystals have unique Brillouin zones and therefore unique energy band structures and orbital hybridisations. Since these characterise material properties, it is clear that every type of crystal has unique behaviours. However, these material properties can be classified under general categories, discussed in the next section.

## 2 Classification of Materials

Using the energy band structure, it is possible to group many materials into distinct categories based on their overall electronic behaviour. Historically, many of these general classifications existed prior to discovery of the energy band structure for

electrons and instead relied upon experimental observations. However, the electronic band structure was found to effectively describe the various material properties [2] and offered a rigorous criteria to classify them. In addition, it created a theoretical framework to discover more exotic emergent quantum properties within the preexisting classifications leading to further subdivisions being defined.

## 2.1 Classical Materials

For any given material, the constituent atoms will possess a specific number of electrons. At absolute zero, these electrons occupy the wavefunctions corresponding to the minimum possible energy for the material. This leads to a distinct cutoff, known as the fermi-energy ( $E_f$ ), where available states above this energy value are unoccupied and states below are fully occupied. At temperatures higher than absolute zero, thermal fluctuations can excite a number of electrons into higher energy states, through phonon-electron interactions. These excited electrons tend to originate from around the fermi-energy since less energy is needed to excite them. After excitation, unoccupied states can remain which may act as quasiparticles, for example a "hole" quasiparticle has analogous properties to a positron [22, 23]. The allowed excitations of electrons and properties of any resultant quasiparticles directly influence material properties and are determined by the energy-momentum defined by the band structure. In the following sections a brief overview is given on how to classify common materials from their band structure.

### 2.1.1 Conductors

In conductors the fermi-energy lies within the continuous energy-band. As a consequence, electrons at the fermi-energy have no threshold of energy required to excite them and this leads to a mix of occupied and unoccupied states around the fermi-energy. As a result, due to the presence of many unoccupied states, electrons can easily redistribute even under the application of weaker electric fields. This can incur a bias in which crystal-momenta states are occupied and this leads to a net total momenta of the electrons. Consequently, electric current is easily produced under an applied electric field, characterising conductors.

In addition, the gradient of the band structure is proportional to the group velocity of the Bloch-electrons as shown in *Eqn. 11*. This group-velocity indicates the transport properties of electrons which affects electrical conductivity ( $\sigma$ ). So, if the fermi-energy lies in a steep region of the energy-momenta relation the electrical conductivity of the material is greater.

$$\vec{v} = \hbar^{-1} \nabla_{\vec{k}} E(\vec{k}) \quad (11)$$

### 2.1.2 (Trivial) Insulators

Insulators have their fermi-energy outside the energy-band. In this case, the fermi-energy lies in a region absent of allowed energy states, known as an "energy bandgap". As a result a substantial threshold energy is required to excite an electron. This results in the electrons occupying all available groundstate, filling the energy bands below the fermi-energy. Hence, insulators are poor at conducting electricity as very large electric fields are required to setup a net momentum bias. Dielectric breakdown gives an indication of how large this electric field must be to make the material conductive.

### 2.1.3 Semiconductors

Semiconductors can be interpreted as the middle ground between conductors and insulators. Like insulators, the fermi-energy is situated within a bandgap but this bandgap is substantially smaller. Thus a threshold energy exists, but is not insurmountable at room-temperatures, resulting in some electrical conductivity. Defining semiconductors explicitly from their bandgap can be somewhat arbitrary depending on the degree of splitting in the energy bands. Therefore, semiconductors can instead be characterised by the ability to dope the material with impurities, altering the fermi-energy.

The doping of semiconductors is one of the greatest applications of condensed matter physics, enabling the wide variety of electronics [24]. Introducing impurities controls conductivity by either increases the number of unoccupied states below the fermi-energy (p-type) or increases the number of electrons (n-type).

Unlike conductors, the group-velocity is smaller in the fermi-energy region due to flatter band structure. This results in poorer conductivity of semiconductors. However, the energy-bands close to the fermi-energy are also approximately parabolic. Consequently, the energy-momentum relation is similar to that of classical massive particles, though it can be anisotropic, as shown in *Eqn. 12*.  $C$  is a constant added to reach the top or bottom of the band.

$$E_{\text{total}}(\vec{k}) = C + \sum_i \frac{\hbar^2 k_i^2}{2m_i} \quad (12)$$

Extending this principle outside of semiconductors, it can be found that many band structures, near the fermi-energy, are better approximated using the relativistic energy [25–27] shown in *Eqn. 13*, with fitting constants  $M$  and the fermi-velocity  $v_{F_i}$ . This indicates that the electrons and holes can travel through the crystal in a way analogous to relativistic particles.

$$E_{\text{total}}(\vec{k}) = \pm \sqrt{M^2 + \hbar^2 \sum_i v_{F_i}^2 k_i^2} \quad (13)$$

In the next section, some semimetals are shown to have band structures which approximate the dispersion relation of ultra-relativistic particles.

#### 2.1.4 (Weyl & Dirac) Semimetals

Semimetals occur when the energy bandgap closes and the respective bands may begin to overlap at the fermi-energy [28]. This can be used as a classification criteria if a material's bandgap reduces to zero under deformation. Semimetals are conductive since there is no energy-threshold for excitation. Furthermore, both the electrons and quasiparticles are responsible for their conductive properties. Particularly, weyl-semimetals and dirac-semimetals [29] are sub-classifications which we will focus on in this description, since these semimetal states arise in our crystal of study: Bismuth-Tellurium-Iodide (BiTeI).

These semimetals are characterised by a linear dispersion relation about the fermi-energy, appearing as dirac-cones [25, 30] in the Brillouin zone. As a consequence, quasiparticles emerge which can be described by the dirac equation, shown in *Eqn. 14*. The energy dispersion of these solutions corresponds to *Eqn. 13* where  $M$  is the quasiparticle's rest energy. However, these quasiparticles are massless if dirac-cones are present in the Brillouin zone. Specifically, at least two possible solutions are exhibited which define the distinction between dirac-semimetals and weyl-semimetals.

$$(i\hbar\gamma^\mu\partial_\mu - m_{\text{eff}}v_D)\Psi(x) = 0 \quad (14)$$

Dirac-semimetals require time-reversal (shown in *Eqn. 15*) and spatial-inversion (shown in *Eqn. 16*) symmetry to be upheld in the crystal. As a consequence of this, the spin-up and spin-down states are degenerate. Thus, any crossing of previously distinct bands, defining a semimetal state, actually includes the crossing of four bands due to the spin-degeneracy. Therefore, the quasiparticles which form are massless dirac fermions.

$$\hat{T}\Psi_{\vec{k},\uparrow} = \Psi_{-\vec{k},\downarrow} = \Psi_{\vec{k},\uparrow} \quad (15)$$

$$\hat{I}\Psi_{\vec{k},\uparrow(\downarrow)} = \Psi_{-\vec{k},\uparrow(\downarrow)} = \Psi_{\vec{k},\uparrow(\downarrow)} \quad (16)$$

Weyl-semimetals occur when either of these symmetries are broken. Spatial-inversion symmetry is broken via spin-orbit interactions, such as the Rashba effect discussed in *Sec. 2.3.1*, resulting in non-degenerate spin states. The presence of spin-orbit interaction is used as a method of distinguishing the semimetal states of BiTeI and is discussed further in *Sec. 3.6*. Due to the splitting of the spin states, the previous dirac-cone involving four band crossings, divides into two weyl-cones with just two band crossings at each. The respective weyl-cones are crossings of exclusively spin-up or spin-down bands. Consequently, emergent quasiparticles are formed known as weyl-monopoles which are a previously unobserved solution to the Dirac equation [31–33]. These pairs of monopoles and antimonopoles can only be created and annihilated along specific symmetry lines of the crystal, thus are protected states and migrate through the crystal symmetrically according to time-reversal symmetry. The monopoles carry the resultant charge from a local  $U(1)$  gauge invariance of the quantum state [34, 35]

shown in *Eqn.* 17, where  $|u(\vec{k})\rangle$  is described in *Eqn.* 4.

$$|u'(\vec{k})\rangle = e^{i\phi(\vec{k})} |u(\vec{k})\rangle \quad (17)$$

Since this is a  $U(1)$  gauge symmetry, a vector potential can be defined as shown in *Eqn.* 18.

$$\vec{A}(\vec{k}) = -i\langle u(\vec{k}) | \nabla_{\vec{k}} | u(\vec{k}) \rangle \quad (18)$$

The curl ( $\vec{B}(\vec{k}) = \nabla \times \vec{A}(\vec{k})$ ), known as berry connection can be integrated across a closed surface to produce the geometric phase [36], also known as the berry phase [37]. This has an associated charge, shown in *Eqn.* 19, which is the charge of the monopoles.

$$\varrho(\vec{k}) = \frac{\nabla_{\vec{k}} \cdot \vec{B}(\vec{k})}{2\pi} = \sum_i q_i \delta(\vec{k} - \vec{k}_i) \quad (19)$$

Where  $q_i = \pm 1$ , whilst  $\delta(\vec{k})$  represents the dirac-delta function. These dirac delta functions are centred on the band crossing points [38] ( $\vec{k}_i$ ) as expected. These weyl-monopoles have either left-handed ( $q_i = -1$ ) or right-handed ( $q_i = +1$ ) chiralities, dependent on the weyl-cone being considered [39]. These weyl-monopoles are more stable, as they do not rely on the symmetries involved in the dirac-semimetals. Due to the spin-orbit splitting, the opposite chiral-charge monopoles are separated in momentum-space by the position of the weyl-cones but join to form characteristic fermi-arcs at the surface [40–42]. These weyl-monopoles are effective charge carriers behaving as massless, ultra-relativistic and therefore highly mobile particles. These can find applications requiring faster conduction in electronics, have many applications in spintronics and can also elucidate the nature of the Higgs mechanism [43]. Weyl-semimetals are considered topologically non-trivial [44] and this premise of analysing materials topologically allows for new classifications of distinct material states. These states exhibit unique, emergent quantum phenomena which will have important and widespread technological applications.

## 2.2 Overview of Topology and its Connection to Material Classification

Topology is the study of general classifications of shapes, which can be homeomorphically deformed into one another. Groups of homeomorphic shapes are characterised by constant topological numbers including genus, orientability and torsion. In recent years, there has been a paradigm shift to use classifications developed in topology and apply it to categorising condensed matter systems. Whereas the energy band structure is usually indicative of a material's classifications, topological insulators are instead distinguished through topological numbers [45–47]

The  $\mathbb{Z}_2$  topological invariant can be applied to condensed matter systems with topological numbers  $\nu_i$ . These represent the number of Kramer pairs [35, 48], which are pairs of degenerate energy states which are the time-reversed states of one another. The possible values of  $\nu_i$  are 0 (trivial) or 1 (topological), therefore a smooth transition between these integers is not possible, resulting in distinct topological categorisations. There are four topological numbers associated with  $\mathbb{Z}_2$ , namely  $\nu_0; (\nu_1, \nu_2, \nu_3)$ . Where  $\nu_0$  is the strong topological index which is resistant to impurities, specially non-magnetic impurities as magnetic fields break the time-reversal symmetry. Whilst  $\nu_{1,2,3}$  are the weak topological indices and are not resilient nor significant in impure crystals.

The  $\mathbb{Z}_2$  topological invariants cannot be changed when a bandgap exists [38]. Consequently, they can be used to further subdivide the previous states using topological categorisation. These topologically non-trivial states arise in insulators, producing topological-insulators with unique properties alongside the aforementioned weyl semimetals. The following section will overview this emergent quantum material classification.

## 2.3 Topological Insulators

Topological insulators are a further sub-classification of insulating materials. Despite displaying similar behaviours to trivial insulators, such as no conductivity in their bulk due to a bandgap, they are distinctly different with at least one non-zero  $\mathbb{Z}_2$  topological number. Therefore, it is not possible to adiabatically transform, analogous to topology's homeomorphic transform, from a trivial to a topological insulator without first closing the bandgap to change the topological index.

The presence of the aforementioned monopoles and antimonopoles prevent the opening of such a bandgap, permitting the topological invariant to change. Subsequently, a semimetallic state [49] is formed between the transition of a trivial insulator to a topological insulator. This separates the two insulating states in phase space, indicating their distinctness. Therefore, the general category of insulating states can be further subdivided. These new categorisations are a result of topological analysis of condensed matter systems.

During the intermediate semimetal stage a rearrangement of the orbital and spin coefficients occur resulting in the topological insulator being a conglomerate of the previously dissimilar bands when the bandgap is formed again. This is known as band inversion [29] and represents one mode of classifying the insulating types, discussed further in *Sec. 3.5*.

At the surface of the material a transition must also occur from a topological to trivial insulator. This is because the vacuum is a trivial insulator, since there exists a bandgap of the minimum energy for particle pair-production. Hence, between the bulk of the topological insulator and the vacuum, the bandgap must close. A resultant semimetal-like gapless state, with linear dispersion, must form. This results in all topological insulators having a conducting surface, with highly mobile and spin-locked electrons, which are topologically protected under deformation. These electrons are also not easily scattered by impurities [50]. These features of topological insulators have a wealth of applications such as in pulsed laser technology [51], spintronics [52] and topological Qubits with greater stability [52,53].

In this project BiTeI is studied as it exhibits the topological insulator state under specific applied hydrostatic pressures. This crystal hosts these emergent quantum phenomena due to a broken spatial-inversion symmetry caused by spin-orbit coupling resulting from the Rashba effect [54].

### 2.3.1 Rashba Effect

Many crystals exhibit the time-reversal symmetry, shown in *Eqn. 15*, and spatial-inversion symmetry, shown in *Eqn. 16*. The time-reversal symmetry states that the wavefunction does not change under application of the time-reversal operator ( $\hat{T}$ ) which flips the spin state and reverses the momentum direction, equivalent to a time-reversal in the evolution of the system. In addition, the spatial-inversion symmetry represents an invariance to the system when mirrored, this is described by the inversion operator ( $\hat{I}$ ) and crystals with this property are labelled as centrosymmetric.

When these two transformations are composed together it is evident that the system must be spin-degenerate, as shown in *Eqn. 20*. Since, both operators commute with the Hamiltonian, they also commute with one another.

$$\hat{I}\hat{T}\Psi_{\vec{k},\uparrow(\downarrow)} = \Psi_{\vec{k},\downarrow(\uparrow)} = \Psi_{\vec{k},\uparrow(\downarrow)} \quad (20)$$

Topological insulators utilise spin-orbit interactions to produce the necessary band inversion characterising them [55]. Typically, this can be achieved by breaking time-reversal symmetry, using an external magnetic field. A resultant spin-degenerate minimum band-gap at  $\vec{k} = 0$  is produced forming a topological insulator [56]. Both the Rashba effect and topological insulators involve significant spin-orbit interactions, which indicates that the Rashba effect offers alternative means for generating topological materials through breaking of spatial-inversion symmetry. These different forms of symmetry broken band structures are illustrated in *Fig. 2*.

This mode of producing a topological insulator is realised in the BiTeI crystal [54,57] used in this project. This crystal under sufficient hydrostatic pressure, exhibits the topological insulator state with topological indices  $\mathbb{Z}_2 : \nu_0; (\nu_1, \nu_2, \nu_3) = 1; (0, 0, 1)$  as opposed to the trivial insulator state with  $\mathbb{Z}_2 : 0; (0, 0, 0)$  [54].

The Rashba effect occurs in BiTeI crystals as Tellurium and Bismuth are heavy ions which produce a strong spin-orbit coupling, leading to band inversion without an external magnetic field. As a consequence, electron spins are no longer degenerate in the Brillouin zone leading to spin-splitting. The Rashba effect is included through a term added to the Hamiltonian shown in *Eqn. 21*. It can be seen that this effect is both spin and momentum dependent, through the pauli matrices  $\sigma_i$  and momentum operator  $\vec{p}$  respectively.

$$H \propto (\vec{\sigma} \times \vec{p}) \cdot \hat{z} = \alpha (\sigma_x p_y - \sigma_y p_x) \quad (21)$$

Uniquely, the Rashba effect in BiTeI leads to several energy bandgap minima, away from the centre of the Brillouin zone. Usually, in centrosymmetric crystals, the bandgap reaches a minimum at the central high-symmetry point. However, in BiTeI

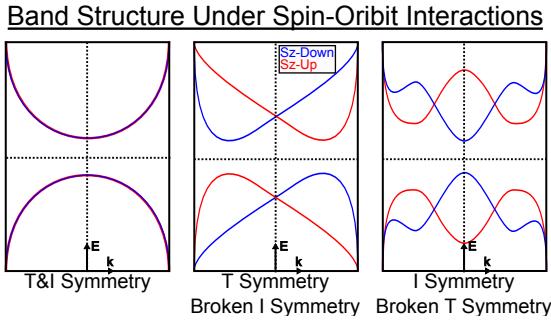


Figure 2: This is a pictorial representation of energy band structure under combinations of the time-reversal ( $T$ ) and spatial-inversion ( $I$ ) symmetries. The spin-states are coloured as blue for spin-down and red for spin-up. It is apparent, in the left diagram, that if both symmetries are present, the crystal is spin-degenerate. Whilst the central crystal represents a non-centrosymmetric crystal such as BiTeI with spin splitting due to a Rashba effect. The rightmost diagram indicates systems with broken time-reversal symmetry imposed by an external magnetic field.

six minima form in the direction of the A-L symmetry line [58], for the topological insulator state, and six minima along the A-H symmetry line for the trivial insulating state. Due to these multiple off-centre bandgap minima, novel properties and applications may be possible for BiTeI, including further spintronic applications [57, 59].

Furthermore, the Rashba effect results in BiTeI becoming a weyl semimetal under lesser pressure due to the spin splitting. However, since the band-crossing points occur away from the symmetry point  $A$ , this poses a problem for identifying the  $\vec{k}_i$  position of the crossing. These crossings are the positions of the monopoles and antimonopoles, as previously discussed. Under time-reversal symmetry the berry curvature is odd,  $\vec{B}(\vec{k}) = -\vec{B}(-\vec{k})$  and therefore the monopoles and antimonopoles are distributed symmetrically about the origin:  $\varrho(\vec{k}) = -\varrho(-\vec{k})$  [38]. After creation on A-H symmetry line, they dissociate and recombine symmetrically towards the A-L symmetry line. In the following sections, methods to identify these bandgap closing points through both grid-search and a novel application of gradient descent are described.

### 3 Methodology and Results of Classifying Topological States for Application to Machine Learning

The possible arrangements for crystals is near infinite, so there likely exists many undiscovered crystalline materials with desirable behaviours allowing for novel technologies. Material development and classification is a time consuming process requiring extensive experimental tests. Due to the vast number of candidate materials to be trialled, it is impractical to naively probe for useful materials without a method of refining the search space to fewer, more promising candidates for experimental verification. Machine learning, and more specifically artificial neural networks, has already been shown to be an effective tool is analogous regimes such as drug discovery [60–64]. The principle is that learned associations from a known subset of crystals can be developed and then applied to the remaining space. Therefore, it is likely that a machine learning approach can suitably, and efficiently, predict the properties of theoretical crystals [65]. As a result, the process of material development would be streamlined. With sufficiently accurate modelling, it may even be possible to specify desirable properties and develop a matching crystal structure from the specification. This is amble motivation to apply similar techniques to condensed matter systems.

This work aims to systematically prove the machine learning approach of modelling condensed matter systems. Particularly, prospective models must be verified against well-studied analytical systems before applying them to the wider unknown search space. Therefore, in this project, we curate a large analytical dataset for properties of Bismuth-Tellurium-Iodide crystals, specifically the classifications of material states, under various deformations to the crystal. Notably, BiTeI crystals exhibits a variety of such states, including the weyl semimetal and topological insulator emergent quantum states. Due to presence of these uniquely quantum mechanical phenomena, this dataset serves as the optimum platform to compare quantum machine learning against classical machine learning techniques in modelling condensed matter systems. As only quantum machine learning may be positioned to appropriately model these systems because of the shared quantum behaviour.

This dataset should consist of atomic orbital hopping coefficients for BiTeI, described in *Eqn. 7*, paired with their resultant material classification, such that a classification neural network can learn mapping relations between the two. This requires a sufficiently large dataset, due to the complicated system dynamics, so that the network can learn good approximate relations. The dataset can be expanded through crystal deformations, discussed in *Sec. 3.3*. Deformations to the crystal manifest in changes in the value of such hopping coefficients and this may result in a change of state. The quality of the trained neural network is limited by the accuracy of the dataset. Therefore, it is desirable to minimise any assumptions to produce a robust dataset which will not impede network generality. The following procedures are performed on every sample in the training dataset. Hence, it is imperative that efficient and accurate classification algorithms are implemented. The development of such methods is the central aim of this project. In the following sections, the algorithms developed to classify materials are motivated, discussed and implemented.

### 3.1 Novel Approach to Highly Parallelised Tight-Binding Equation Solvers

As discussed, in this project we will be using the tight-binding equation to model the BiTeI condensed matter system to develop an analytic dataset. The initial information provided is the hopping coefficients between neighbouring atomic orbitals, generated using density-functional-theory (DFT). Two samples of these hopping coefficients are provided, one when the BiTeI is known to be in a trivial insulating state and one when it is in a topological insulating state. These represent the crystal at two different hydrostatic pressures [54] and it is expected that these are separated by a weyl semimetal, and maybe also a dirac semimetal state. Therefore, classification of the four involved states is desired.

The DFT algorithm produces hopping coefficients for the electron originating in an atomic orbital, in the central unit cell, and transitioning to itself or other neighbouring atomic orbitals for all cells up to a cutoff distance. Hence, it is assumed that long-range interactions offer negligible corrections to the electronic band structure. Specifically, for BiTeI the cutoff is defined such that 1155 cells, closest to the central unit cell, are considered.

Within each cell the  $P_x$ ,  $P_y$  and  $P_z$  orbitals for Bismuth ( $6P$ ), Tellurium ( $5P$ ) and Iodide ( $5P$ ) significantly contribute so are considered alongside the electron's 2-component spin. This results in 18 atomic orbitals for every combination. Therefore, the electron can originate in one of 18 initial orbitals and jump to one of 18 final state orbitals, yielding a total of 324 permutations per cell. Further, considering the 1155 neighbouring cells produces 374420 complex hopping coefficients contributing to the electronic band structure.

Using *Eqn. 3*, the vector  $\vec{a} \in \mathbb{Z}^3$  takes values  $\alpha \equiv a_x \in [-6, 6]$ ,  $\beta \equiv a_y \in [-5, 5]$ ,  $\gamma \equiv a_z \in [-5, 5]$  in the DFT calculation, representing each neighbouring cell's displacement from the central cell at  $\vec{a} = \vec{0}$ . Not all these combinations of  $\vec{a}$  exist due to the neighbour cutoff. Furthermore,  $P_i \in ([0, 18] \cap \mathbb{Z})$  and  $P_f \in ([0, 18] \cap \mathbb{Z})$  denote the initial state and final state atomic orbital, where the orbitals are enumerated as shown in *Tab. 1*. Finally, for every permutation there is a complex-valued hopping term. These parameters can act as indices for an array representation of the hopping coefficients ( $T \in \mathbb{C}^{6 \times 5 \times 5 \times 18 \times 18}$  and indexed  $T_{a_1, a_2, a_3, P_i, P_j}$ ) with unknown hopping coefficients set to 0. Using an array structure enables repurposing of array-optimised machine learning libraries.

0 Te- $P_{x\uparrow}$	1 Te- $P_{y\uparrow}$	2 Te- $P_{z\uparrow}$	3 Bi- $P_{x\uparrow}$	4 Bi- $P_{y\uparrow}$	5 Bi- $P_{z\uparrow}$	6 I- $P_{x\uparrow}$	7 I- $P_{y\uparrow}$	8 I- $P_{z\uparrow}$
9 Te- $P_{x\downarrow}$	10 Te- $P_{y\downarrow}$	11 Te- $P_{z\downarrow}$	12 Bi- $P_{x\downarrow}$	13 Bi- $P_{y\downarrow}$	14 Bi- $P_{z\downarrow}$	15 I- $P_{x\downarrow}$	16 I- $P_{y\downarrow}$	17 I- $P_{z\downarrow}$

Table 1: Orbital Enumeration

Using this, various means of computing the tight-binding equation, shown in *Eqn. 8*, can be implemented. Although Python, considered as a slow computing language, is used in the project, it is merely for the accessing the front-end of the PyTorch library which utilises the graphics processing unit (GPU) for optimisation. Consequently, the following approach is executed in back-end CUDA code ran on the GPU. This allows very fast computation of array operations. Thus, we leverage this to compute the energy band structure in an efficient and novel manner.

Due to the array being discretised and unable to store functions, it is not possible to generate a closed form solution for the band structure at continuous crystal momentum values. Therefore, in the initial formulation of this procedure, the band

structure is sampled at a single value for  $\vec{k}$ . This is a direct implementation of *Eqn.* 8, where singular value decomposition is used to determine the  $n$  eigenvalues ( $E_n$ ) and eigenvectors ( $\psi_i^n$ ) for the resultant 2D hamiltonian. The energy eigenvalues are then sorted in increasing order ( $E_0 \leq E_1 \dots \leq E_n$ ), for each value of  $\vec{k}$ . This produces a sampling of the desired energy band structure and dispersion relations. Moreover, the Einstein summation convention [66] can be used to directly specify efficient array operations to be executed on a GPU. Therefore, for convenience of implementation, the convention will be used in all following algebra. This is demonstrated in *Eqn.* 22 describing this implementation of the tight-binding equation.

$$\hat{H}_{ij}\psi_{nj} = (e^{ia_{\alpha\beta\gamma l}r_{lm}k_m} T_{\alpha\beta\gamma ij}) \psi_{nj} = E_n \psi_{ni} \quad (22)$$

Where  $r_{ij}$  is the array of lattice vectors stacked row-wise and  $a_{\alpha\beta\gamma l}$  is an array lookup of considered cell offsets, such that  $a_{\alpha\beta\gamma} = (\alpha, \beta, \gamma)$ . The full singular value decomposition of the hamiltonian into the energy eigenvalues and eigenvectors is shown in *Eqn.* 23, where  $\delta_{ij}$  is the kronecker delta.

$$(e^{ia_{\alpha\beta\gamma l}r_{lm}k_m} T_{\alpha\beta\gamma ij} - E_n \delta_{ij}) \psi_{nj} = 0_{ni} \quad (23)$$

Furthermore, it was essential to make this calculation fully differentiable to enable a novel technique of locating bandgap minima in the Brillouin zone, discussed further in *Sec. 3.2.2*. The jacobian representing the derivative of the tight binding equation is shown in *Eqn.* 24 and is also proportional to the direction dependent group velocity of *Eqn.* 11.

$$\frac{\partial E_n}{\partial k_\varrho} = ia_{\alpha\beta\gamma l} r_{l\varrho} e^{ia_{\alpha\beta\gamma l} r_{l\varrho} k_\varrho} \psi_{ni}^* T_{\alpha\beta\gamma ij} \psi_{nj} \quad (24)$$

Despite the GPU optimisation, this method remains relatively slow as the computation needs to be calculated sequentially for small interval sampling of  $\vec{k}$  to appropriately resolve the continuous behaviour of the band structure. This sequential processing is an inefficient use of the large number of GPU cores. If sufficient video-RAM is accessible, then this process should be parallelised in  $\vec{k}$ . The implementation of multiple parallel computations of the tight binding model made up a significant proportion of this project. The parallelisation of *Eqn.* 22 in  $\vec{k}$  is shown in *Eqn.* 25. Where  $p$  discrete samples of the Brillouin zone are stacked row-wise to form the matrix  $k_{pm}$ . Thus, the tight-binding equation can be calculated for a large number of  $\vec{k}$  samples at once, by exploiting the GPU optimisation of array operations, making this method exceedingly efficient.

$$E_{pn}\psi_{pni} = e^{ia_{\alpha\beta\gamma l}r_{lm}k_{pm}} T_{\alpha\beta\gamma ij} \psi_{pnj} \quad (25)$$

Finally, it is intended that multiple hamiltonians will be evaluated to show the various material classifications which BiTeI can exhibit. If an adiabatic transformation is performed between the two DFT datasets, then a large set of hamiltonians will be generated. In a similar way to how the discrete  $\vec{k}$  samples were computed in parallel, this procedure can be repeated for the various hamiltonians stacked in their first index:  $q$  hamiltonians ( $T_{\alpha\beta\gamma ij}$ ) produce a six-dimensional array ( $T_{q\alpha\beta\gamma ij}$ ). However, to make each hamiltonian independent under gradient descent update by  $\partial E_n / \partial k_{pm}$  requires the array  $k_{pm}$  to be tiled  $q$  times producing  $k_{qpm}$ . This leads to the final einstein summation expression for a highly-parallelised approach to computing the tight-binding equation, as shown in *Eqn.* 26.

$$E_{qpn}\psi_{qpn} = e^{ia_{\alpha\beta\gamma l}r_{lm}k_{qpm}} T_{q\alpha\beta\gamma ij} \psi_{qpnj} \quad (26)$$

The respective derivative for this equation is required for later gradient analysis of the energy band structure. This derivative is shown in *Eqn.* 27 and represents the local evaluation of electronic band structures gradients.

$$\frac{\partial E_{qpn}}{\partial k_{sh\varrho}} = i(\delta_{qs}\delta_{ph})(a_{\alpha\beta\gamma l}r_{l\varrho}) (\psi_{qpn}^* T_{q\alpha\beta\gamma ij} \psi_{qpnj}) e^{ia_{\alpha\beta\gamma l}r_{l\varrho}k_{qpe}} \quad (27)$$

The independence of the hamiltonians and discrete  $\vec{k}$  samples is shown through the two kronecker deltas in *Eqn.* 27.

Overall, this approach effectively leverages GPU acceleration by computing the desired tight-binding equations in parallel. This novel technique offers a large improvement in computation time compared to similar approaches. However, it is limited

by the video-RAM capacity, since all the arrays must be loaded into the memory. Yet with new, purpose-built GPU's for machine learning, this constraint is not a impediment to the method. The implementation of this technique was shared with our group, to allow the efficient calculation of the required tight-binding results. This work made up a substantial backbone for further calculations in the project. An example of the BiTeI band structure produced using this technique is shown in *Fig. 3*.

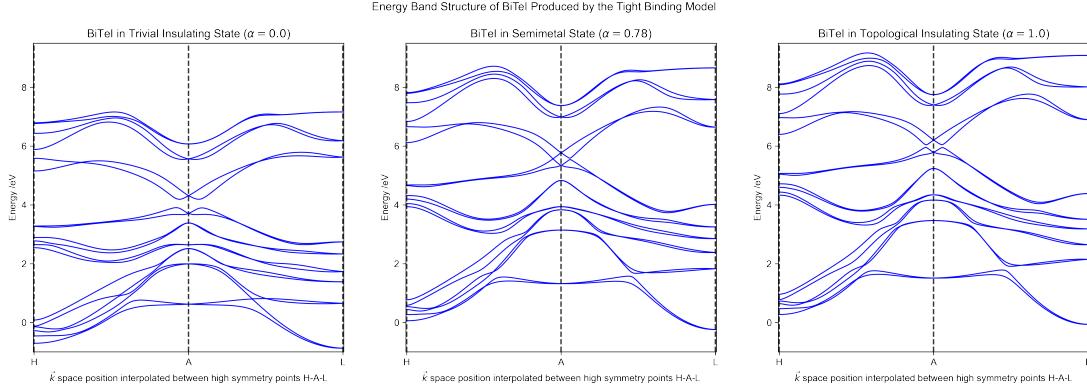


Figure 3: The energy band structures for BiTeI in three different states are shown, generated using the tight-binding equation. Left to right are the trivial insulating, semimetal and topological insulating states respectively, all evaluated along the H-A-L symmetry lines [58]. As expected 18 continuous bands are formed.

## 3.2 Material Classification Procedure

Using the parallel tight-binding equation solver, the nature of the Brillouin zone can now be determined for specific values of  $\vec{k}$ . With this, the various hamiltonians can be categorised into disjoint sets for trivial insulators, topological insulators and semimetals states of BiTeI. To distinguish between insulating and semimetallic states requires determination of the minimum bandgap between the valence and conduction bands. The fermi-energy level for BiTeI lies between the 12<sup>th</sup> and 13<sup>th</sup> bands, making these the valence and conduction bands respectively.

Therefore, the location corresponding to their closest approach ( $\vec{k}_{\min}$ ), in the Brillouin zone, is desired to separate the two states. Further splitting of trivial and topological insulating states can also be achieved through the associated eigenvectors at this minima. The details of this classification procedure are outlined in *Secs. 3.4* and *3.5*. In either case, the determination of the bandgap minima is required. Grid-search is a common technique [67,68] to determine parameters such a  $\vec{k}_{\min}$ , however we also present a novel application of the gradient descent algorithm to determine the bandgap minima.

### 3.2.1 Determining Minimum Bandgap by Grid Search and Random Sampling

Grid search can be described as a brute-force search technique for determining the Brillouin zone position of the minimum bandgap energy ( $\vec{k}_{\min}$ ) between the valence and conduction bands. It involves discretising the Brillouin zone up to a specified resolution, then applying the tight-binding equation at each  $\vec{k}$  value to determine the 12<sup>th</sup> and 13<sup>th</sup> energy eigenvalues. These are then subtracted from one another, shown in *Eqn. 28*, and the position of the minima is determined to be the value of  $\vec{k}$  corresponding to the smallest value of  $E_{BG}$ . For higher precision on  $\vec{k}_{\min}$ , a further sampling can be performed on a subspace surrounding the previously found minima. Though this makes the technique vulnerable to anomalous local minima. An analogous procedure can be used with a point cloud, of randomly sampled  $\vec{k}$  values within the Brillouin zone.

$$E_{BG}(\vec{k}) = E_{13}(\vec{k}) - E_{12}(\vec{k}) \quad (28)$$

Using the parallelised tight-binding equation, this method is reasonably efficient due to bulk computation of  $\vec{k}$  samples. However, the number of  $\vec{k}$  samples grows cubically with increased resolution. If repeated sequential sampling is performed, then on order of one hundred samples must be computed for each order-of-magnitude improvement in precise. It is preferable

to minimise any sequential computation performed in determining the minima for efficiency. Instead an application of the gradient descent algorithm, in finding  $\vec{k}_{\min}$ , is expected to have faster and more accurate convergence.

### 3.2.2 Gradient descent - A Novel Approach for Determining the Bandgap Minima

Gradient descent has proved to be an extraordinarily effective technique for parameter optimisation [69, 70]. This technique is an improvement over grid search, especially in higher dimensional parameter-spaces. It requires less initial  $\vec{k}$  samples to converge on more accurate minima, in fewer sequential steps. In addition, it has a variety of variations in its procedure for different applications. For example, the ADAM optimiser [71] is utilised in this project for even faster convergence when the position of minima are unknown. The elementary form of gradient descent consists of the parameter update shown in *Eqn. 29* for a given cost-function  $C$ . Where  $\eta$  is a small value known as the learning rate. Since all prior calculations are differentiable the gradients, such as from *Eqn. 27*, are easily implemented into *Eqn. 29*.

$$\vec{k}' = \vec{k} - \eta \nabla_{\vec{k}} C \quad (29)$$

The energy bandgap itself can be directly minimised. However, it is known that a linear dispersion relation of the bands is probable near the minima, leading to poor convergence due to equal update magnitudes. Thus, the bandgap energy is squared to ensure the desired convergence by an increasingly smaller update near the minima, as shown in *Eqn. 30*. Furthermore, it is commonplace to use stochastic gradient descent to optimise a mapping. However, this is undesirable when minimising an input parameter such as  $\vec{k}$  as it leads to the updates being, incorrectly, inversely scaled by the number of  $\vec{k}$  samples.

$$k'_{sh\varrho} = k_{sh\varrho} - \eta \frac{\partial (E_{sh(n=13)} - E_{sh(n=12)})^2}{\partial k_{sh\varrho}} \quad (30)$$

Local minima are a concern when using gradient descent, since parameters can converge on incorrect  $\vec{k}$  values. The overall risk of this is minimised by using the ADAM optimiser, since it utilises a momentum term in the algorithm. In addition, by using multiple  $\vec{k}$  samples and taking the minima of these, the likelihood of a local minima being found is reduced. Moreover, since the energy band-structure is the sum of finite hopping contributions we expect it to be reasonably smooth, containing finite minima as opposed to infinite fractal minima which would disrupt the technique.

Consequently, the novel application of gradient descent has been a resounding success in our project, enabling fast determination of bandgap minima in order to classify materials, as shown in *Fig. 4*. It has been implemented using PyTorch, alongside the discussed tight-binding approach, and used in the production and analysis of all our results.

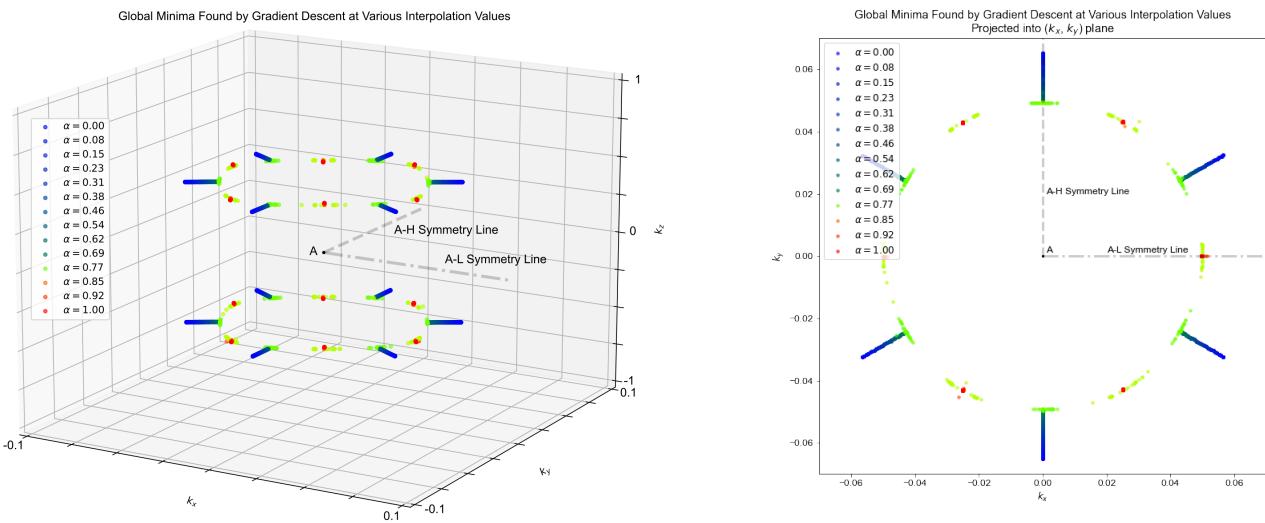


Figure 4: Plotted are the global minima found using the gradient descent procedure, at various values of the interpolation parameter  $\alpha$  discussed in *Sec. 3.3.1*. Left shows the minima as a 3D plot, showing all global minima, whilst right shows these projected into the  $k_x$  and  $k_y$  plane. The colour corresponds to the value of  $\alpha$  as shown in the key.

Multiple global minima are expected in the Brillouin zone for BiTeI due to the absence of spatial-inversion symmetry. Therefore, several distribution options are useful when initialising  $\vec{k}_{qpm}$ . For most data collection, a uniform random distribution across the entire Brillouin zone is desirable as it leads to better estimates of the location of the minimum energy bandgap and a larger learning rate can be used. This is the preferred method when the location of the minima is unknown. However, for later spin behaviour analysis, it is favourable to initialise all  $\vec{k}_{qpm}$  at a single point with a small learning rate. This is to encourage convergence onto specific global minima within a twelfth of the Brillouin zone. The smaller learning rate also confines the  $\vec{k}$  distribution better. Initialising  $\vec{k}_{qpm}$  in this manner allows the evolution of spin behaviour for a specific minima to be observed when varying the interpolation value. If the solutions are not confined to this region, then large shifts in spin components are observed due to spin splitting by the Rashba effect. This latter procedure was used to locate the minimas for *Fig. 11*. Checks should be done to ensure the resultant distribution of the global minima is known and ruled out as a source of drastic eigenvector changes.

### 3.3 Dataset Expansion

Training a neural network to model the given condensed matter system requires an extensive dataset, such that it can learn to relate the input information to the desired output. Neural networks typically learn relations better when producing a strongly abstractive map. We define abstraction as any map which lowers the intrinsic dimension of data, whilst transformative and generative<sup>1</sup> describe maps which keep constant or increase the intrinsic dimension respectively. The intrinsic dimension is a measure of the minimal degrees of freedom needed to span the embedded data, whilst the extrinsic dimension is the dimension of the embedding space. The dataset, so far, includes BiTeI at two different pressures corresponding to a trivial insulating state and a topological insulating state. This is insufficient to train an effective neural network, therefore the dataset should be expanded to improve the resultant model's generality.

#### 3.3.1 Interpolation

One method to expand the dataset is linear interpolation between the two known sets of hopping parameters. This is equivalent to varying the hydrostatic pressure between the two samples and is required to produce examples of the weyl semimetal state in the dataset. This procedure is shown in *Eqn. 31* and characterised by a mixing parameter  $\alpha \in (\mathbb{R} \cap [0, 1])$ .

$$T'(\alpha) = T^{\text{Trivial}}(1 - \alpha) + T^{\text{Topological}}\alpha \quad (31)$$

The resultant hopping coefficients ( $T'$ ) can be used to generate new samples for the dataset. Furthermore, since the hamiltonians cannot be adiabatically transformed from a trivial insulating state to a topological insulating state without first crossing a semimetal state, then we would expect two critical values for  $\alpha$ . The first is the boundary between the trivial insulator and the semimetal state, representing the monopole-antimonopole creation point, whilst the second critical value is their destruction on the semimetal to topological insulator transition. A linear dependence between the minimum bandgap energy and interpolation parameter is expected in each of the insulating states. Using samples of  $\alpha$  a linear regression can be performed, for values in each insulating state, to find the critical values of  $\alpha$  which are the root of each respective fit.

The dataset generated via interpolation produces a sampling of a 1D continuous manifold of input data. We categorise this into a disjoint set of 4 possible elements: trivial insulators, topological insulators, weyl semimetal and dirac semimetal. This represents 4, 0-dimensional points which can be connected using a minimum of a 1D manifold. Therefore, the network is only provided a reformative map to ascertain relations from, which is sub-optimal. Therefore, a further technique is required to expand the dataset.

#### 3.3.2 Random Noise

Generating samples perturbed by random noise will increase the input's intrinsic dimension, depicted in *Fig. 5*. This will provide more varied samples to reduce network overfitting [72] and consequently improve the model's generality. These perturbed hamiltonians are also labelled using the same procedure as the interpolated samples. Therefore, it is not assumed

---

<sup>1</sup>Due to insufficient input degrees of freedom, such a dataset requires the network to be a multi-valued function which is not possible. However, the network can still approximate this process by folding discretely sampled manifolds so that they appear to have increased in intrinsic dimension.

that a consistent classification is upheld in a localised region around a previously classified sample. This is especially important when considering any discontinuous phase change. Since a neural network is limited by the quality of the dataset, the lack of such assumptions should improve the modelling.

However, unconstrained noise can result in unphysical hamiltonians. In addition, excessive network parameters would be required to model these cases. Therefore, two forms of constrained hermitian perturbations can be produced: spatial-inversion symmetric and time-reversal symmetric. These represent a variety of deformations to, or external fields applied, to BiTeI.

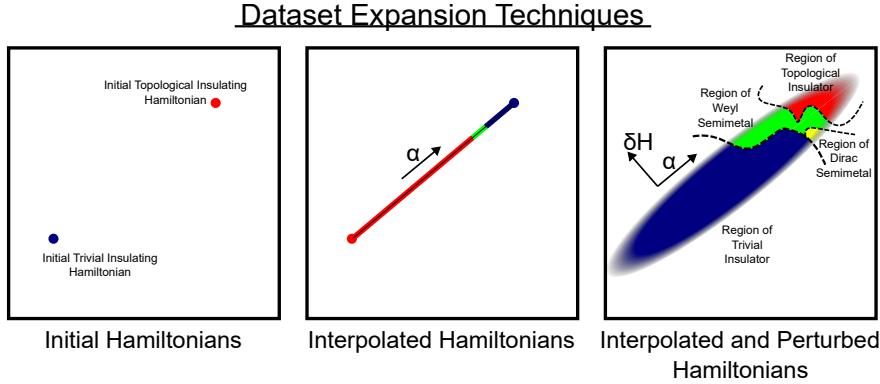


Figure 5: This illustration depicts a 2D slice of the possible hamiltonian space. Left shows the initial dataset with just two samples, thus two 0-dimensional points embedded in the space. Centre shows how the dataset can be expanded by interpolating between the initial hamiltonians, producing a sampling of a 1D continuous manifold embedded in the space. Right shows how the dataset is then further expanded using constrained perturbations producing a much higher dimensional manifold. The colours depict the disjoint material classifications and the dashed lines indicate the boundaries between these.

Spatial-inversion symmetric perturbations can be generated using the condition shown in *Eqn. 32*.

$$\delta H_I(\vec{k}) = \hat{P}(\delta H_I(-\vec{k})) \hat{P}^{-1} \quad (32)$$

Where  $\hat{P}$  is defined by  $u(\vec{k}) = \hat{P}u(-\vec{k})$ , from *Eqn. 4*. This corresponds to a diagonal matrix operator with elements  $\pm 1$ . In addition, time-reversal perturbations must abide by the condition shown in *Eqn. 33* [34].

$$\delta H_T(\vec{k}) = \hat{\sigma}_y (\delta H_T(\vec{k}))^T \hat{\sigma}_y \quad (33)$$

Where  $\delta H$  is the perturbation to the hamiltonian in the representation shown in *Eqn. 34*.

$$\delta H(\vec{k}) = \begin{pmatrix} \delta H_{\uparrow\uparrow}(\vec{k}) & \delta H_{\uparrow\downarrow}(\vec{k}) \\ \delta H_{\downarrow\uparrow}(\vec{k}) & \delta H_{\downarrow\downarrow}(\vec{k}) \end{pmatrix} \quad (34)$$

Uniform, random matrices can be drawn from distributions respecting these symmetries and then normalised. These random perturbations can be scaled and added elementwise to an interpolated hamiltonian before classification. This procedure suitably expands the dataset to have a very high intrinsic dimension, improving the neural network's performance. However, it has been shown [34] that adding inversion-symmetric perturbations changes the interpolation critical values, this is depicted in *Fig. 6*.

As a consequence, the material transitions cannot be identified using critical values of the interpolation parameter, found by aforementioned linear regression of the minimum energy bandgap. Further, generalising this to a hyperplane fit would not suitably account for a non-linear boundary and any further fitting would be needlessly contrived and overfitted. Therefore, a more robust procedure is required to generalise to the classification of randomly perturbed hamiltonians. This is achieved using a simpler approach of analysing the energy bandgap.

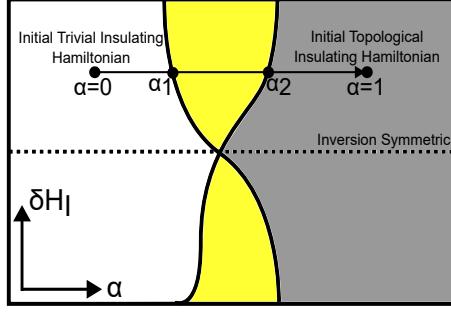


Figure 6: This diagram [34, 38] illustrates how the addition of spatial-inversion asymmetric noise alters the critical values for the interpolation parameter  $\alpha$ . In the spatial-inversion symmetric case, it can be seen that there is a single critical value, whereas inversion-asymmetric case there are two. White shading indicates a trivial insulator, whilst grey and yellow indicate a topological insulator and semimetal states respectively. A line is drawn indicating the interpolation path between the two initial hamiltonians.

### 3.4 Energy Eigenvalue Analysis

By definition, the two forms of insulating states are expected to have a non-zero bulk energy bandgap between the valence and conduction bands, whilst the semimetal states are gapless. Due to this behaviour, the insulating and semimetal states are separable by measuring the minimum bandgap energy obtained using the procedure discussed in *Sec. 3.2.2*.

When applying increasing hydrostatic pressure, by increasing  $\alpha$  from *Eqn. 31*, the BiTeI transitions from a trivial insulator to a semimetal, then to a topological insulator. Under this deformation the energy bandgap decreases whilst in the trivial insulating state, until the gap closes to form the gapless semimetal state at a critical value of  $\alpha_1$ . Since BiTeI is a non-centrosymmetric crystal, we would expect it to exist in this semimetal state for a finite range of  $\alpha$  values, thus there exists an upper bound critical values of  $\alpha_2$ . This was directly observed in *Fig. 7*, with a linear fit finding values of  $\alpha_1 = (0.784 \pm 0.001)$  and  $\alpha_2 = (0.796 \pm 0.001)$ .

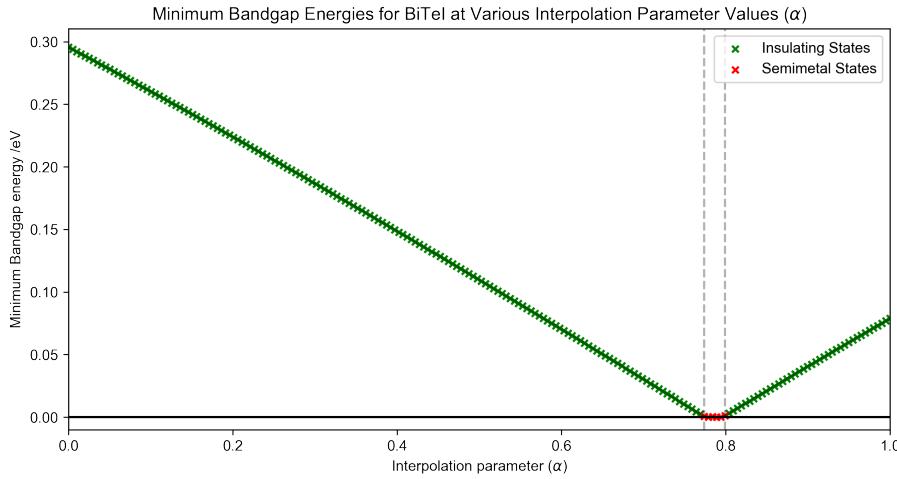


Figure 7: This shows the minimum energy bandgap ( $E_{BG}(\vec{k}_0)$ ) between the valence and conduction bands of BiTeI at various interpolation parameter values. The samples classification as semimetals ( $E_{BG}(\vec{k}_0) < \epsilon$ ) are shown in red whilst insulators ( $E_{BG}(\vec{k}_0) \geq \epsilon$ ) are shown in green. The grey dashed lines indicate the critical values  $\alpha_1$  and  $\alpha_2$ .

Further hydrostatic pressure results in band inversion and the separation of the bands, producing a topological insulator. Since the tight-binding equation assumes an infinitely extending crystal, the gapless conductive boundary of this topological insulator will not be produced in the modelled band structure nor in the value of the minima, shown in *Fig. 7*. Consequently, it is possible to define a threshold energy value which distinguishes the insulating states from gapless semimetal states. Due to the imprecision of estimating the, often single point, bandgap minima alongside any floating point errors, the determined minima may sometimes be small but non-zero for the semimetal state. Therefore, a threshold value of  $\epsilon = 2\text{meV}$  is used.

This procedure produces classifications for general insulating and semimetallic states. However, due to the minimum bandgap energy being a surjective but non-injective function of the interpolation parameter, there is insufficient information to produce the further sub-classifications of the semimetallic and insulating states. To determine whether the insulating states are topological or trivial requires analysis of the associated eigenvectors.

### 3.5 Eigenvector Analysis and Spin Behaviour

The eigenvalues and eigenvectors provide a complete basis of the hamiltonian so should be sufficient for determining further classifications. Calculating the topological numbers associated with  $\mathbb{Z}_2$  would be a robust way to categorise the different states although it is too slow computationally. Therefore an alternative procedure is required to efficiently generate a very large dataset. We use a cosine similarity measure of the band composition to achieve this. The transition from trivial to topological insulator is characterised by a band inversion, shown in *Fig. 8*. This process can only occur in the gapless semimetal state. Therefore, we would expect the eigenvectors, indicating the orbital hybridisation of the bands, to remain roughly constant and unique for each of the trivial and topological hamiltonians with small perturbations, shown in *Fig. 9*.

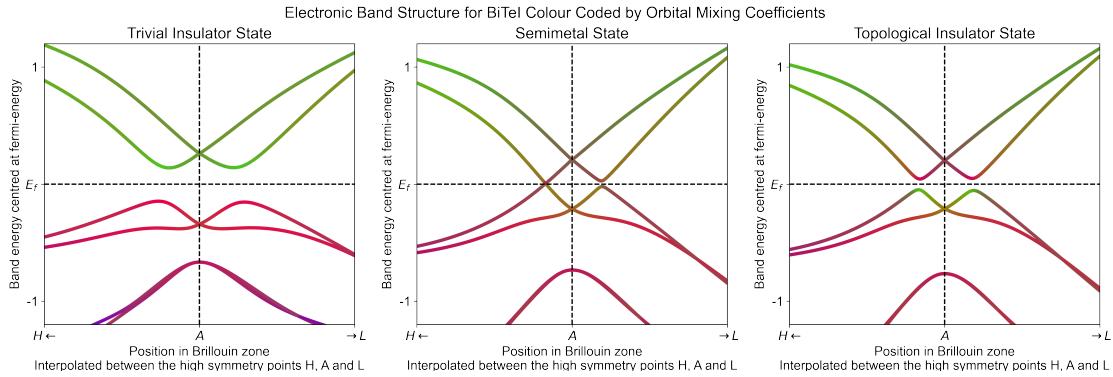


Figure 8: The valence and conduction band structure for BiTeI is shown along a section of the crystal symmetry lines H-A-L. The colours indicate the composition of each band in terms of the respective atomic orbitals, where red indicates Tellurium, green for Bismuth and blue for Iodide. The colour component represents the sum over all orbital angular momenta and spin configurations states. Left indicates the BiTeI crystal in its trivial insulating state, middle is the semimetal state and right is the topological insulator state. The band-inversion is clearly demonstrated by the mixing of the components in the semimetal and topological insulator state.

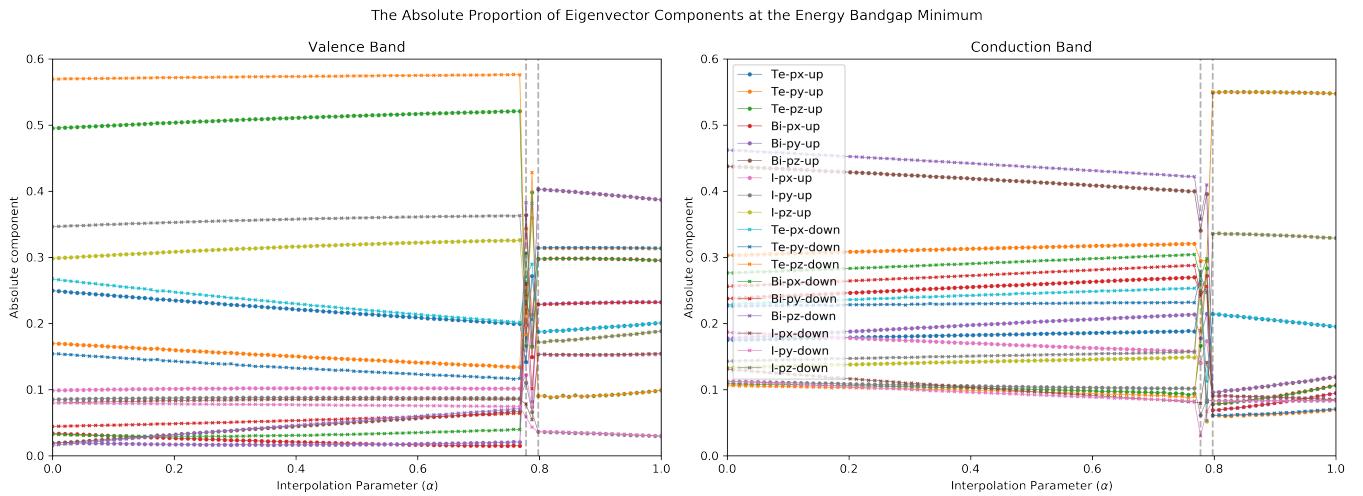


Figure 9: The absolute value of the eigenvector components is shown, taken at the energy bandgap minima for the valence band (left) and conduction band (right). It can be seen that band inversion occurs as expected, since the various components drastically change value before and after the critical values shown by the dashed, vertical grey lines. Spin-splitting is observed, as the graph represents one global minima rather than the average between a time-reversal symmetry linked pair. The idealised vectors are taken at  $\alpha = 0$  and  $\alpha = 1$ .

Consequently, we can define two vectors representing the idealised band composition evaluated at the bandgap minima. These are taken to be the eigenvectors for the  $\alpha = 0$  trivial ( $|\beta_{n,\text{Tri.}}\rangle$ ) and  $\alpha = 1$  topological ( $|\beta_{n,\text{Top.}}\rangle$ ) insulators. Since, the global minima vary in spin properties, then the idealised vectors should be spin-independent. The components of the idealised vectors are shown in *Fig. 10*.

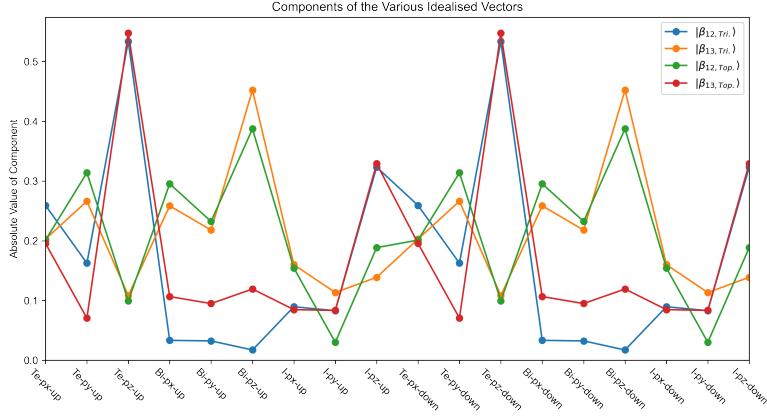


Figure 10: The components of the various idealised vectors are displayed. The spin-independence can be observed in the repetition of components for spin-up and spin-down pairs. Tellurium- $P_z$  and Iodine- $P_z$  have large components in  $|\beta_{12,\text{Tri.}}\rangle$  and  $|\beta_{13,\text{Top.}}\rangle$ , whilst Bismuth- $P_z$  is a large component in the other pair of idealised vectors. This is indicative of band-inversion and suggests the  $P_z$  orbitals are primarily involved in this state change.

The eigenvector at the energy bandgap minima ( $|\phi_n\rangle$ ) and the idealised vectors should be absolute-valued and normalised to one. Furthermore, both valence and conduction bands contribute to the location of the bandgap minima, so both of their eigenvectors should be considered. These are bands  $n = 12$  and  $n = 13$  respectively for BiTeI. Using these vectors, a dot-product measures the angular separation between the measured eigenvector and the idealised state. These are used to provide two coefficients representing the trivial and topological character shown in *Eqns. 35* and *36* accordingly.

$$C_{\text{Tri.}} = \frac{\langle \beta_{12,\text{Tri.}} | \phi_{12} \rangle + \langle \beta_{13,\text{Tri.}} | \phi_{13} \rangle}{2} \quad (35)$$

$$C_{\text{Top.}} = \frac{\langle \beta_{12,\text{Top.}} | \phi_{12} \rangle + \langle \beta_{13,\text{Top.}} | \phi_{13} \rangle}{2} \quad (36)$$

The components of spin represented in the bands also drastically changes in the two insulating states due to band inversion, as shown in *Fig. 11*. These spin-components are determined by taking the expectation values of each pauli matrices with the summed, spin-up and spin-down components of the eigenvector.

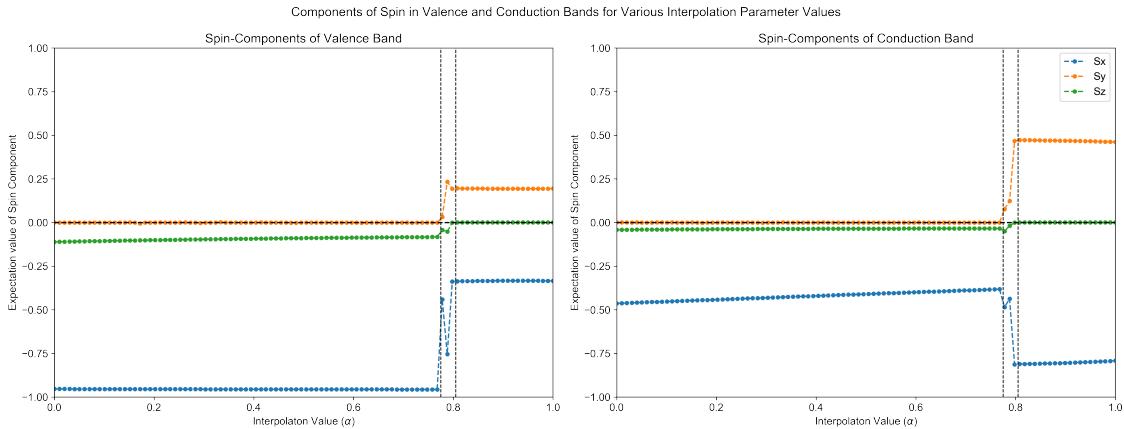


Figure 11: Shows the spin expectation values evaluated at the energy bandgap minima for various  $\alpha$ . A distinct change in the spin-components is observed in the trivial and topological insulating states. The grey dashed lines indicate the boundaries of the semimetal state determined from eigenvalue analysis.

It can be seen that  $S_y$  is zero when BiTeI is a trivial insulator and non-zero when it is a topological insulator, whilst  $S_z$  shows the opposite behaviour. Therefore, conditions could be developed to determine the state using this spin behaviour. However, this is a potentially unreliable and more computationally intensive mode of determining the insulating state than the previous orbital components. This is because the spin depends on which minima is found by gradient descent, whilst the orbital analysis does not, so this method does not generalise well. Therefore, we proceeded with a spin-independent method of determining the insulating state. However, the time-reversal symmetric and Rashba effect properties of BiTeI are demonstrable in the spin behaviour as shown in *Fig. 12*.

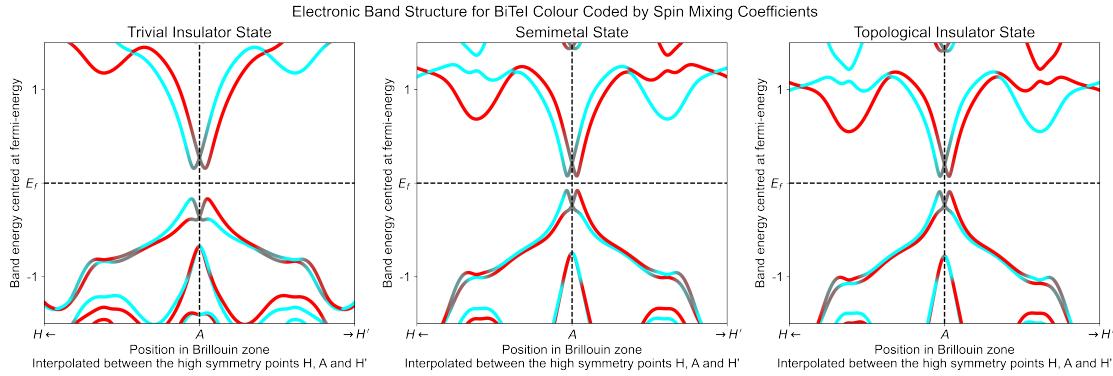


Figure 12: The valence and conduction band structure for BiTeI is shown along a section of the crystal symmetry lines H-A-H'. Red indicates the average contribution by spin-up eigenvector components whilst cyan indicates the spin-down components. Left, middle and right shows BiTeI in its trivial insulating, semimetal and topologically insulating states respectively. It can be seen that BiTeI exhibits spin-splitting through spatial-inversion symmetry breaking, caused by the Rashba effect.

Furthermore, *Fig. 13* shows the characteristic chiral monopoles and antimonopoles of a weyl semimetal and their spin properties. It can be seen that the initial six global minima of the trivial insulator state divide into twelve minima in the semimetal state, indicating the creation and location of the chiral monopoles.

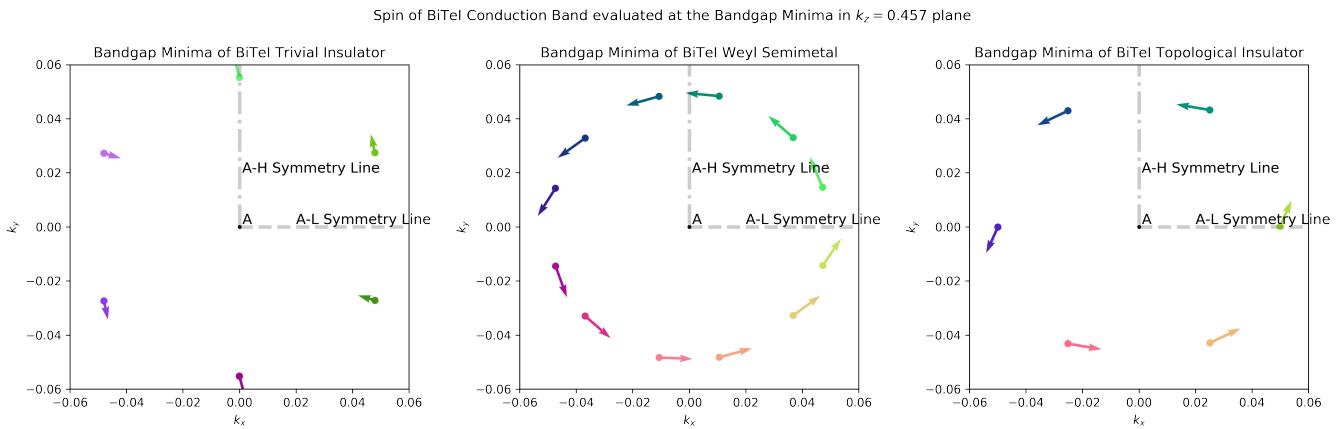


Figure 13: Shows the spins, as arrows and colours, of the conduction band eigenvectors evaluated at the bandgap minima for various values of the interpolation value  $\alpha \in \{0.4, 0.786, 1\}$  corresponding to three BiTeI states. The arrows indicate the  $\langle S_x \rangle$ , also indicated by red colour, and  $\langle S_y \rangle$ , also indicated by green colour, of the spin components, whilst  $\langle S_z \rangle$  is indicated by the blue colour component. In both the trivial insulator (left) and topological insulator (right) six global minima are observed near to the plane  $k_z = 0.457$  along the A-H and A-L lines respectively. It can be seen that six weyl momopoles and six antimonopoles are created in the weyl semimetal state as expected (centre). It can be observed that their in-plane spins are orthogonal to their in-plane momentum. These pairs dissociate and then annihilate along the A-L symmetry line of the material to form a topological insulator state.

The behaviour of weyl monopoles trajectories within the Brillouin zone is more evident in *Fig. 14*. This shows the location where creation and annihilation of the monopoles occurs. Unexpectedly, the minima of the topological insulator are shown not to exist along the A-L and A-L' symmetry lines, but rather in the plane defined by A, L and  $\Gamma$ .

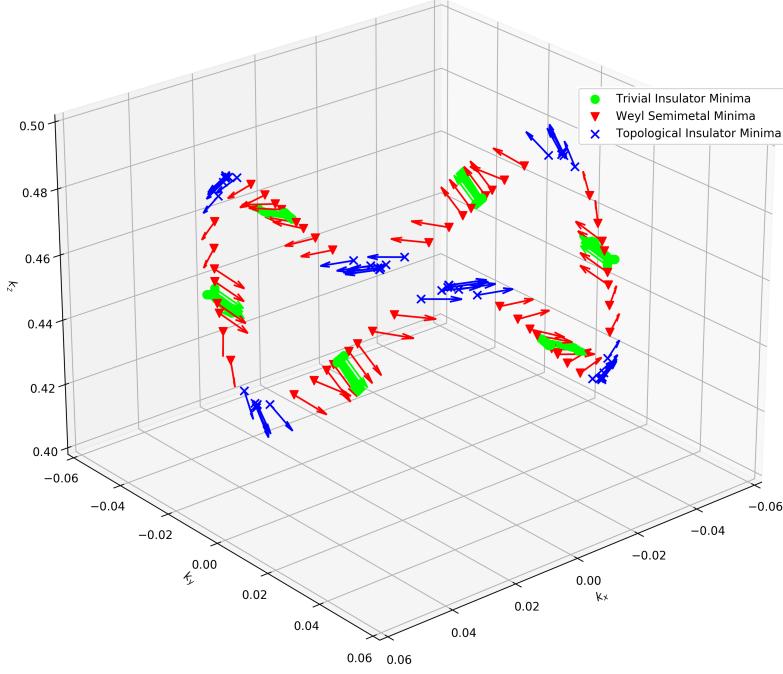


Figure 14: Shows the 3D arrangement of global minima at various values of  $\alpha \in [0, 1]$ , with colours indicating the classified states of the minima and arrows indicating spin expectation. It can be seen that all minima for trivial insulator states occur along  $A-H$  and  $A-H'$  symmetry lines. The red states indicate the centre of dirac cones in the semimetal phase, they are grouped into sets of twelve minima corresponding to the monopoles. These are shown to be created along the  $A-H$  and  $A-H'$  lines and then traverse through the Brillouin zone before annihilating at topological insulator minima. The topological minima are arranged out of the  $k_z = 0.457$ , giving a three-dimensional structure.

Returning to classification by *Eqns. 35* and *36* characters, the insulating state is determined to be the maximum value of these coefficients. This was found to be a successful strategy of classify the trivial and topological insulators, as shown in *Fig. 15*. The correct states are determined, without any outliers, at high resolution of the interpolation parameter.

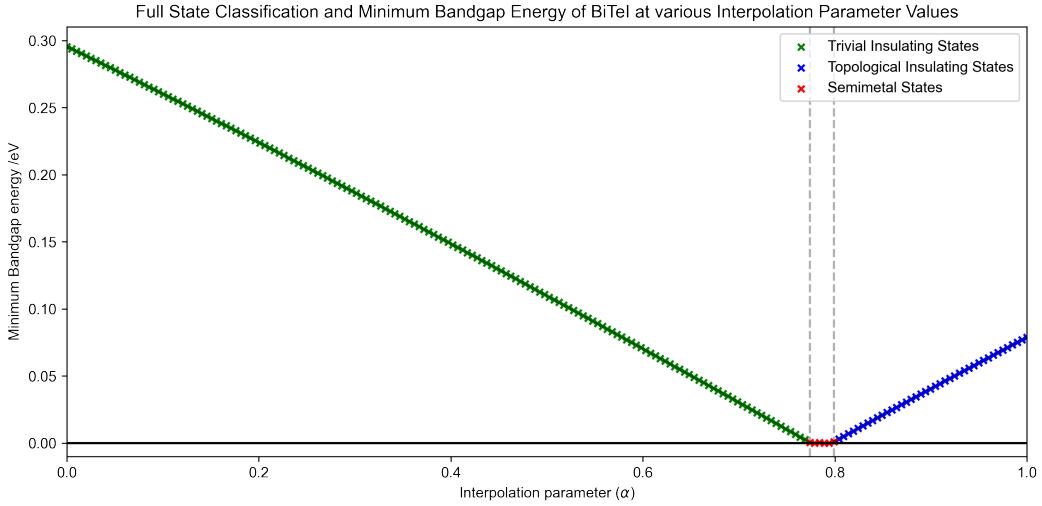


Figure 15: This shows the classifications of BiTeI states plotted in various colours. The trivial and topological insulating states are distinguished using the maximum of *Eqns. 35* and *36*. These classifications are shown on a plot of the minimum bandgap energies against the interpolation value to better illustrate the state transitions of BiTeI under hydrostatic pressure.

Therefore the classification of the semimetal, trivial insulator and topological insulator states of BiTeI has been demonstrated using the electronic band structure. However, the separation of weyl and dirac semimetals must still be implemented. It can be achieved using a analysis of the spin-orbit terms.

### 3.6 Analysis of Spin-Orbit Terms

To distinguish between the weyl and dirac semimetal states, requires directly analysing the hamiltonian. A dirac semimetal consists of a 4-fold degeneracy at the band crossing point. As discussed, this requires both spatial-inversion symmetry and time-reversal symmetry for dirac semimetals to arise. The Rashba effect is responsible for the spin-orbit coupling which breaks inversion symmetry. Therefore, a dirac semimetal will not be observed in the interpolated only hamiltonians, but may arise in a perturbed hamiltonian where a centrosymmetric BiTeI could occur.

Therefore, the presence of a spin-orbit interaction can be used to determine between a weyl and dirac semimetal. Once a hamiltonian is classified as a semimetal from the aforementioned procedures, then the presence of non-zero spin-orbit terms indicates it is a weyl semimetal. This is described by *Eqn. 34*, when  $(H_{\uparrow\downarrow}(\vec{k}) \neq 0) \cup (H_{\downarrow\uparrow}(\vec{k}) \neq 0)$ . Otherwise, it is a dirac semimetal. The implementation of this final condition concludes our classification procedure for BiTeI, allowing a dataset of BiTeI's hamiltonians and their respective classifications to be produced for application to machine learning.

## 4 Conclusion

The novel approaches developed to analyse Bismuth-Tellurium-Iodide crystals were successful at categorising the expected states and matched results from other analyses. Generalised techniques were developed using the energy bandgap, eigenvector components and spin-orbit interaction, to robustly classify BiTeI, under different deformations, into the four described states. The overall methodology is summarised in *Fig. 16*. The parallel tight-binding equation solver and gradient descent were very successful and pivotal in the production of all the results. These techniques utilised new developments in GPU acceleration and parameter-optimisation to efficiently perform the crucial calculations. In addition, the methods are robust to additional noise, allowing a larger and more varied analytical dataset to be developed. This preemptively reduces the neural network overfitting problem. Moreover, this approach is general, allowing a similar methodology to be easily applied to other crystalline materials with minor modifications. Therefore, it is feasible to efficiently create large datasets of many crystalline materials, enabling the development and benchmarking of models which operate upon a variety of materials. Overall, it is hoped that this will streamline research into material development by identifying more promising candidates and deformations which yield exciting properties to then experimentally verify.

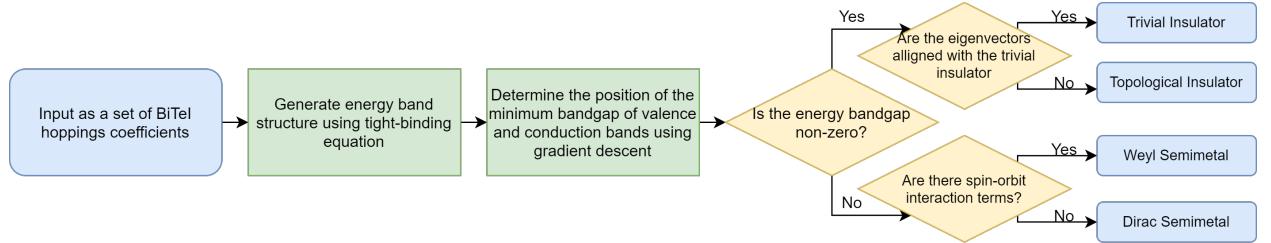


Figure 16: Shows a schematic approach to the procedure, where initial hopping coefficients are classified into the four states BiTeI is expected to exhibit.

The procedures correctly identify the weyl and dirac semimetal states as well as the trivial and topological insulating states exhibited by BiTeI. The respective gapless and gapped-bulk properties were observed in the energy band structures. These also offered insight into the process of band inversion, required to form topological insulators. Additionally, the weyl semimetal's monopole and antimonopole quasiparticles were observable as twelve global minima, demonstrating their existence between a range of hydrostatic pressures for the non-centrosymmetric BiTeI. It was shown that these are created, then migrate symmetrically through the Brillouin zone before being annihilated in different pairs to form the topological insulator. The presence of both band-inversion and Rashba spin-splitting was demonstrated in the eigenvector analysis, alongside the evolution of spin and orbital components with hydrostatic pressure. Furthermore, trivial and topological insulators were shown to be distinct states which cannot be adiabatically transformed into one another without forming a gapless weyl semimetal state. Consequently, the developed methods fulfilled our primary aim of this project to develop a standardised dataset and give insight into BiTeI under various deformations. However, there remains some caveats to the procedure, discussed below.

If sufficiently large perturbations are introduced, the fast idealised vector analysis may be ineffective as the proportions of orbitals represented in the valence and conduction band structure will begin differ without changing the material's state. Therefore, for this method to remain precise, only small perturbations should be used to expand the dataset or a re-sampling of the idealised vectors should be undertaken.

Furthermore, the classification list we developed is likely not exhaustive in multiple ways. A nodal semimetal [73] is a further semimetal classification which may occur with certain perturbations. These are characterised by lines of band crossings and would be classified as a weyl semimetal by our model. Despite this, inclusion of further divisions of the weyl semimetal state were out-of-scope for this project due to time constraints. In addition, there exists fifteen sub-classifications of the topological insulator state, characterised by combinations of the  $\mathbb{Z}_2$  invariant topological numbers  $\nu_0; (\nu_1, \nu_2, \nu_3)$ . These are all grouped collectively under a single topological insulator state in our approach.

Finally, there may exist undocumented states in the possible hamiltonian phase space. By definition, these are not identifiable using our approach of developing a dataset. However, indirect evidence of these undocumented states, and further sub-classifications, may be observed in the machine learning model. The cost function of a machine learning model may grow anomalously large when one of these undocumented states are encountered. This can occur when a state is misrepresented in the dataset, so the model may struggle to classify it. If this behaviour is observed, then further investigation will be undertaken to determine the characteristics of these states and whether they are novel.

Overall, these deficiencies in the approach are acceptable for the fast and efficient classifier we required for generating a large dataset. The nodal semimetal and finer classification of topological insulators should not impede the machine learning model, as they are still represented under the broader classifications used. In addition, assumptions used in the production of the dataset are made clear. Thus, the dataset is fully documented making it a useful and standardised platform to verify future work in this area. Using this dataset we can now proceed with developing quantum machine learning and classical machine learning models in the continuation of this project.

## 4.1 Outlook of Work

### 4.1.1 Benchmarking Supervised Learning Models

This project work has enabled the development of a large analytical dataset pairing BiTeI hamiltonians to their resultant material classification, making it suitable for application to supervised learning [74]. In the immediate follow-up work to this project, the dataset will be partitioned into training and testing subsets, allowing new machine learning techniques to be proven against this well-studied system. There have been several attempts [75–77] at applying machine learning to condensed matter systems and this dataset offers a further avenue for verifying such approaches, particularly in 3D topological systems. Once a machine learning model is benchmarked using this dataset, the technique can confidently be applied to further systems which display similar characteristics. The dataset can also be used to expand the number of training samples for the model in these cases.

### 4.1.2 Quantum Machine Learning Approach

Quantum machine-learning is an avenue of modelling condensed matter systems which may be particularly successful due to the shared quantum nature [78]. This approach would allow faster computation of quantum behaviours, such as entanglement involved in the emergent material phenomena. The dataset produced in this project can be used to assess the effectiveness of such an approach. Therefore, it can be determined whether quantum machine learning techniques better model condensed matter systems than the classical machine learning approaches. Furthermore, this dataset includes examples of emergent quantum phenomena, so it is the perfect dataset for comparing such models.

### 4.1.3 Generalising for Many Crystalline Materials

When developing this approach, a universal crystal embedding procedure arose which can result in a single machine learning model being applied to a wide variety of crystalline materials. This extension of the procedure is very promising and will be explored in a follow up to this work. The general form of artificial neural networks required for property predictions are

known as classification networks. These are the most common application of neural networks. Consequently, a wide variety of preexisting neural network architectures exist for this form of problem. Benchmarking should begin on the most primitive yet general neural network techniques, such as fully-connected neural networks, before testing more advanced architectures with narrower applications, such as convolution or transformer networks. However, these more advanced architectures are purpose-built for their respective tasks such as computer vision or natural language processing. Therefore, to achieve the optimal results might require novel neural network architectures designed specifically for condensed matter system. The combination of a universal crystal embedding and purpose-built neural networks would enable mappings learnt from one material to be usefully applied to another, including those without analytical solutions. This approach, if proved successful on well-studied materials, could streamline development of new materials. Moreover, if it becomes sufficiently well-trained, the model could be used to specify desirable properties and deduce a physical crystal featuring these properties. The prospects of this extended approach are exciting with potentially many applications.

## References

- [1] A. Bolotin, “Computational complexity and the interpretation of a quantum state vector,” 2013.
- [2] N. W. Ashcroft and N. D. Mermin, *Solid state physics*. Cengage, 2021.
- [3] F. Bloch, “Über die quantenmechanik der elektronen in kristallgittern,” *Zeitschrift für Physik*, vol. 52, pp. 555–600, Jul 1929.
- [4] V. L. Gurevich and A. Thellung, “Quasimomentum in the theory of elasticity and its conservation,” *Phys. Rev. B*, vol. 42, pp. 7345–7349, Oct 1990.
- [5] M. Born and R. Oppenheimer, “Zur quantentheorie der moleküle,” *Annalen der Physik*, vol. 389, no. 20, pp. 457–484, 1927.
- [6] P. Debye, “Zur theorie der spezifischen wärmen,” *Annalen der Physik*, vol. 344, no. 14, pp. 789–839, 1912.
- [7] A. V. Anda, S. S. Makler, R. G. Barrera, and H. M. Pastawski, “Electron - phonon effects on transport in mesoscopic heterostructures,” *Brazilian Journal of Physics*, vol. 24, pp. 330–336, 1994.
- [8] P. Babilotte, E. Morozov, P. Ruello, D. Mounier, M. Edely, J.-M. Breteau, A. Bulou, and V. Gusev, “Physical mechanism of coherent acoustic phonons generation and detection in gaas semiconductor,” *Journal of Physics: Conference Series*, vol. 92, p. 012019, dec 2007.
- [9] I. E. Tamm, “The electrodynamics of anisotropic media in the special theory of relativity,” 1923.
- [10] F. Giustino, “Electron-phonon interactions from first principles,” *Rev. Mod. Phys.*, vol. 89, p. 015003, Feb 2017.
- [11] J. Zhou, H. D. Shin, K. Chen, B. Song, R. A. Duncan, Q. Xu, A. A. Maznev, K. A. Nelson, and G. Chen, “Direct observation of large electron-phonon interaction effect on phonon heat transport,” *Nature Communications*, vol. 11, p. 6040, Nov 2020.
- [12] J. L. Mañes, “Symmetry-based approach to electron-phonon interactions in graphene,” *Phys. Rev. B*, vol. 76, p. 045430, Jul 2007.
- [13] L. N. Cooper, “Bound electron pairs in a degenerate fermi gas,” *Phys. Rev.*, vol. 104, pp. 1189–1190, Nov 1956.
- [14] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, “Microscopic theory of superconductivity,” *Phys. Rev.*, vol. 106, pp. 162–164, Apr 1957.
- [15] J. E. Lennard-Jones, “The electronic structure of some diatomic molecules,” *Trans. Faraday Soc.*, vol. 25, pp. 668–686, 1929.

- [16] I. G. Kaplan, “Pauli exclusion principle and its theoretical foundation,” 2019.
- [17] D. Rayner and W. Hartree, “Self-consistent field, with exchange, for beryllium,” *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences*, vol. 150, no. 869, pp. 9–33, 1935.
- [18] J. C. Slater, “A simplification of the hartree-fock method,” *Phys. Rev.*, vol. 81, pp. 385–390, Feb 1951.
- [19] P. Hohenberg and W. Kohn, “Inhomogeneous electron gas,” *Phys. Rev.*, vol. 136, pp. B864–B871, Nov 1964.
- [20] J. C. Slater and G. F. Koster, “Simplified lcao method for the periodic potential problem,” *Phys. Rev.*, vol. 94, pp. 1498–1524, Jun 1954.
- [21] Brillouin, L., “Les électrons libres dans les métaux et le rôle des réflexions de bragg,” *J. Phys. Radium*, vol. 1, no. 11, pp. 377–400, 1930.
- [22] W. Kohn, “Electrons, positrons and holes,” *Materials Research Bulletin*, vol. 5, no. 8, pp. 641–654, 1970.
- [23] M. Rohlfing and S. G. Louie, “Electron-hole excitations and optical spectra from first principles,” *Phys. Rev. B*, vol. 62, pp. 4927–4944, Aug 2000.
- [24] S. S. Li, *Semiconductor physical electronics*, pp. 247–503. Springer Science & Business Media, 2012.
- [25] A. K. Geim and K. S. Novoselov, “The rise of graphene,” *Nature Materials*, vol. 6, pp. 183–191, Mar 2007.
- [26] W. Zawadzki, “Semirelativity in semiconductors: A review,” *Journal of Physics: Condensed Matter*, vol. 29, no. 37, p. 373004, 2017.
- [27] L. Hrivnák, “Relativistic analogies in direct-gap semiconductors,” *Progress in Quantum Electronics*, vol. 17, no. 3, pp. 235–271, 1993.
- [28] N. Armitage, E. Mele, and A. Vishwanath, “Weyl and dirac semimetals in three dimensional solids,” *Reviews of Modern Physics*, vol. 90, 05 2017.
- [29] A. Bansil, H. Lin, and T. Das, “Colloquium: Topological band theory,” *Reviews of Modern Physics*, vol. 88, jun 2016.
- [30] J. Wang, S. Deng, Z. Liu, and Z. Liu, “The rare two-dimensional materials with Dirac cones,” *National Science Review*, vol. 2, pp. 22–39, 01 2015.
- [31] H. Weyl, “Elektron und gravitation. i,” *Zeitschrift für Physik*, vol. 56, pp. 330–352, May 1929.
- [32] S.-Y. Xu, I. Belopolski, N. Alidoust, M. Neupane, G. Bian, C. Zhang, R. Sankar, G. Chang, Z. Yuan, C.-C. Lee, S.-M. Huang, H. Zheng, J. Ma, D. S. Sanchez, B. Wang, A. Bansil, F. Chou, P. P. Shibayev, H. Lin, S. Jia, and M. Z. Hasan, “Discovery of a weyl fermion semimetal and topological fermi arcs,” *Science*, vol. 349, pp. 613–617, aug 2015.
- [33] L. Lu, Z. Wang, D. Ye, L. Ran, L. Fu, J. D. Joannopoulos, and M. Soljačić, “Experimental observation of weyl points,” *Science*, vol. 349, pp. 622–624, aug 2015.
- [34] S. Murakami, “Phase transition between the quantum spin hall and insulator phases in 3d: emergence of a topological gapless phase,” *New Journal of Physics*, vol. 9, pp. 356–356, sep 2007.
- [35] C. Kane, “Chapter 1 - topological band theory and the z2 invariant,” in *Topological Insulators* (M. Franz and L. Molenkamp, eds.), vol. 6 of *Contemporary Concepts of Condensed Matter Science*, pp. 3–34, Elsevier, 2013.
- [36] S. Pancharatnam, “Generalized theory of interference, and its applications,” *Proceedings of the Indian Academy of Sciences - Section A*, vol. 44, pp. 247–262, Nov 1956.
- [37] M. V. Berry, “Quantal phase factors accompanying adiabatic changes,” *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, vol. 392, no. 1802, pp. 45–57, 1984.

- [38] S. Murakami and S. ichi Kuga, “Universal phase diagrams for the quantum spin hall systems,” *Physical Review B*, vol. 78, oct 2008.
- [39] M. Z. Hasan, G. Chang, I. Belopolski, G. Bian, S.-Y. Xu, and J.-X. Yin, “Weyl, dirac and high-fold chiral fermions in topological quantum matter,” *Nature Reviews Materials*, vol. 6, pp. 784–803, apr 2021.
- [40] A. P. Sakhy, C.-Y. Huang, G. Dhakal, X.-J. Gao, S. Regmi, X. Yao, R. Smith, M. Sprague, B. Singh, H. Lin, S.-Y. Xu, F. Tafti, A. Bansil, and M. Neupane, “Observation of fermi arcs and weyl nodes in a non-centrosymmetric magnetic weyl semimetal,” 2022.
- [41] S. Jia, S.-Y. Xu, and M. Z. Hasan, “Weyl semimetals, fermi arcs and chiral anomalies,” *Nature Materials*, vol. 15, pp. 1140–1144, oct 2016.
- [42] K. Deng, G. Wan, P. Deng, K. Zhang, S. Ding, E. Wang, M. Yan, H. Huang, H. Zhang, Z. Xu, J. Denlinger, A. Fedorov, H. Yang, W. Duan, H. Yao, Y. Wu, S. Fan, H. Zhang, X. Chen, and S. Zhou, “Experimental observation of topological fermi arcs in type-ii weyl semimetal mote2,” *Nature Physics*, vol. 12, pp. 1105–1110, Dec 2016.
- [43] Z. Hasan, “What can you do with a weyl semimetal?,” Oct 2016.
- [44] B. Yan and C. Felser, “Topological materials: Weyl semimetals,” *Annual Review of Condensed Matter Physics*, vol. 8, no. 1, pp. 337–354, 2017.
- [45] A. A. Soluyanov and D. Vanderbilt, “Computing topological invariants without inversion symmetry,” *Phys. Rev. B*, vol. 83, p. 235401, Jun 2011.
- [46] L. Fu and C. L. Kane, “Topological insulators with inversion symmetry,” *Phys. Rev. B*, vol. 76, p. 045302, Jul 2007.
- [47] A. Tanaka and S. Takayoshi, “A short guide to topological terms in the effective theories of condensed matter,” *Science and Technology of Advanced Materials*, vol. 16, p. 014404, feb 2015.
- [48] H. A. Kramers, “Théorie générale de la rotation paramagnétique dans les cristaux,” *Proc. Acad. Amst*, vol. 33, no. 6, 1930.
- [49] S.-Q. SHEN, W.-Y. SHAN, and H.-Z. LU, “TOPOLOGICAL INSULATOR AND THE DIRAC EQUATION,” *SPIN*, vol. 01, pp. 33–44, jun 2011.
- [50] P. Roushan, J. Seo, C. V. Parker, Y. S. Hor, D. Hsieh, D. Qian, A. Richardella, M. Z. Hasan, R. J. Cava, and A. Yazdani, “Topological surface states protected from backscattering by chiral spin texture,” *Nature*, vol. 460, pp. 1106–1109, Aug 2009.
- [51] C.-K. Lee, J. H. Lee, B.-T. Zhang, and J.-L. He, “Chapter 2 - topological insulators and applications,” in *2D Materials for Nanophotonics* (Y. M. Jhon and J. H. Lee, eds.), Nanophotonics, pp. 81–138, Elsevier, 2021.
- [52] J. E. Moore, “The birth of topological insulators,” *Nature*, vol. 464, pp. 194–198, Mar 2010.
- [53] T. W. Schmitt, M. R. Connolly, M. Schleenvoigt, C. Liu, O. Kennedy, J. M. Chávez-Garcia, A. R. Jalil, B. Bennemann, S. Trellenkamp, F. Lentz, E. Neumann, T. Lindström, S. E. de Graaf, E. Berenschot, N. Tas, G. Mussler, K. D. Petersson, D. Grützmacher, and P. Schüffelgen, “Integration of topological insulator josephson junctions in superconducting qubit circuits,” *Nano Letters*, vol. 22, pp. 2595–2602, mar 2022.
- [54] M. S. Bahramy, B.-J. Yang, R. Arita, and N. Nagaosa, “Emergence of non-centrosymmetric topological insulating phase in bitei under pressure,” *Nature Communications*, vol. 3, p. 679, Feb 2012.
- [55] C. L. Kane and E. J. Mele, “Quantum spin hall effect in graphene,” *Phys. Rev. Lett.*, vol. 95, p. 226801, Nov 2005.
- [56] J. S. Lee, A. Richardella, R. D. Fraleigh, C.-x. Liu, W. Zhao, and N. Samarth, “Engineering the breaking of time-reversal symmetry in gate-tunable hybrid ferromagnet/topological insulator heterostructures,” *npj Quantum Materials*, vol. 3, p. 51, Oct 2018.

- [57] K. Ishizaka, M. S. Bahramy, H. Murakawa, M. Sakano, T. Shimojima, T. Sonobe, K. Koizumi, S. Shin, H. Miyahara, A. Kimura, K. Miyamoto, T. Okuda, H. Namatame, M. Taniguchi, R. Arita, N. Nagaosa, K. Kobayashi, Y. Murakami, R. Kumai, Y. Kaneko, Y. Onose, and Y. Tokura, “Giant rashba-type spin splitting in bulk bsite,” *Nature Materials*, vol. 10, pp. 521–526, Jul 2011.
- [58] C. Bradley and A. Cracknell, *The mathematical theory of symmetry in solids: representation theory for point groups and space groups*, pp. 95–119. Oxford University Press, 2010.
- [59] S. Datta and B. Das, “Electronic analog of the electro-optic modulator,” *Applied Physics Letters*, vol. 56, no. 7, pp. 665–667, 1990.
- [60] J. Vamathevan, D. Clark, P. Czodrowski, I. Dunham, E. Ferran, G. Lee, B. Li, A. Madabhushi, P. Shah, M. Spitzer, and S. Zhao, “Applications of machine learning in drug discovery and development,” *Nature Reviews Drug Discovery*, vol. 18, pp. 463–477, Jun 2019.
- [61] H. Gerdes, P. Casado, A. Dokal, M. Hijazi, N. Akhtar, R. Osuntola, V. Rajeeve, J. Fitzgibbon, J. Travers, D. Britton, S. Khorsandi, and P. R. Cutillas, “Drug ranking using machine learning systematically predicts the efficacy of anti-cancer drugs,” *Nature Communications*, vol. 12, p. 1850, Mar 2021.
- [62] A. Lavecchia, “Machine-learning approaches in drug discovery: methods and applications,” *Drug Discovery Today*, vol. 20, no. 3, pp. 318–331, 2015.
- [63] S. Dara, S. Dhamercherla, S. S. Jadav, C. M. Babu, and M. J. Ahsan, “Machine learning in drug discovery: A review,” *Artificial Intelligence Review*, vol. 55, pp. 1947–1999, Mar 2022.
- [64] S. Hochreiter, G. Klambauer, and M. Rarey, “Machine learning in drug discovery,” *Journal of Chemical Information and Modeling*, vol. 58, no. 9, pp. 1723–1724, 2018. PMID: 30109927.
- [65] V. Stanev, K. Choudhary, A. G. Kusne, J. Paglione, and I. Takeuchi, “Artificial intelligence for search and discovery of quantum materials,” *Communications Materials*, vol. 2, p. 105, Oct 2021.
- [66] A. Einstein, “The Foundation of the General Theory of Relativity,” *Annalen Phys.*, vol. 49, no. 7, pp. 769–822, 1916.
- [67] P. Liashchynskyi and P. Liashchynskyi, “Grid search, random search, genetic algorithm: A big comparison for nas,” 2019.
- [68] J. Bergstra and Y. Bengio, “Random search for hyper-parameter optimization.,” *Journal of machine learning research*, vol. 13, no. 2, 2012.
- [69] M. del Rosario, “An overview of stochastic gradient descent in machine learning,” 2019.
- [70] S. Ruder, “An overview of gradient descent optimization algorithms,” 2016.
- [71] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” 2014.
- [72] S. Salman and X. Liu, “Overfitting mechanism and avoidance in deep neural networks,” 2019.
- [73] A. A. Burkov, M. D. Hook, and L. Balents, “Topological nodal semimetals,” *Physical Review B*, vol. 84, dec 2011.
- [74] V. Nasteski, “An overview of the supervised machine learning methods,” *HORIZONS.B*, vol. 4, pp. 51–62, 12 2017.
- [75] Y. Zhang and E.-A. Kim, “Quantum loop topography for machine learning,” *Physical Review Letters*, vol. 118, may 2017.
- [76] P. Zhang, H. Shen, and H. Zhai, “Machine learning topological invariants with neural networks,” *Physical Review Letters*, vol. 120, feb 2018.

- [77] G. Pilania, C. Wang, X. Jiang, S. Rajasekaran, and R. Ramprasad, “Accelerating materials property predictions using machine learning,” *Scientific Reports*, vol. 3, p. 2810, Sep 2013.
- [78] S. Stanisic, J. L. Bosse, F. M. Gambetta, R. A. Santos, W. Mruczkiewicz, T. E. O’Brien, E. Ostby, and A. Montanaro, “Observing ground-state properties of the fermi-hubbard model using a scalable algorithm on a quantum computer,” *Nature Communications*, vol. 13, p. 5743, Oct 2022.