

Deep learning approach to searching flat bands in 3D materials

Mikolaj J. Gawkowski
10636642

School of Physics and Astronomy, The University of Manchester, Manchester M13 9PL, United Kingdom

(This work was done in collaboration with Mateusz Czajka, 10639338)

(Dated: January 7, 2024)

Electron-electron interactions play a crucial role in solid state and condensed matter physics, as they give rise to fascinating phenomena such as superconductivity, magnetism, and Wigner crystals. These interactions are particularly enhanced in flat bands within band structures. The search for materials with flat bands that host exotic physics has been greatly facilitated by the creation of online databases containing vast arrays of different materials. In this study, we used supervised machine learning to automate the classification of 3D materials with flat bands. Specifically, we used a convolutional neural network to identify materials, taken from the Materials Project database, that exhibit flat bands. We found 30,298 materials with some flat segments of their band structures next to the Fermi level of the 60,548 materials considered. Furthermore, from this subset of materials, we identified 15,597 plane-flat materials with flat bands throughout their Brillouin zones as well as 5,651 high-quality plane-flat materials that are extremely flat in the vicinity of the Fermi level. Our research enables an efficient approach to identifying new material species that could contribute to breakthroughs in the physics of strongly correlated materials.

I. INTRODUCTION

The electronic interactions in materials are influenced by their energy spectra, which can be described using band theory and are well-represented through band structure diagrams. Some electrons exhibit dispersionless spectra, indicating that their energies E are independent of momentum \mathbf{k} , i.e. $\nabla_{\mathbf{k}}E \approx 0$ which is equivalent to constant energy within a certain bandwidth. The suppression of electronic kinetic energy renders these electrons more susceptible to electron-electron interactions [1]. The regions where many-body effects prevail over kinetic energy are commonly referred to as 'flat bands'.

The presence of flat bands can lead to various interesting physical phenomena. An example of this can be seen in twisted bilayer graphene, where the unconventional superconductivity observed is believed to be a result of the topological non-triviality of these flat bands [2, 3]. The study of the topological non-triviality of flat bands is an ongoing area of research, and it has been shown that flat bands can be characterised by a non-zero Chern number [4]. Additionally, when the flat bands are partially filled, fractional quantum Hall states can arise at high temperatures [5]. Hence, it is anticipated that flat bands exhibiting nontrivial topology will have a significant impact on achieving the integer and fractional quantum Hall effect and fractional Chern insulator phase [6–8].

Furthermore, flat bands are expected to contribute to the emergence of various phenomena, including Wigner crystallisation [9], non-Fermi liquid behaviour [10], the fractional quantum Hall effect [11], Hubbard ferromagnetism [12, 13], and more. Many of these phenomena have been reported in two-dimensional (2D) materials. Furthermore, similar phenomena can be identified in the electronic structures of three-dimensional (3D) materials. Recent reports include the discovery of 3D flat bands in the pyrochlore metal CaNi_2 , with the pyrochlore system

itself attracting attention as a potential host for superconductivity [14, 15]. It has also been demonstrated that single crystals of $\text{Ca}(\text{Rh}_{1-x}\text{Ru}_x)_2$ exhibit superconductivity [16]. Moreover, research has shown that stacking van der Waals materials enables the engineering of 3D moiré flat bands with unconventional superconducting states, realised as a function of twist angle in graphitic systems, WSe_2 , and hexagonal boron nitride [17]. Flat bands in 3D materials are also expected to exhibit perfect ferromagnetism [18–21].

Recent research suggests the effectiveness of machine learning (ML) methods in the identification of flat electronic bands in 2D materials [22]. In this study, we extend this approach to 3D materials, leveraging a key advantage of this expansion: the greater availability of 3D materials compared to their 2D counterparts. As a result, we can predict a considerable number of new materials that may be classified as flat bands, potentially exhibiting exotic physics. This methodology also facilitates more meaningful statistical analyses given the abundance of available 3D materials.

Recent advances in ML have significantly eased the traditionally challenging task of manually searching for flat bands in 3D materials. This is largely attributed to the emergence of online materials science databases such as Materials Project [23], Aflow [24], and the Topological Materials Database [25–27]. These databases grant access to vast repositories of potential 3D materials, thereby facilitating more statistically significant and accurate predictions. To classify materials, we employ band structure images sourced from an online database. This approach presents an alternative to prior research [28, 29], where the screening of materials with flat bands relied on parameterized bands and predefined bandwidth. Such methods are susceptible to overlooking certain flat bands, especially when considering band crossings [22]. Our method provides an efficient and accurate means

of classifying materials with flat bands. The increasing availability of online databases containing flat-band materials [30, 31] allows us to benchmark our results, identifying both the strengths and weaknesses of our approach.

A convolutional neural network (CNN) trained on a designated training set obtained from the Materials Project website was employed to identify a set of flat-band materials across the entire database. The Materials Project database was accessed using an application programming interface (API), providing not only access to band structures but also to corresponding density of states and other useful properties. The identified flat-band materials underwent subsequent statistical analysis, aiming to pinpoint materials with the most favorable flat bands—those with the potential to host exotic physics, such as superconductivity.

II. THEORY

A. Neural networks

Here, we provide a brief introduction to fundamental concepts in neural networks, focusing on a specific subset known as convolutional neural networks.

To define a neural network, we must first introduce the concept of a neuron. A neuron is a non-linearity applied to functions of the form [32]:

$$f(\mathbf{x}) = A_1x_1 + \dots + A_nx_n + b, \quad (1)$$

where the input $\mathbf{x} = (x_1, x_2, \dots, x_n)$ is a vector of scalar inputs. This input is then passed through a function of the form:

$$T(\mathbf{x}) = \phi \left(\sum_i W_i x_i + b \right) = \phi(\mathbf{W} \cdot \mathbf{x} + \mathbf{b}), \quad (2)$$

with $\phi(\dots)$ representing the neuron and weights \mathbf{W} , and bias \mathbf{b} . A typical non-linearity, also referred to as an activation function, is the sigmoid function given by:

$$\sigma(x) = \frac{1}{1 + e^{-x}}. \quad (3)$$

A neural network is composed of multiple neurons connected in a graph and is typically organised in layers. A schematic of a neural network is shown in Fig. 1. These layers include an input layer, one or more hidden layers, and an output layer. It is common to refer to a network as an N -layer network if it consists of N layers, excluding the input layer. To make predictions with a neural network, values for the parameters \mathbf{W} and \mathbf{b} need to be determined, and a function is defined to map the network's output to a prediction. These parameters, known as trainable parameters, determine the size of the neural network and play a crucial role in its predictive abilities.

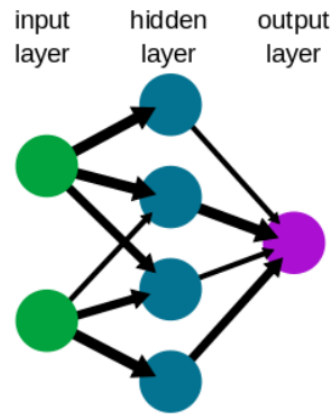


FIG. 1: Neural network with 1 input layer of 2 neurons, and one hidden layer of 4 neurons. The hidden layer has $2 \cdot 3$ weights, and 4 biases. The output layer has 4 weights and 1 bias. In total, this network has 15 learnable parameters.

1. Convolutional neural networks

CNNs operate similarly to standard neural networks, featuring neurons with learnable weights and biases. The key distinction lies in CNNs being specifically designed to process grid-like data, such as images [33]. This specialization is achieved through convolution operations, introducing translational invariance, where filters are independent of location. This design significantly reduces the number of parameters, making the network more manageable to optimize and less dependent on data size.

In CNNs, neurons are arranged in several dimensions: channels, width, height, and the number of filters in the simplest 2D case. Similarly to standard neural networks, CNNs consist of multiple layers, with each layer transforming activations or outputs from the previous layer through a differentiable function. Commonly used layers include convolution layers, pooling layers, and fully connected layers, all contributing to a comprehensive convolutional layer.

In CNNs, convolutions are a key operation, with the convolution of two functions g and h defined as

$$f(x) = (g * h)(x) = \int g(x)h(x' - x)dx'. \quad (4)$$

Here, $g * h$ denotes the convolution of the two functions. In the context of CNNs, g is termed the input, h is the filter, and f is referred to as activation. Eq. (4) assumes that g and h are continuous functions. In a more practical setting, a discrete convolution is often used:

$$f(x) = (g * h)(x) = \sum_{x'} g(x)h(x - x'). \quad (5)$$

Here, $x, x' \in \mathbb{R}^n$ for both continuous and discrete cases.

This convolution operation allows the CNN to learn filters that recognise visual information, such as edges or orientations. As the network progresses to higher layers, these filters can combine to recognise more complex patterns. Each convolution layer in the network comprises an entire set of such filters.

B. Fundamentals of wave mechanics

In this segment, we lay the groundwork by delving into key principles of wave mechanics crucial for our exploration of band structures and density functional theory (DFT). Our focus here intentionally excludes considerations of time-dependent interactions among atoms and molecules, allowing us to pivot to the examination of a time-independent Schrödinger equation. Specifically, for an isolated atomic or molecular system with N electrons within the Born-Oppenheimer nonrelativistic approximation, this equation is expressed as [34]:

$$\hat{H}\Psi = E\Psi, \quad (6)$$

where E is the electronic energy, $\Psi = \Psi(\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_N)$ is the wave function with coordinates \mathbf{q}_i comprising space coordinates \mathbf{r}_i and spin coordinates s_i , and \hat{H} is the Hamiltonian operator, given by

$$\hat{H} = \sum_i^N -\frac{1}{2}\nabla_i^2 + \sum_i^N v(\mathbf{r}_i) + \sum_{i<j}^N \frac{1}{r_{ij}}, \quad (7)$$

where $v(\mathbf{r}_i) = -\sum_\alpha \frac{Z_\alpha}{r_{i\alpha}}$ is the potential acting on the electron i from the nuclei of charges Z_α . Here, we are using atomic units, where $e = m_e = \hbar = 1$.

We can write Eq. (7) more compactly as

$$\hat{H} = \hat{T} + \hat{V}_{ne} + \hat{V}_{ee}, \quad (8)$$

where

$$\hat{T} = \sum_i^N -\frac{1}{2}\nabla_i^2 \quad (9)$$

is the kinetic energy operator,

$$\hat{V}_{ne} = \sum_i^N v(\mathbf{r}_i) \quad (10)$$

is the electron-nucleus attraction energy operator, and

$$\hat{V}_{ee} = \sum_{i<j}^N \frac{1}{r_{ij}} \quad (11)$$

is the electron-electron repulsion energy operator.

Expectation value of the Hamiltonian is given by

$$E[\Psi] = \langle \hat{H} \rangle = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \quad (12)$$

where we used Dirac bra-ket notation. We also have that $E[\Psi] \geq E_0$ with $E_0 = \min_\Psi E[\Psi]$ where we used the variational principle $\delta E[\Psi] = 0$. We can express the variational principle as a problem involving the method of Lagrange multipliers where we extremize $\langle \Psi | \hat{H} | \Psi \rangle$ subject to the normalization constraint $\langle \Psi | \Psi \rangle$. This is equivalent to the following statement

$$\delta[\langle \Psi | \hat{H} | \Psi \rangle - E \langle \Psi | \Psi \rangle] = 0, \quad (13)$$

with E the Lagrange multiplier. This procedure allows us to obtain the ground-state energies and other properties by starting from N and $v(\mathbf{r})$. Note that this method works for the excited states if and only if the excited-state wave function is orthogonal to the exact ground-state wave function.

Suppose Ψ is approximated as an antisymmetrized product of N orthonormal spin orbitals $\psi_i(\mathbf{q})$, each a product of a spatial orbital $\phi(\mathbf{r})$ and a spin function $\sigma(s) = \alpha(s)$ or $\beta(s)$, the Slater determinant

$$\Psi_{\text{HF}} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{q}_1) & \dots & \psi_N(\mathbf{q}_1) \\ \vdots & & \vdots \\ \psi_1(\mathbf{q}_N) & \dots & \psi_N(\mathbf{q}_N) \end{vmatrix} \quad (14)$$

where we call Ψ_{HF} the Hartree-Fock wave function. The Hartree-Fock approximation relies on minimizing the following energy functional

$$E_{\text{HF}} = \sum_i^N H_i + \frac{1}{2} \sum_{i,j=1}^N (J_{ij} - K_{ij}), \quad (15)$$

with

$$H_i = \int \psi_i^*(\mathbf{q}) \left[-\frac{1}{2}\nabla^2 + v(\mathbf{q}) \right] \psi_i(\mathbf{q}) d\mathbf{q}, \quad (16)$$

$$J_{ij} = \iint \psi_i^*(\mathbf{q}_1) \psi_i(\mathbf{q}_1) \frac{1}{r_{12}} \psi_j^*(\mathbf{q}_2) \psi_j(\mathbf{q}_2) d\mathbf{q}_1 d\mathbf{q}_2, \quad (17)$$

$$K_{ij} = \iint \psi_i^*(\mathbf{q}_1) \psi_j(\mathbf{q}_1) \frac{1}{r_{12}} \psi_i(\mathbf{q}_2) \psi_j^*(\mathbf{q}_2) d\mathbf{q}_1 d\mathbf{q}_2. \quad (18)$$

These integrals are all real, and $J_{ij} \geq K_{ij} \geq 0$. The J_{ij} are called Coulomb integrals, K_{ij} are called exchange integrals. We also have that $J_{ii} = K_{ii}$.

Minimization of Eq. (15) yields the Hartree-Fock equations [34]

$$\hat{F}\psi_i(\mathbf{q}) = \sum_{j=1}^N \epsilon_{ij} \psi_j(\mathbf{q}), \quad (19)$$

where

$$\hat{F} = -\frac{1}{2}\nabla^2 + v + \hat{g}, \quad (20)$$

and $\hat{g} = \hat{j} - \hat{k}$ with,

$$\hat{j}(\mathbf{q}_1)f(\mathbf{q}_1) = \sum_{k=1}^N \int \psi_k^*(\mathbf{q}_2)\psi_k(\mathbf{q}_2) \frac{1}{r_{12}} f(\mathbf{q}_1) d\mathbf{q}_2, \quad (21)$$

$$\hat{k}(\mathbf{q}_1)f(\mathbf{q}_1) = \sum_{k=1}^N \int \psi_k^*(\mathbf{q}_2)f(\mathbf{q}_2) \frac{1}{r_{12}} \psi_k(\mathbf{q}_1) d\mathbf{q}_2 \quad (22)$$

and the Hermitian matrix ϵ contains Lagrange multipliers associated with the orthonormalization condition of the orbitals ψ_i . Eq. (19) must be solved iteratively, since the orbitals that solve the equation, appear in the operator \hat{F} . For this reason, the Hartree-Fock method is a 'self-consistent' approach. The Hartree-Fock solution serves as the primary starting point for many methods aiming to offer a more accurate description of the many-electron system, for both atoms and molecules. However, its main weakness lies in neglecting electronic correlation, a consequence of using Ψ_{HF} as the electronic wave function.

We can also define the electron density $\rho(\mathbf{r})$ in terms of Ψ

$$\rho(\mathbf{r}_1) = N \int \dots \int |\Psi(\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_N)|^2 ds_1 d\mathbf{q}_2 \dots d\mathbf{q}_N, \quad (23)$$

and it must integrate to the total number of electrons

$$\int \rho(\mathbf{r}) d\mathbf{r} = N. \quad (24)$$

1. Band structures

Here we give a brief overview of the mathematical representation of electrons within a periodic potential. This approach, known as the one-electron approximation, focuses on electrons that occupy a lattice. A key outcome of this exploration is the segmentation of the energy spectrum of one-electron states into bands, with gaps between them where no allowed states exist. This configuration is commonly termed an energy band structure.

The Schrödinger equation for electrons in a periodic potential is given by:

$$\hat{H}\Psi(\mathbf{r}) = \left[-\frac{1}{2}\nabla^2 + \hat{V}(\mathbf{r}) \right] \Psi(\mathbf{r}) = E\Psi(\mathbf{r}), \quad (25)$$

where $\hat{V}(\mathbf{r})$ includes parts of the electron-electron interaction as well as the potential of the ions. We also assume that this Hamiltonian is invariant under primitive translations.

We now introduce an operator $T_{\mathbf{R}_l}$ defined as

$$T_{\mathbf{R}_l}f(\mathbf{r}) = f(\mathbf{r} + \mathbf{R}_l), \quad (26)$$

that is, $T_{\mathbf{R}_l}$ is a translation operator whose effect is to displace the position \mathbf{r} by \mathbf{R}_l . From the assumption of the translation invariance of the Hamiltonian, we can write

$$T_{\mathbf{R}_l}(H\Psi_n) = T_{\mathbf{R}_l}(E_n\Psi_n) \quad (27)$$

and hence

$$H(T_{\mathbf{R}_l}\Psi_n) = E_n(T_{\mathbf{R}_l}\Psi_n). \quad (28)$$

This implies that $T_{\mathbf{R}_l}\Psi_n$ and Ψ_n are eigenfunctions of the same eigenvalue E_n . If E_n has only one eigenfunction Ψ_n , then $T_{\mathbf{R}_l}\Psi_n$ must, apart from some factor, be equal to Ψ_n . Additionally, since $|T_{\mathbf{R}_l}\Psi_n|^2$ must be equal to $|\Psi_n|^2$, this factor must have the absolute value of 1:

$$T_{\mathbf{R}_l}\Psi_n = \lambda^{(1)}\Psi_n \quad \text{with} \quad |\lambda^{(1)}|^2 = 1. \quad (29)$$

We can see that $\lambda^{(1)}$ are eigenvalues of $T_{\mathbf{R}_l}$ and that they can be written as $\lambda^{(1)} = \exp(i\mathbf{k} \cdot \mathbf{R}_l)$ where \mathbf{k} is the wave vector.

If, in turn, E_n is degenerate, we can still find \mathbf{k} for every state Ψ such that Ψ has the eigenvalue $\exp(i\mathbf{k} \cdot \mathbf{R}_l)$. We can combine these results in the famous Bloch's theorem [35].

Theorem 1 (Bloch's theorem). *The nondegenerate solutions of the Schrödinger equation given in Eq. (25) and suitably chosen linear combinations of the degenerate solutions are at the same time eigenfunctions $\Psi_n(\mathbf{k}, \mathbf{r})$ of the translational operator $T_{\mathbf{R}_l}$ with the eigenvalues $\exp(i\mathbf{k} \cdot \mathbf{R}_l)$:*

$$T_{\mathbf{R}_l}\Psi_n(\mathbf{k}, \mathbf{r}) = \exp(i\mathbf{k} \cdot \mathbf{R}_l)\Psi_n(\mathbf{k}, \mathbf{r}). \quad (30)$$

The $\Psi_n(\mathbf{k}, \mathbf{r})$ are also referred to as the Bloch states. We can also see that the eigenvalues E_n also depend on \mathbf{k} , i.e., $E_n = E_n(\mathbf{k})$. We now establish a reciprocal lattice \mathbf{K}_m in the space of the vector \mathbf{k} . We also have that $\mathbf{K}_m \cdot \mathbf{R}_l$ is an integer multiple of 2π . Operating on $\Psi(\mathbf{k} + \mathbf{K}_m, \mathbf{r})$ with $T_{\mathbf{R}_l}$ gives

$$T_{\mathbf{R}_l}\Psi(\mathbf{k} + \mathbf{K}_m, \mathbf{r}) = \exp(i\mathbf{k} \cdot \mathbf{R}_l)\Psi(\mathbf{k} + \mathbf{K}_m, \mathbf{r}). \quad (31)$$

This implies that we can have multiple values of \mathbf{k} associated with one $\Psi(\mathbf{r})$. All such points in \mathbf{k} -space are equivalent:

$$\Psi_n(\mathbf{k}, \mathbf{r}) = \Psi_n(\mathbf{k} + \mathbf{K}_m, \mathbf{r}). \quad (32)$$

This means that for every fixed \mathbf{k} , $E_n(\mathbf{k})$ gives a discrete energy spectrum. For a given n , $E_n(\mathbf{k})$ within the Brillouin zone—defined as the primitive cell in the reciprocal lattice—is a continuous and (apart from degenerate points) differentiable function of \mathbf{k} [36]. Such $E_n(\mathbf{k})$ is then called an energy band. A set of all bands, i.e., the complete function $E_n(\mathbf{k})$, is referred to as the band structure. The pictorial representation of band structures involves plotting $E_n(\mathbf{k})$ against the high-symmetry points of the Brillouin zone, with these high-symmetry points also referred to as critical points [37]. An example band structure is shown in Fig. 5a, where Γ corresponds to the center of the Brillouin zone, and other letters represent other points of symmetry within the Brillouin zone. In the band structure diagram, the value of $E - E_f = 0$ is referred to as the Fermi level, with E_f being the Fermi energy.

C. Density functional theory

The majority of Materials Project data are obtained through density functional theory, an electronic structure theory extensively applied in condensed matter physics and materials science. DFT facilitates the exploration of the electronic structure of many-body systems. At its core, DFT is based on the concept that a complex N -electron wave function and its associated Schrödinger equation can be replaced with the considerably simpler electron density. In DFT, the entire ground state wave function is treated as a functional of this electron density [34]. The primary advantage of DFT over the Hartree-Fock method is its inclusion of correlation energies, albeit approximately.

DFT is based on two fundamental theorems due to Hohenberg and Kohn [38]:

Theorem 2. *The external potential $v(\mathbf{r})$ is determined, within a trivial additive constant, by the electron density $\rho(\mathbf{r})$.*

Proof. The proof is given in App. A. ■

Theorem 3. *For a trial density $\bar{\rho}(\mathbf{r})$, such that $\bar{\rho}(\mathbf{r}) \geq 0$ and $\int \bar{\rho}(\mathbf{r}) d\mathbf{r} = N$,*

$$E_0 \leq E_v[\bar{\rho}]. \quad (33)$$

Proof. The proof is given in App. A. ■

Put differently, the first theorem establishes that the external potential is uniquely determined by the electron density, while the second equation asserts that the input density serves as the true ground-state density if and only if the functional determining the ground-state energy, yields the lowest energy.

The Hohenberg-Kohn equations offer the advantage of directly handling electronic density rather than wave functions. However, these equations have a limitation—they do not provide a practical method for solving the electronic structure of many-electron systems.

This is where the Kohn-Sham equations become crucial. Formulated by Kohn and Sham in 1965, they have since revolutionized electronic structure calculations by offering a practical and computationally feasible approach. Kohn and Sham’s innovative idea was to introduce a kinetic energy functional $T_s[\rho]$, which is a functional of the electron density $\rho(\mathbf{r})$ and it describes N non-interacting electrons. After functional differentiation, $T_s[\rho]$ leads to a set of one-particle equations, and the solutions to these equations are the Kohn-Sham orbitals.

The Kohn-Sham orbitals $\{\psi_i\}$ satisfy the Kohn-Sham

equations given by [39]:

$$\left[-\frac{1}{2}\nabla^2 + v_{\text{eff}}\right] \psi_i = \epsilon_i \psi_i, \quad (34)$$

$$v_{\text{eff}}(\mathbf{r}) = v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{xc}(\mathbf{r}), \quad (35)$$

$$\rho(\mathbf{r}) = \sum_i \sum_s |\psi_i(\mathbf{r}, s)|^2, \quad (36)$$

where ϵ_i are the corresponding orbital energies, $v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})}$ with $E_{xc}[\rho]$ the exchange-correlation functional describing the difference between $T_s[\rho]$ and the exact formula for the ground-state kinetic energy.

At this point, the problem has been reduced to finding the best approximation to the unknown functional $E_{xc}[\rho]$. It is also important to point to the fact that Kohn-Sham orbitals do not represent actual electronic wave functions, but are often associated with Bloch energy bands for interacting systems [40].

Electron density in DFT can be evaluated using a self-consistent cycle. This cycle starts by assuming a certain electronic density, which is then plugged into the Kohn-Sham equations. From the Kohn-Sham equations, we can get a new electronic density, which we can compare with our initial guess. If the two are more or less the same, we are done. If the two densities are different, we need to appropriately modify our initial density. This way, we can evaluate a self-consistent electronic structure.

Another important aspect of DFT is the usage of pseudopotentials. Pseudopotentials are based on the concept of replacing the entire electronic system with only the valence electrons. These valence electrons are the most significant ones in terms of participating in bonding and exhibiting interesting physics phenomena. By "freezing" the core electrons, pseudopotentials effectively substitute the Coulomb potential with an effective potential. This approach offers the advantage of enabling the utilisation of plane-wave basis sets, which greatly simplify calculations in computational simulations such as DFT.

III. METHODS

A. Materials Project database

The Materials Project stands as a comprehensive open online database, consisting of over 600,000 entries encompassing crystal structures, molecules, band structures, density of states, tensor properties, X-ray absorption near-edge structure, extended X-ray absorption fine structure, and charge densities. Notably, approximately 60,000 materials include band structure data, as not all entries have corresponding information available. The Materials Project database, along with other material databases, employs DFT to generate various material properties, such as electronic structures and density of states. To perform DFT calculations, Materials Project

utilises the Vienna Ab Initio Simulation Package software [41]. In DFT calculations, Materials Project also utilises pseudopotentials.

The exchange-correlation functional in Materials Project is a combination of generalised gradient approximation (GGA) and GGA+U, or a mixture of GGA, GGA+U, and r2SCAN [42–44]. The projector augmented wave (PAW) method, which is a generalisation of pseudopotentials, is used for GGA and GGA+U calculations to model core electrons, with an energy cutoff of 520 eV [45]. For all computations, a baseline k-point mesh of 1000 divided by the number of atoms in the cell is utilised. The Monkhorst-Pack method is employed for selecting k-points, with a γ -centered approach for hexagonal cells, and the tetrahedron method is used for k-point integration. r2SCAN also utilizes the PAW method for modeling core electrons, with an energy cutoff of 680 eV, and k-point grids ranging from $0.22/\text{\AA}$ to $0.44/\text{\AA}$. All calculations are performed at 0 K and 0 atm and are performed with spin polarization on and with magnetic ions in a high-spin ferromagnetic initialization. Input structures are sourced from many different places, including the Inorganic Crystal Structure Database (ICSD) [46]. The cell and atomic positions are relaxed in the calculation two times in consecutive runs. When multiple crystal structures are present for a single chemical composition, Materials Project evaluates all unique structures determined by an affine mapping technique [47].

Because DFT is formulated to calculate ground-state properties, it does not allow for a completely rigorous analysis of the properties of excited states, and there is no guarantee that Kohn-Sham eigenvalues will correspond to physical observables. In practise, however, band structures provide useful and applicable information, and in general, band structures calculated using DFT match those obtained experimentally. The error in the band dispersion accuracy has been reported to range from 0.1 to 0.4 eV [48]. However, predicted band gaps are usually severely underestimated, with the disagreement between simulations and experiment being as high as 40% [49] and the errors originate mostly from the approximations to the exchange-correlation functional. One way of interpreting DFT band structures is to shift the conduction bands in energy by a constant amount so that the band gaps match known experimental observations [42]. One of the methods that could improve the DFT evaluation of the band gaps relies on using hybrid functionals or using ML techniques [50, 51].

Band structure data from the Materials Project are stored in JavaScript Object Notation (JSON) files, accessible through an API. MPRester, a Python client provided by the Materials Project, facilitates easy access to these data. Complementing this, Python Materials Genomics (Pymatgen) is a robust open-source Python library designed for materials analysis, with a focus on electronic structure analyses like density of states and band structures.

To gather data, we utilize the Materials Project API to

download pertinent information for each available band structure from the website. This includes details such as Fermi energy, k-points, bands, and reciprocal lattice information. Discarding extraneous data from JSON files not only reduces their size, but also accelerates the download process. Following retrieval, these JSON files are converted into Portable Network Graphics files for further analysis.

B. Convolutional neural network and image segmentation

The CNN used for identifying flat segments in band structure images consists of six layers, as illustrated in Fig. 2, following the architecture described in [22]. With a total of 6.9 million learnable parameters, the network processes 96x96 pixel images and generates a flatness score in the range $[0,1]$. The CNN was trained using MATLAB and accessed through Python using the MATLAB Engine API for Python.

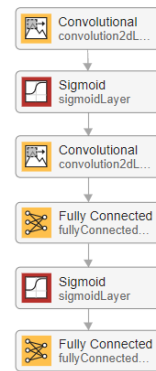


FIG. 2: Layers of the convolutional neural network used for assigning flatness scores of band structures.

The initial layer involves 30 channels of 10×10 filters with a $[1,1]$ stride. Subsequently, the second layer adopts the sigmoid activation function, chosen for its superior performance given the network’s output range of $[0,1]$. The third layer is another convolutional layer, featuring 12 channels of 3×3 filters with a $[1,1]$ stride. Following this, a fully connected layer with 80 output nodes is employed. This is succeeded by another sigmoid layer, leading to a final fully connected layer with a single node. The network ultimately returns a real number within the interval $[0,1]$, serving as the flatness measure for each segment within the band structure.

Table I provides a comprehensive overview of the training parameters utilised in the CNN training process, including batch size, number of epochs, loss function, and optimisation algorithm.

The training dataset for the network comprised approximately 1900 band structure images sourced from the Materials Project database. Subsequently, the network underwent testing using approximately 7000 image

TABLE I: Training parameters used in the training of the CNN along with their values as given here [22].

Training parameter	Value
batch size	128
number of epochs	10
gradient decay factor	0.9
loss function	'mean squared error'
optimization algorithm	'adam'
learning rate	0.00005
L2Regularization	0.003
LearnRateDropFactor	0.1

segments, also obtained from the same database. The predictions in the test set showed a commendable match of 92.1% with manual selection. Notably, false negative predictions accounted for 3.3% of cases, while false positive were observed in 4.6% of instances [22].

Following the retrieval of images from the Materials Project database, a segmentation process was applied. Each image was segmented along the high-symmetry points, focusing on an energy range of [-1,1] eV with respect to the Fermi level for this study. The chosen energy range was further divided into four segments of 0.5 eV each. Consequently, each image can be conceptualised as an array of segments with four rows, and the number of columns matching the count of the directions in the Brillouin zone (number of intervals between the high-symmetry points).

These segmented arrays are then introduced into the network. The network assigns a real number from the range [0,1] to each segment, where 0 signifies the absence of flatness, and 1 indicates a fully flat segment. Subsequently, a flatness score threshold is defined to classify whether a particular segment is flat or not. Of particular interest in this study is the identification of materials featuring flat bands, specifically those with at least one row (bandwidth) where all segments are flat. In other words, we focus on identifying plane-flat materials exhibiting flat bands across all planes of their reciprocal spaces. An example of a plane-flat band structure is shown in Fig. 5.

For each material, we introduce a compound flatness score given by

$$\mathcal{F} = \max_i F_i, \quad (37)$$

where

$$F_i = \frac{\sum_{j=1}^{N_i} f_{ij}}{N_i}. \quad (38)$$

F_i represents the averaged flatness score for each row, with each segment having its flatness score f_{ij} and N_i being the number of segments in the i^{th} row. In cases where rows have identical flatness scores, priority is given to the row closer to the Fermi level. The compound flatness score is then used to establish a threshold value for the flatness criterion.

In principle, plane-flat materials can be defined in two ways. Material may be deemed plane-flat if it contains at least one row of flat segments, meaning all segments in a given row have flatness scores greater than a specified flatness criterion, denoted as ϕ . Mathematically, this is equivalent to representing the band structure \mathbf{B} as a 4-row array with n columns. Each element \mathbf{B}_{ij} , where $i = 1, 2, 3, 4$ and $j = 1, 2, \dots, n$, must satisfy

$$\exists \mathbf{B}_{i,*} \text{ such that } \forall \mathbf{B}_{ij} \in \mathbf{B}_{i,*}, f_{ij} \geq \phi. \quad (39)$$

Here, $\mathbf{B}_{i,*}$ denotes the i^{th} row of \mathbf{B} .

An alternative approach to defining a plane-flat material is based on its compound flatness score, denoted as \mathcal{F} , satisfying $\mathcal{F} \geq \phi$. In other words, the material's compound flatness criterion should be at least as large as a specified threshold value, represented by the flatness criterion. However, this method has certain drawbacks when compared to the first one. It might categorise a band as plane-flat even if one or more of its segments are not flat. For instance, in Fig. 3, a flatness prediction is observed where all but one segment is flat. The compound flatness score for this case is $\mathcal{F} \approx 0.9638$, and if the flatness criterion is set at 0.96, it could classify this as a flat row despite not being flat in all directions of the Brillouin zone.

On the other hand, the disadvantage of using the definition based on materials with at least one fully flat row is that it tends to overlook materials and may fail to identify those that are indeed plane-flat. For example, setting $\phi = 0.99$ for the material shown in Fig. 4 would not consider it as plane-flat since it contains two segments for which $f_{ij} < 0.99$. According to this definition, the corresponding row is not fully composed of flat segments, despite the fact that the row is extremely flat. If we instead choose the other definition, the material's compound flatness score $\mathcal{F} \approx 0.996 \geq 0.99$, allowing us to classify this as a plane-flat material.

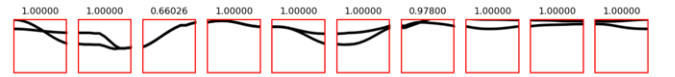


FIG. 3: A part of the flatness prediction where all but one segment is flat and the band not being extremely flat.

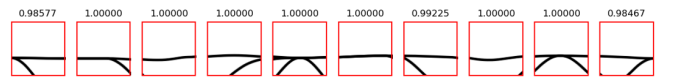


FIG. 4: A portion of the flatness prediction shows that not all segments are classified as flat within the given energy bandwidth, despite the overall flatness of the band.

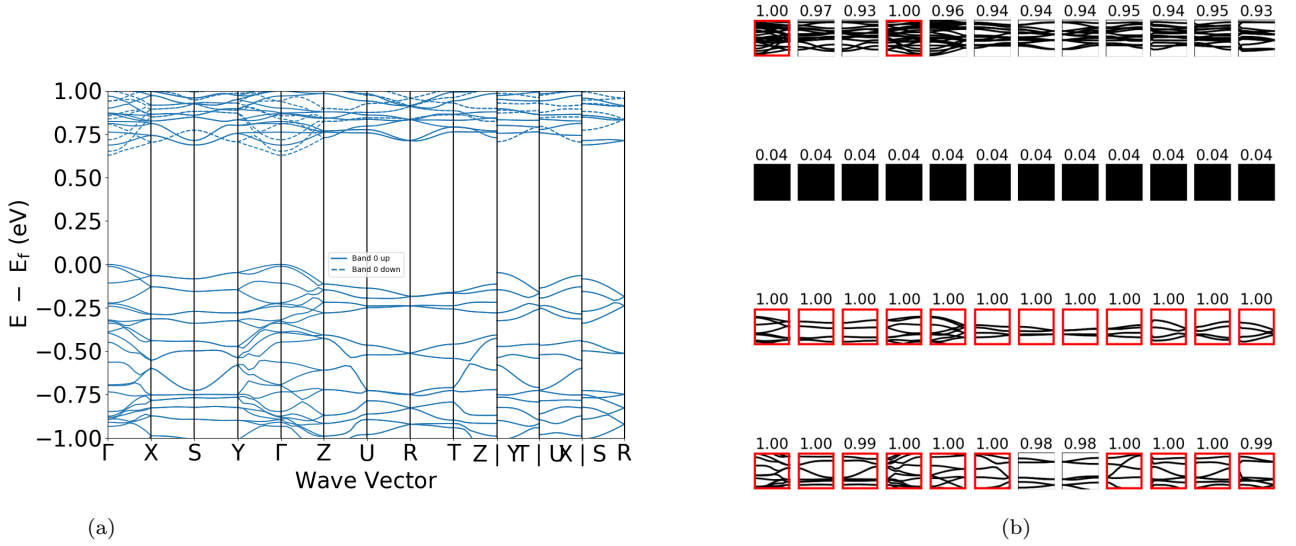


FIG. 5: An example of a plane-flat material. (a) Band structure of $\text{Li}_5\text{TiV}_3\text{O}_8$ (mp-759911). (b) Segmentation of the band structure image of $\text{Li}_5\text{TiV}_3\text{O}_8$ horizontally into four 0.5 eV energy strips, and vertically along high symmetry points in k-space. Black squares correspond to the fully white regions of the band structure plot. The segments with their flatness scores satisfying $f_{ij} \geq 0.99$ are highlighted in red.

IV. RESULTS

In this work, we leverage the Materials Project database as our primary source for band structure images. Subsequently, these images undergo a segmentation process before being fed into the trained convolutional neural network. The segmentation plays a crucial role in isolating and identifying flat segments within the band structures. Through horizontal segmentation, we can discern whether a flat-band material is uniformly flat in all planes or if it exhibits flatness only in one specific direction, between two high-symmetry points.

Our focus is on examining band segments near the Fermi level. We initially selected the energy range $[-1, 1]$ eV with respect to the Fermi level, which was then divided into four energy bandwidths of 0.5 eV each. The distribution of compound flatness scores is illustrated in Fig. 6.

We observe a prominent peak near a compound flatness score of 1, suggesting the adoption of a large flatness criterion. Starting with $\phi = 0.9$, we assessed the number of materials for which their compound flatness scores satisfy $\mathcal{F} \geq 0.9$. For this specific value of ϕ , we identified 30,298 materials meeting the given criterion. We then modified our approach and calculated the number of materials with at least one fully flat row. Using this method, we identified 25,336 plane-flat materials. However, even with this flatness criterion, we concluded that to pinpoint the best plane-flat materials, featuring flat bands throughout their entire Brillouin zones, we needed to further increase the flatness criterion to $\phi = 0.99$. This ensured that the identified flat bands were extremely flat. The distribution of materials, along with their corre-

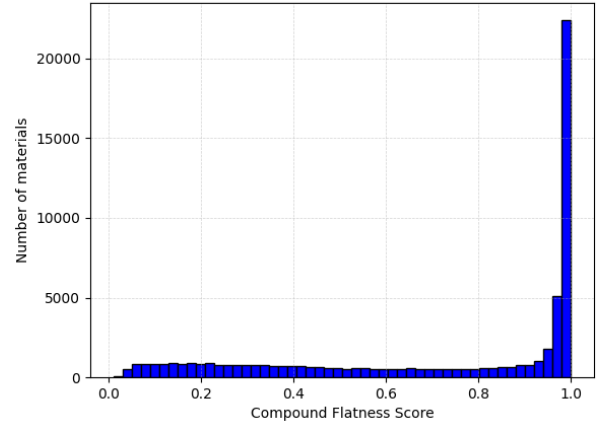


FIG. 6: Histogram displaying compound flatness scores, as given by Eq. (37), for all 60,548 considered materials.

sponding flatness scores in the range $[0.9, 1]$, is presented in Fig. 7, where materials with a compound flatness score $\mathcal{F} \geq 0.99$ are also highlighted in the figure.

We determined that the number of plane-flat materials with compound flatness scores $\mathcal{F} \geq 0.99$ is 15,597. From this subset, we applied an additional condition on their band structures to identify materials with high-quality flat bands. Specifically, we considered band structures featuring at least one row composed entirely of flat segments. This filtering yielded 5,651 plane-flat materials with bands exceptionally flat in the vicinity of the Fermi level. It is noteworthy this number is significantly lower compared to the total number of plane-flat materials with

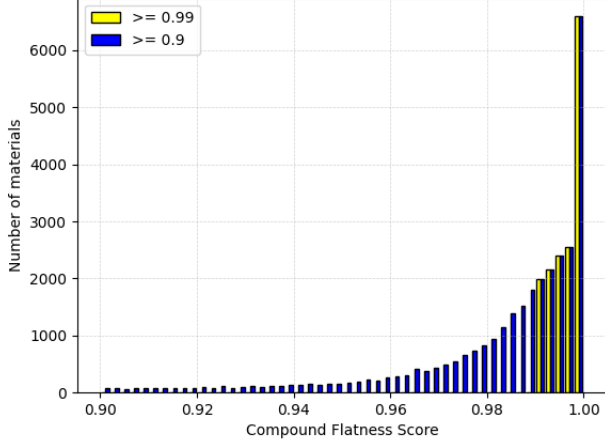
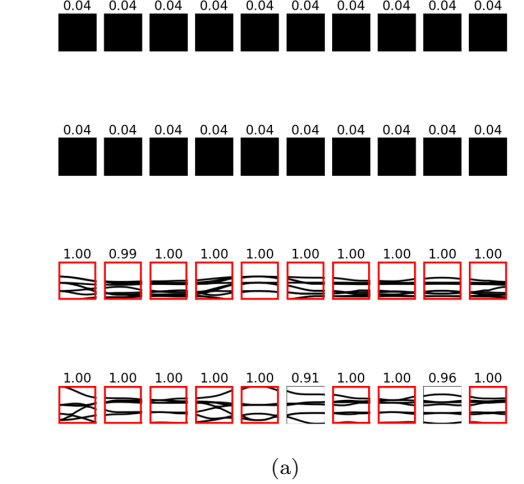


FIG. 7: Histogram showing compound flatness scores for the interval $[0.9, 1]$, which contains 30,298 materials. The materials with compound flatness scores $\mathcal{F} \geq 0.9$ are indicated in blue, and the materials with compound flatness scores $\mathcal{F} \geq 0.99$ are also indicated in yellow.

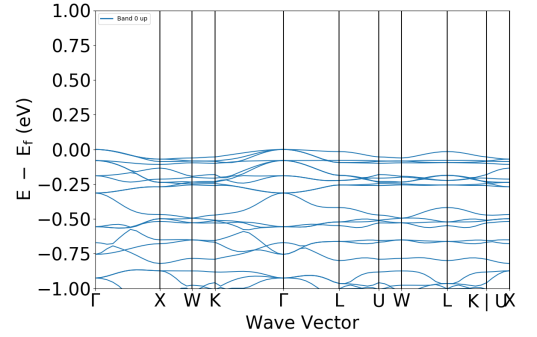
compound flatness scores in the interval $[0.9, 1]$.

An analysis of the difference between the two approaches was also performed and it was found that given $\phi = 0.9$ the fractions of materials with at least one fully flat row and materials with their compound scores above ϕ were most similar for this value of the flatness criterion. It was found that the number of materials with their $\mathcal{F} \geq 0.9$ was 30,298, giving a fraction of around 50% of the total number of materials, while the number of materials with at least one fully flat row was 25,336, which gives a fraction of around 41.8%. This suggests that taking this range gives the best agreement between the two definitions, however, this does not mean that taking $\phi = 0.9$ yields the best flat-band materials. We can also see that the number of materials with their $\mathcal{F} \geq 0.9$ roughly corresponds to half of the total number of materials. An example of a material classified as plane-flat using the two described definitions is shown in Fig. 8.

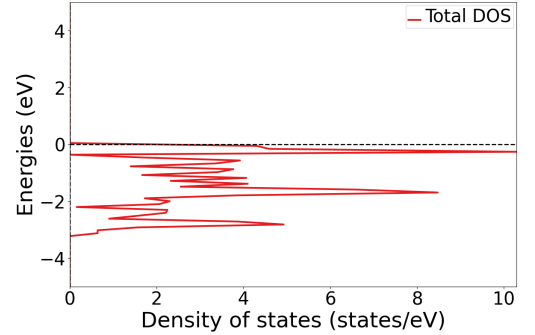
In this study, we are most interested in examining bands that are flat in the vicinity of the Fermi level. For this reason, we also explored the number of materials with at least one fully flat row in the range $[-0.5, 0.5]$ eV with respect to the Fermi level. Plane-flat materials in this range would provide us with materials of even higher quality. Considering this range as $[-0.5, 0.5]$ eV, we found that the number of materials with at least one fully flat row in this range is 3,871 out of the total 60,548 materials with band structure data available from the Materials Project database. However, by narrowing down this range, we might overlook many other potential plane-flat materials. Therefore, considering the energy range as $[-1, 1]$ eV provides a better correspondence between the quality of flat-band materials and their quantity.



(a)



(b)



(c)

FIG. 8: An example of a material that is considered plane-flat under both methods of defining a plane-flat material. (a) Segmentation of band structure image of KY_3F_{10} (mp-2943) horizontally into four 0.5 eV energy strips, and vertically along high symmetry points in k-space. Black squares correspond to the fully white regions of the band structure plot. The segments with their flatness scores satisfying $f_{ij} \geq 0.99$ are highlighted in red. (b) Band structure of KY_3F_{10} . (c) The total density of states of KY_3F_{10} .

V. DISCUSSION

Our supervised machine learning framework, leveraging a trained convolutional neural network, excels in identifying high-quality plane-flat three-dimensional materials. The success of our approach is evident in the identification of flat-band materials, encompassing both theoretical and experimentally observed materials. Remarkably, a lot of our plane-flat materials align with those identified in other studies [30]. By cross-referencing with the cited research work, which covers around 1800 entries, we identified up to 800 materials sharing the classification of flat-band materials. The inter-database correspondence is made feasible by utilizing ICSDs (Inorganic Crystal Structure Database entries) associated with Materials Project entries. As additional flat-band databases emerge, the potential for cross-checking and validation increases.

Enhancing the quality of identified flat-band materials may involve introducing additional criteria concerning bands near the Fermi level. One such condition could be the presence of a density of states peak in the Fermi level vicinity, with a width smaller than the considered energy bandwidth. Another criterion might be a reduction of the considered bandwidth range. Additionally, the demand for well-defined crystallographic structures and ease of synthesis could be imposed on flat-band materials. Such criteria would yield a subset of plane-flat materials characterized by even higher quality, making them amenable to experimental verification.

Future work will focus on identifying flat band sublattices by leveraging our supervised machine learning framework for sublattice extraction. This approach provides a pathway for validating a conjecture that proposes

the use of flat band element sublattices, rather than the entire crystal structure, for material clustering purposes [22]. Numerous studies support this conjecture, illustrating that elemental sublattices adhering to specific lattice and orbital symmetries are instrumental in the emergence of flat bands [52, 53]. Sublattice identification can be accomplished through element-projected density of states; flat bands are often associated with peaks in the DOS, as demonstrated in Fig. 8c. Furthermore, our future investigations will delve into the application of unsupervised machine learning for density-based clustering. This method seeks to categorize materials with comparable structural fingerprints, capitalizing on insights obtained from the findings presented in this work.

VI. CODE AVAILABILITY

The code generated for this study is available in the form of Python and Jupyter Notebook files on the GitHub repository at <https://github.com/mikolgawek/MPhys>.

VII. ACKNOWLEDGEMENTS

M.J.G. would like to thank Mateusz Czajka for his suggestions and help on matters related to CNNs. M.J.G. would like to thank Prof. Artem Mishchenko and Dr Anupam Bhattacharya for their supervision and help with this study. M.J.G. would also like to thank Prof. Artem Mishchenko for his assistance with the database downloads. M.J.G. expresses gratitude to Dr Patrick Huck from Berkeley Lab for his suggestions on API usage as well as to Prof. Nicolas Regnault from Princeton for his valuable inputs on the Materials Project ICSDs.

-
- [1] M. Tovmasyan, S. Peotta, P. Törmä, and S. D. Huber, Effective theory and emergent SU(2) symmetry in the flat bands of attractive hubbard models, *Phys. Rev. B* **94**, 245149 (2016).
 - [2] R. Bistritzer and A. H. MacDonald, Moiré bands in twisted double-layer graphene, *Proceedings of the National Academy of Sciences* **108**, 12233 (2011).
 - [3] D.-S. Ma, Y. Xu, C. S. Chiu, N. Regnault, A. A. Houck, Z. Song, and B. A. Bernevig, Spin-orbit-induced topological flat bands in line and split graphs of bipartite lattices, *Phys. Rev. Lett.* **125**, 266403 (2020).
 - [4] K. Sun, Z. Gu, H. Katsura, and S. Das Sarma, Nearly flatbands with nontrivial topology, *Phys. Rev. Lett.* **106**, 236803 (2011).
 - [5] E. Tang, J.-W. Mei, and X.-G. Wen, enHigh-temperature fractional quantum hall states, *Phys Rev Lett* **106**, 236802 (2011).
 - [6] T. Neupert, L. Santos, C. Chamon, and C. Mudry, Fractional quantum hall states at zero magnetic field, *Phys. Rev. Lett.* **106**, 236804 (2011).
 - [7] D. N. Sheng, Z.-C. Gu, K. Sun, and L. Sheng, Fractional quantum hall effect in the absence of landau levels, *Nature Communications* **2**, 389 (2011).
 - [8] N. Regnault and B. A. Bernevig, Fractional chern insulator, *Phys. Rev. X* **1**, 021014 (2011).
 - [9] H. Li, S. Li, E. C. Regan, D. Wang, W. Zhao, S. Kahn, K. Yumigeta, M. Blei, T. Taniguchi, K. Watanabe, S. Tongay, A. Zettl, M. F. Crommie, and F. Wang, Imaging two-dimensional generalized wigner crystals, *Nature* **597**, 650 (2021).
 - [10] P. Kumar, S. Peotta, Y. Takasu, Y. Takahashi, and P. Törmä, Flat-band-induced non-fermi-liquid behavior of multicomponent fermions, *Phys. Rev. A* **103**, L031301 (2021).
 - [11] D. C. Tsui, H. L. Stormer, and A. C. Gossard, Two-dimensional magnetotransport in the extreme quantum limit, *Phys. Rev. Lett.* **48**, 1559 (1982).
 - [12] A. Mielke and H. Tasaki, Ferromagnetism in the hubbard model, *Communications in Mathematical Physics* **158**, 341 (1993).
 - [13] A. Mielke, Ferromagnetic ground states for the hubbard model on line graphs, *Journal of Physics A: Mathematical and General* **24**, L73 (1991).

- [14] H. Aoki, Theoretical possibilities for flat band superconductivity, *Journal of Superconductivity and Novel Magnetism* **33**, 2341 (2020).
- [15] H. Sakai, K. Yoshimura, H. Ohno, H. Kato, S. Kambe, R. E. Walstedt, T. D. Matsuda, Y. Haga, and Y. Onuki, Superconductivity in a pyrochlore oxide, $\text{Cd}_2\text{Re}_2\text{O}_7$, *Journal of Physics: Condensed Matter* **13**, L785 (2001).
- [16] J. P. Wakefield, M. Kang, P. M. Neves, D. Oh, S. Fang, R. McTigue, S. Y. Frank Zhao, T. N. Lamichhane, A. Chen, S. Lee, S. Park, J.-H. Park, C. Jozwiak, A. Bostwick, E. Rotenberg, A. Rajapitamahuni, E. Vescovo, J. L. McChesney, D. Graf, J. C. Palmstrom, T. Suzuki, M. Li, R. Comin, and J. G. Checkelsky, Three-dimensional flat bands in pyrochlore metal $\text{Cd}_2\text{Re}_2\text{O}_7$, *Nature* **623**, 301 (2023).
- [17] L. Xian, A. Fischer, M. Claassen, J. Zhang, A. Rubio, and D. M. Kennes, Engineering three-dimensional moiré flat bands, *Nano Letters* **21**, 7519 (2021).
- [18] A. Mielke, Ferromagnetic ground states for the hubbard model on line graphs, *Journal of Physics A: Mathematical and General* **24**, L73 (1991).
- [19] A. Mielke, Ferromagnetism in the hubbard model on line graphs and further considerations, *Journal of Physics A: Mathematical and General* **24**, 3311 (1991).
- [20] H. Tasaki, From Nagaoka's Ferromagnetism to Flat-Band Ferromagnetism and Beyond: An Introduction to Ferromagnetism in the Hubbard Model, *Progress of Theoretical Physics* **99**, 489 (1998).
- [21] and and, Exotic electronic states in the world of flat bands: From theory to material, *Chinese Physics B* **23**, 077308 (2014).
- [22] A. Bhattacharya, I. Timokhin, R. Chatterjee, Q. Yang, and A. Mishchenko, Deep learning approach to genome of two-dimensional materials with flat electronic bands, *npj Computational Materials* **9**, 101 (2023).
- [23] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, Commentary: The Materials Project: A materials genome approach to accelerating materials innovation, *APL Materials* **1**, 011002 (2013).
- [24] C. Toher, C. Oses, D. Hicks, E. Gossett, F. Rose, P. Nath, D. Usanmaz, D. C. Ford, E. Perim, C. E. Calderon, J. J. Plata, Y. Lederer, M. Jahnátek, W. Setyawan, S. Wang, J. Xue, K. Rasch, R. V. Chepulskii, R. H. Taylor, G. Gomez, H. Shi, A. R. Supka, R. A. R. Al Orabi, P. Gopal, F. T. Cerasoli, L. Liyanage, H. Wang, I. Siloi, L. A. Agapito, C. Nyshadham, G. L. W. Hart, J. Carrete, F. Legrain, N. Mingo, E. Zurek, O. Isayev, A. Tropsha, S. Sanvito, R. M. Hanson, I. Takeuchi, M. J. Mehl, A. N. Kolmogorov, K. Yang, P. D'Amico, A. Calzolari, M. Costa, R. D. Gennaro, M. B. Nardelli, M. Fornari, O. Levy, and S. Curtarolo, The aflow fleet for materials discovery, in *Handbook of Materials Modeling : Methods: Theory and Modeling*, edited by W. Andreoni and S. Yip (Springer International Publishing, Cham, 2018) pp. 1–28.
- [25] M. G. Vergniory, B. J. Wieder, L. Elcoro, S. S. P. Parkin, C. Felser, B. A. Bernevig, and N. Regnault, All topological bands of all nonmagnetic stoichiometric materials, *Science* **376**, eabg9094 (2022).
- [26] B. Bradlyn, L. Elcoro, J. Cano, M. G. Vergniory, Z. Wang, C. Felser, M. I. Aroyo, and B. A. Bernevig, Topological quantum chemistry, *Nature* **547**, 298 (2017).
- [27] M. G. Vergniory, L. Elcoro, C. Felser, N. Regnault, B. A. Bernevig, and Z. Wang, A complete catalogue of high-quality topological materials, *Nature* **566**, 480 (2019).
- [28] J. Duan, D.-S. Ma, R.-W. Zhang, Z. Zhang, C. Cui, W. Jiang, Z.-M. Yu, and Y. Yao, Inventory of high-quality flat-band van der waals materials (2022), arXiv:2204.00810 [cond-mat.mtrl-sci].
- [29] H. Liu, S. Meng, and F. Liu, Screening two-dimensional materials with topological flat bands, *Phys. Rev. Mater.* **5**, 084203 (2021).
- [30] N. Regnault, Y. Xu, M.-R. Li, D.-S. Ma, M. Jovanovic, A. Yazdani, S. S. P. Parkin, C. Felser, L. M. Schoop, N. P. Ong, R. J. Cava, L. Elcoro, Z.-D. Song, and B. A. Bernevig, Catalogue of flat-band stoichiometric materials, *Nature* **603**, 824 (2022).
- [31] D. Călugăru, A. Chew, L. Elcoro, Y. Xu, N. Regnault, Z.-D. Song, and B. A. Bernevig, General construction and topological classification of crystalline flat bands, *Nature Physics* **18**, 185 (2022).
- [32] J. Teuwen and N. Moriakov, Chapter 20 - convolutional neural networks, in *Handbook of Medical Image Computing and Computer Assisted Intervention*, The Elsevier and MICCAI Society Book Series, edited by S. K. Zhou, D. Rueckert, and G. Fichtinger (Academic Press, 2020) pp. 481–501.
- [33] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, Convolutional neural networks: an overview and application in radiology, *Insights into Imaging* **9**, 611 (2018).
- [34] R. G. Parr and Y. Weitao, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, 1995).
- [35] C. Kittel, *Introduction to Solid State Physics* (Wiley, 2004).
- [36] O. Madelung, *Introduction to solid-state theory* (Springer-Verlag Berlin, Berlin, 1978).
- [37] H. Ibach and H. Lüth, The electronic bandstructure of solids, in *Solid-State Physics: An Introduction to Principles of Materials Science* (Springer Berlin Heidelberg, Berlin, Heidelberg, 1996) pp. 129–154.
- [38] P. Hohenberg and W. Kohn, Inhomogeneous electron gas, *Phys. Rev.* **136**, B864 (1964).
- [39] W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, *Phys. Rev.* **140**, A1133 (1965).
- [40] P. Kratzer and J. Neugebauer, The basics of electronic structure theory for periodic systems, *Frontiers in Chemistry* **7**, 10.3389/fchem.2019.00106 (2019).
- [41] G. Kresse and J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, *Phys. Rev. B* **54**, 11169 (1996).
- [42] A. Jain, G. Hautier, C. J. Moore, S. Ping Ong, C. C. Fischer, T. Mueller, K. A. Persson, and G. Ceder, A high-throughput infrastructure for density functional theory calculations, *Computational Materials Science* **50**, 2295 (2011).
- [43] A. Jain, G. Hautier, C. J. Moore, S. P. Ong, C. C. Fischer, T. Mueller, K. A. Persson, and G. Ceder, A high-throughput infrastructure for density functional theory calculations, *Computational Materials Science* **50**, 2295 (2011).
- [44] R. Kingsbury, A. S. Gupta, C. J. Bartel, J. M. Munro, S. Dwaraknath, M. Horton, and K. A. Persson, Performance comparison of $r^2\text{SCAN}$ and scan meta-gga density functionals for solid materials via an automated, high-throughput computational workflow, *Phys. Rev. Mater.* **6**, 013801 (2022).

- [45] P. E. Blöchl, Projector augmented-wave method, Phys. Rev. B **50**, 17953 (1994).
- [46] G. Bergerhoff, R. Hundt, R. Sievers, and I. D. Brown, The inorganic crystal structure data base, Journal of Chemical Information and Computer Sciences **23**, 66 (1983).
- [47] R. Hundt, J. C. Schön, and M. Jansen, Cmpz— an algorithm for the efficient comparison of periodic structures, Journal of Applied Crystallography **39**, 6 (2006).
- [48] R. W. Godby, M. Schlüter, and L. J. Sham, Self-energy operators and exchange-correlation potentials in semiconductors, Phys. Rev. B **37**, 10159 (1988).
- [49] J. P. Perdew, Density functional theory and the band gap problem, International Journal of Quantum Chemistry **28**, 497 (1985).
- [50] J. Heyd, J. E. Peralta, G. E. Scuseria, and R. L. Martin, enEnergy band gaps and lattice parameters evaluated with the Heyd-Scuseria-Ernzerhof screened hybrid functional, J Chem Phys **123**, 174101 (2005).
- [51] P. Borlido, J. Schmidt, A. W. Huran, F. Tran, M. A. L. Marques, and S. Botti, Exchange-correlation functionals for band gaps of solids: benchmark, reparametrization and machine learning, npj Computational Materials **6**, 96 (2020).
- [52] S. Zhang, M. Kang, H. Huang, W. Jiang, X. Ni, L. Kang, S. Zhang, H. Xu, Z. Liu, and F. Liu, Kagome bands disguised in a coloring-triangle lattice, Phys. Rev. B **99**, 100404 (2019).
- [53] H. Nakai and C. Hotta, Perfect flat band with chirality and charge ordering out of strong spin-orbit interaction, Nature Communications **13**, 579 (2022).
- [54] J. P. Wakefield, M. Kang, P. M. Neves, D. Oh, S. Fang, R. McTigue, S. Y. Frank Zhao, T. N. Lamichhane, A. Chen, S. Lee, S. Park, J.-H. Park, C. Jozwiak, A. Bostwick, E. Rotenberg, A. Rajapitamahuni, E. Vescovo, J. L. McChesney, D. Graf, J. C. Palmstrom, T. Suzuki, M. Li, R. Comin, and J. G. Checkelsky, Three-dimensional flat bands in pyrochlore metal SnTe , Nature **623**, 301 (2023).

Appendix A: Hohenberg-Kohn theorems

Theorem (Hohenberg-Kohn 1). *The external potential $v(\mathbf{r})$ is determined, within a trivial additive constant, by the electron density $\rho(\mathbf{r})$.*

Proof. Consider the electron density $\rho(\mathbf{r})$ for the nondegenerate ground state of an N -electron system. We know that $\rho(\mathbf{r})$ determines both N and $v(\mathbf{r})$, and consequently, all properties. Now, suppose that there are two different potentials, $v(\mathbf{r})$ and $v'(\mathbf{r})$, that yield the same ρ for the ground state. We would then have two Hamiltonians, \hat{H} and \hat{H}' , whose ground-state densities are the same, but the ground-state wave functions $|\Psi\rangle$ and $|\Psi'\rangle$ would be different. Taking $|\Psi'\rangle$ for the \hat{H} problem and using

$E[\Psi] \geq E_0$, we obtain:

$$\begin{aligned} E_0 &< \langle \Psi' | \hat{H} | \Psi' \rangle \\ &= \langle \Psi' | \hat{H}' | \Psi' \rangle + \langle \Psi' | \hat{H} - \hat{H}' | \Psi' \rangle \\ &= \langle \Psi' | \hat{H}' | \Psi' \rangle + \langle \Psi' | \hat{V} - \hat{V}' | \Psi' \rangle \\ &= E'_0 + \int [v(\mathbf{r}) - v'(\mathbf{r})] d\mathbf{r}. \end{aligned} \quad (\text{A1})$$

Now, we take $|\Psi\rangle$ for \hat{H}' :

$$\begin{aligned} E_0 &< \langle \Psi | \hat{H}' | \Psi \rangle \\ &= \langle \Psi | \hat{H} | \Psi \rangle + \langle \Psi | \hat{H}' - \hat{H} | \Psi \rangle \\ &= \langle \Psi | \hat{H} | \Psi \rangle + \langle \Psi | \hat{V}' - \hat{V} | \Psi \rangle \\ &= E_0 + \int [v'(\mathbf{r}) - v(\mathbf{r})] d\mathbf{r}. \end{aligned} \quad (\text{A2})$$

Finally, we add the two equations to get $E_0 + E'_0 < E'_0 + E_0$, which is a contradiction.

Hence, the external potential is uniquely determined by the electron density. \blacksquare

Theorem (Hohenberg-Kohn 2). *For a trial density $\bar{\rho}(\mathbf{r})$, such that $\bar{\rho}(\mathbf{r}) \geq 0$ and $\int \bar{\rho}(\mathbf{r}) d\mathbf{r} = N$,*

$$E_0 \leq E_v[\bar{\rho}]. \quad (\text{A3})$$

Proof. Note that the first Hohenberg-Kohn theorem assures that $\bar{\rho}$ determines its own \bar{v} , Hamiltonian \hat{H} , and wave function $\bar{\Psi}$, which can be taken as a trial wave function for the problem of interest having external potential v . Thus,

$$\langle \bar{\Psi} | \hat{H} | \bar{\Psi} \rangle = \int \bar{\rho}(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} + F_{\text{HK}}[\bar{\rho}] = E_v[\bar{\rho}] \geq E_v[\rho] \geq E_0. \quad (\text{A4})$$

\blacksquare

Appendix B: Kohn-Sham theorem

Theorem (Kohn-Sham). *Introduction of the noninteracting kinetic energy functional*

$$T_s = \sum_i^N \langle \psi_i | -\frac{1}{2} \nabla_i^2 | \psi_i \rangle \quad (\text{B1})$$

into the energy expression

$$E[\rho] = T_s[\rho] + J[\rho] + E_{xc}[\rho] + \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} \quad (\text{B2})$$

leads, upon functional differentiation, to a collection of one-particle equations whose solutions are the

Kohn–Sham orbitals:

$$\left[-\frac{1}{2}\nabla^2 + v_{\text{eff}}\right] \psi_i = \epsilon_i \psi_i, \quad (\text{B3})$$

$$v_{\text{eff}}(\mathbf{r}) = v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{xc}(\mathbf{r}), \quad (\text{B4})$$

$$\rho(\mathbf{r}) = \sum_i^N \sum_s |\psi_i(\mathbf{r}, s)|^2. \quad (\text{B5})$$

Proof. The energy functional $E[\rho] = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F[\rho]$ can be rewritten as

$$\begin{aligned} E[\rho] &= T_s[\rho] + J[\rho] + E_{xc}[\rho] + \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} \\ &= \sum_i^N \sum_s \int \psi_i^*(\mathbf{q}) \left(-\frac{1}{2}\nabla^2\right) \psi_i(\mathbf{q})d\mathbf{r} + J[\rho] \\ &\quad + E_{xc}[\rho] + \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r}, \end{aligned} \quad (\text{B6})$$

and the electron density as

$$\rho(\mathbf{r}) = \sum_i^N \sum_s |\psi_i(\mathbf{r}, s)|^2. \quad (\text{B7})$$

Thus, we expressed the energy in terms of N orbitals. We also require the N orbitals to be continuous-to have finite kinetic energy, and to be square integrable-to guarantee normalisation. We can then perform the variational minimisation of $E[\rho]$ in the space of orbitals $\{\psi_i\}$. We also need to constrain the orbitals to be orthonormal, i.e.

$$\int \psi_i^*(\mathbf{q})\psi_j(\mathbf{q})d\mathbf{q} = \delta_{ij}. \quad (\text{B8})$$

We now define the functional of N orbitals

$$\Omega[\psi_i] = E[\rho] - \sum_i^N \sum_j^N \epsilon_{ij} \int \psi_i^*(\mathbf{q})\psi_j(\mathbf{q})d\mathbf{q}, \quad (\text{B9})$$

where $E[\rho]$ is the functional of the ψ_i , and the ϵ_{ij} are the Lagrange multipliers for the constraints expressed in Eq. (B8). For $E[\rho]$ to be a minimum, it is necessary that

$$\delta\Omega[\{\psi_i\}] = 0 \quad (\text{B10})$$

which leads to the equation

$$\hat{h}_{\text{eff}}\psi_i = \left[-\frac{1}{2}\nabla^2 + v_{\text{eff}}\right] \psi_i = \sum_j^N \epsilon_{ij}\psi_j \quad (\text{B11})$$

with the effective potential $v_{\text{eff}}(\mathbf{r})$ determined from the density through

$$\begin{aligned} v_{\text{eff}}(\mathbf{r}) &= v(\mathbf{r}) + \frac{\delta J[\rho]}{\delta \rho(\mathbf{r})} + \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})} \\ &= v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{xc}(\mathbf{r}) \end{aligned} \quad (\text{B12})$$

with the exchange-correlation potential

$$v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})}. \quad (\text{B13})$$

Since \hat{h}_{eff} is a Hermitian operator, ϵ_{ij} is a Hermitian matrix and can be diagonalized by a unitary transformation of the orbitals which leaves the Hamiltonian of Eq. (B11) invariant. The Kohn-Sham orbital equations are thus obtained in their canonical form

$$\left[-\frac{1}{2}\nabla^2 + v_{\text{eff}}\right] \psi_i = \epsilon_i \psi_i, \quad (\text{B14})$$

$$v_{\text{eff}}(\mathbf{r}) = v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{xc}(\mathbf{r}), \quad (\text{B15})$$

$$\rho(\mathbf{r}) = \sum_i^N \sum_s |\psi_i(\mathbf{r}, s)|^2. \quad (\text{B16})$$

These equations are nonlinear and must be solved iteratively. The total energy can be determined from the resultant density via Eq. (B6) or from the formula

$$E = \sum_i^N \epsilon_i - \frac{1}{2} \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + E_{xc}[\rho] - \int v_{xc}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}. \quad (\text{B17})$$

Here

$$\begin{aligned} \sum_i^N \epsilon_i &= \sum_i^N \langle \psi_i | -\frac{1}{2}\nabla^2 + v_{\text{eff}}(\mathbf{r}) | \psi_i \rangle \\ &= T_s[\rho] + \int v_{\text{eff}}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} \end{aligned} \quad (\text{B18})$$

which means that, just as in the Hartree-Fock theory, the total electronic energy is not the sum of the orbital energies. ■