In our previous work titled " Topological Quantization of Superfluid Stiffness in Dirac Materials", we demonstrated the mechanism of the quantization of the superfluid stiffness in Dirac materials. Our findings revealed that the superfluid stiffness in Dirac materials remains quantized as long as the number of Berry monopoles is conserved in the synthetic space. Specifically, in the case of charge-neutral and Zeeman field neutral conditions, the superfluid stiffness(D_{total}) is equal to $\frac{\Delta}{\pi}$ per cone (we will derive a more general D_{total} later). Therefore the quantity of superfluid stiffness is determined by the number of Dirac cones present. This raises a intriguing question: can we establish a mapping from the value of superfluid stiffness to the number of Dirac cones? Such a mapping would have practical implication allowing experimentalists to determine the number of Dirac cones by measuring the superfluid stiffness of real materials, like graphene.

In this note, we aim to address above question through a supervised machine learning approach. We train our neural network by feeding superfluid stiffness data obtained from the Dirac model that may encompass higher order Dirac cones, enabling it to predict the Hamiltonian's total absolute vorticity. The total absolute vorticity is equal to the number of Dirac cones times their corresponding order. After fully training our network, we test our model's performance using superfluid stiffness data for graphene which is a 2D material. Our model achieve a high accuracy, which validates the effectiveness of our method.

1 Dirac model and higher order band touching points

In this section, we will introduce the Dirac model and derive the corresponding superfluid stiffness. It's important to note that we utilize Dirac model's superfluid stiffness as input for our neural network during training process.

The Hamiltonian of the Dirac model in momentum space can be written as

$$\hat{h} = (ak_x + ibk_y)^n \hat{\sigma}_+ + (ak_x - ibk_y)^n \hat{\sigma}_- + \mu \tag{1}$$

$$= |\widetilde{k}|^n \cos(n\theta)\hat{\sigma}_x - |\widetilde{k}|^n \sin(n\theta)\hat{\sigma}_y + \mu \tag{2}$$

$$= |\widetilde{k}|^n e^{in\theta\hat{\sigma}_z}\hat{\sigma}_x + \mu, \tag{3}$$

where μ is chemical potential, n is a positive integer, $\hat{\sigma}$ denote Pauli matrices in sub-lattice space, $|\widetilde{k}| = \sqrt{(ak_x)^2 + (bk_y)^2}$ and $\tan(n\theta) = \frac{(ak_x - ibk_y)^n}{(ak_x + ibk_y)^n}$. The eigen-vector of the Hamiltonian is read as

$$|u_\pm\rangle = \frac{1}{\sqrt{2}} \left(\begin{array}{c} 1 \\ \pm e^{-in\theta} \end{array} \right), \tag{4}$$

whose corresponding eigen-value is presented as $\epsilon_\pm = \pm|\widetilde{k}|^n + \mu$.

Now let's consider that superconductivity is introduced into the system. The super current flow in the direction of gradient of superconducting phase ϕ proportional to superfluid stiffness, such as $J_i = D_{ij}\nabla_i\phi$, where superfluid stiffness D is a symmetric matrix

$$D = \left(\begin{array}{cc} D_{xx} & D_{xy} \\ D_{yx} & D_{yy} \end{array} \right) \tag{5}$$

The element of superfluid stiffness can be read as:

$$D_{ij} = 2\Delta^2 \iint \frac{dk_x}{2\pi}\frac{dk_y}{2\pi} \sum_{\alpha,\beta=\pm} \frac{\langle u_\alpha|\partial_i\hat{h}|u_\beta\rangle\langle u_\beta|\partial_j\hat{h}|u_\alpha\rangle}{E_\alpha^2 - E_\beta^2} \left(\frac{f(E_\beta)}{E_\beta} - \frac{f(E_\alpha)}{E_\alpha} \right), \tag{6}$$

where $i,j \in \{x,y\}$, $f(E) = \Theta(B + E) - \Theta(B - E) = \Theta(E - |B|)$ and $E_\pm = \sqrt{\epsilon_\pm^2 + \Delta^2}$. Utilizing Eq. (6), we immediately obtain the superfluid stiffness of the Dirac model,

$$D_{xx} = \frac{a}{b}(D_{\text{intra}} + D_{\text{inter}}), \tag{7}$$

$$D_{yy} = \frac{b}{a}(D_{\text{intra}} + D_{\text{inter}}), \tag{8}$$

$$D_{xy} = D_{yx} = 0, \tag{9}$$

where

$$D_{\text{intra}} = n \frac{\sqrt{\mu^2 + \Delta^2}}{2\pi} \left\{ \Theta(|\Delta| - |B|) + \Theta(|B| - |\Delta|)\Theta(\sqrt{\mu^2 + \Delta^2} - |B|) \left[1 - \frac{|\mu|}{\sqrt{\mu^2 + \Delta^2}} \frac{|B|}{\sqrt{B^2 - \Delta^2}} \right] \right\},$$

$$D_{\text{inter}} = \frac{n}{|\mu|} \frac{\Delta^2}{2\pi} \left\{ \Theta(|\Delta| - |B|) \ln(\frac{\sqrt{\mu^2 + \Delta^2} + |\mu|}{|\Delta|}) + \Theta(|B| - |\Delta|)\Theta(\sqrt{\mu^2 + \Delta^2} - |B|) \ln(\frac{\sqrt{\mu^2 + \Delta^2} + |\mu|}{\sqrt{B^2 - \Delta^2} + |B|}) \right\}.$$

when $\mu \to 0$, $D_{\text{inter}} = n\frac{|\Delta|}{2\pi}\Theta(|\Delta| - |B|)$. For zero Zeeman field and chemical potential, $D_{xx} = D_{yy} = n\frac{|\Delta|}{\pi}$ which aligns in our initial statement.

In order to eliminate the effect of anisotropy, we are interested at the determinant of superfluid stiffness

$$|D| = D_{xx}D_{yy} - DxyDyx \tag{10}$$
$$= D_{\text{total}}^2, \tag{11}$$

which does not include a, b. Here $D_{\text{total}} = D_{\text{intra}} + D_{\text{inter}}$. By considering the determinant, we can effectively capture the overall behavior and characteristics of the superfluid stiffness, while cancels the influence of the inherent anisotropy present in the system. It is worth to notice that $|D| = D_{\text{total}}^2$ appears because of $D_{xy} = D_{yz} = 0$ in Dirac model, for more general case we need to calculate Eq. (10). For convenience, in the following discussion we set $\Delta = 1$. Figure. (1) illuminates the amplitude of $|D|$ with mu and B varying in the range of $(-3, 3)$. Notably, one may observe that $|D|$ changes discontinuously at $|B| = \Delta$, specifically when $mu = 0$ $|D|$ experiences an abrupt drop $\frac{1}{\pi^2}$ (see inset figure). This discontinuity may serves as an crucial signal in identifying the so called total absolute vorticity.
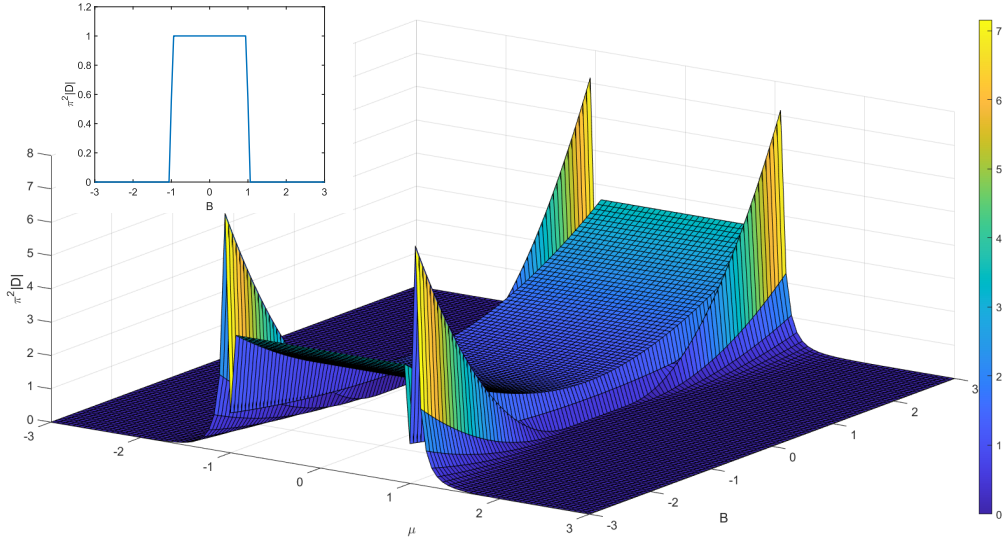

Figure 1: Amplitude of $|D|$ with mu and B varying in the range of $(-3, 3)$. Inset figure shows variation of $|D|$ with B varying and $mu = 0$.

1.1 Total absolute vorticity

The total absolute vorticity is equal to the number of Dirac cones times their corresponding order, which is our neural network's main prediction. One can rewrite above Dirac model Hamiltonian as a vector form

$$\hat{h} = \mathbf{H} \cdot \boldsymbol{\sigma}, \tag{12}$$

where $\mathbf{H} = \left(|\widetilde{k}|^n \cos(n\theta), |\widetilde{k}|^n \sin(n\theta) \right)$. The angle of vector $\mathbf{H}$ is denoted by $\phi = \arctan(-\tan(n\theta)) = -n\theta$. The total absolute vorticity is written as

$$w = \frac{1}{2\pi} \left| \oint d\phi \right| = n. \tag{13}$$

2

2 Parabolic model

In previous section, we compute the contribution of a single Dirac cone to superfluid stiffness. In this section, we study the contribution of a trivial crossing point of the Fermi level to the superfluid stiffness. Without loss generality, we study the parabolic model whose Hamiltonian can be written as

$$\hat{h}_{\mathrm{p}} = \left[(ak_x)^2 + (bk_y)^2 \right] \hat{\sigma}_x + \mu \tag{14}$$

whose eigen-vectors are

$$u_{\pm}^{\mathrm{p}} = \left(\begin{array}{c} 1 \\ \pm 1 \end{array} \right) \tag{15}$$

corresponds to eigen-value $\epsilon_{\pm}^{\mathrm{p}} = \pm |\widetilde{k}|^2 + \mu$. Applying Eq. (6), the contribution to the superfluid stiffness is

$$D_{xx}^{\mathrm{p}} = \frac{a}{b} D_{\mathrm{intra}}, \tag{16}$$

$$D_{yy}^{\mathrm{p}} = \frac{b}{a} D_{\mathrm{intra}}, \tag{17}$$

$$D_{xy}^{\mathrm{p}} = D_{yx}^{\mathrm{p}} = 0. \tag{18}$$

For trivial case, the interband part of superfluid stiffness is equal to zero.

Since the parabolic model's Hamiltonian only have contribution to $\hat{\sigma}_x$, then the corresponding vorticity is equal to zero.

3 Training data set

Before constructing the input data set, it's important to emphasize the primary objective of our neural network. Our neural network aims to predict two key aspects: firstly, whether the crossing point of the Fermi level corresponds to a Dirac cone. Secondly, the number of Dirac cones within a small window as the parameters B and μ vary.

To achieve this goal, we begin by selecting a interested central chemical potential, denoted as μ_0 which serves as the reference point. And then compute the determinant of the superfluid stiffness within a defined window varying μ in the range of $(\mu_0 - \mu_{\mathrm{off}}, \mu_0 + \mu_{\mathrm{off}})$ and B in $(-B_{\mathrm{off}}, B_{\mathrm{off}})$. Finally, a matrix labeled by the total absolute vorticity is obtained. By varying the central potential and total absolute vorticity, we can generate a set of labeled matrices that serve as the input data for our neural network. Each labeled matrix corresponds to a different combination of total absolute vorticity and central chemical potential. Fig. (2) shows a example of training data.
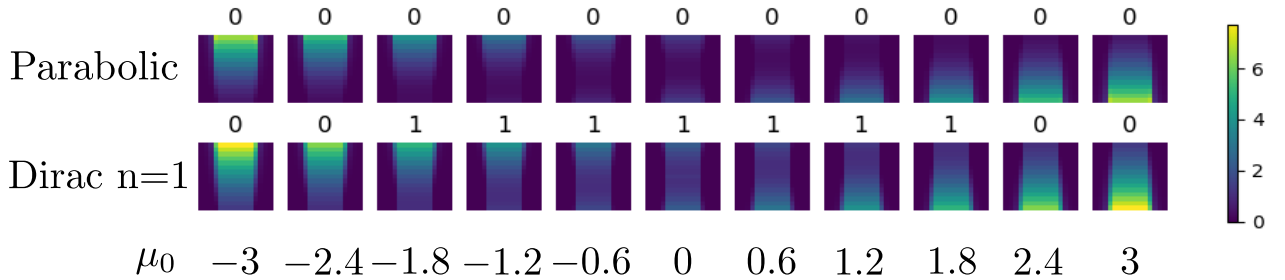

Figure 2: Training data set example. The first row illustrates the determinant of superfluid stiffness for a parabolic model with μ_0 varying in the range of $(-3, 3)$, $\mu_{\mathrm{off}} = B_{\mathrm{off}} = 2$. The second row represents the determinant for a Dirac model($n = 1$) where μ_0 varies in the same range. The corresponding vorticity is shown above the matrix.

3

4 Graphene model

The Hamiltonian of Graphene in momentum space can be written as

$$\hat{h}_{\mathrm{g}} = - t \left[\cos(\frac{1}{2}k_x + \frac{\sqrt{3}}{2}k_y) + \cos(\frac{1}{2}k_x - \frac{\sqrt{3}}{2}k_y) + \cos(-k_x) \right] \hat{\sigma}_x$$

$$- t \left[\sin(\frac{1}{2}k_x + \frac{\sqrt{3}}{2}k_y) + \sin(\frac{1}{2}k_x - \frac{\sqrt{3}}{2}k_y) + \sin(-k_x) \right] \hat{\sigma}_y + \mu, \tag{19}$$

where $t \approx 2.8 eV$ is hopping amplitude.

The eigen-energy have formula $\epsilon_{\pm}^{\mathrm{g}} = \pm t \sqrt{3 + 2\cos(\sqrt{3}k_y) + 4\cos(\frac{\sqrt{3}}{2}k_y)\cos(\frac{3}{2}k_x)} + \mu$ and its corresponding eigen-vector can be read as

$$u_{\pm}^{\mathrm{g}} = \frac{1}{\sqrt{2}} \left(\begin{array}{c} 1 \\ \pm i e^{i\theta_{\mathrm{g}}} \end{array} \right), \tag{20}$$

where

$$\tan(\theta_{\mathrm{g}}) = \frac{\sin(\frac{1}{2}k_x + \frac{\sqrt{3}}{2}k_y) + \sin(\frac{1}{2}k_x - \frac{\sqrt{3}}{2}k_y) + \sin(-k_x)}{\cos(\frac{1}{2}k_x + \frac{\sqrt{3}}{2}k_y) + \cos(\frac{1}{2}k_x - \frac{\sqrt{3}}{2}k_y) + \cos(-k_x)} \tag{21}$$

Applying Eq. (6), one obtain the determinant of graphene's superfluid stiffness.

4.1 Total absolute vorticity of graphene

Since the hopping amplitude is much larger than the induced superconductivity in graphene $\Delta \approx 1 meV$, the Hamiltonian can be approximately described by two Dirac cones

$$\hat{h}_{\mathrm{g}} \approx v_F \left(q_x \hat{\sigma}_x + q_y \hat{\rho}_z \hat{\sigma}_y \right), \tag{22}$$

where $\mathbf{q}$ is the momentum measured relatively to the Dirac points, v_F is the Fermi velocity, given by $v_F = \frac{3}{2}t$, $\hat{\rho}$ are Pauli matrices in valley space.

The total absolute vorticity is the sum of absolute vorticity of two containing Dirac cones, i.e. $w_{\mathrm{g}} \approx 2$. The determinant of graphene would be labeled as 2 if Dirac cones present within the defined window, else would be labeled as 0. By varying hopping amplitude and central chemical potential, one gather a set of test date.

5 Neural network

The neural network work flow and the structure is shown in Fig. (3). One input the determinant of superfluid stiffness to the neural network which output the total absolute vorticity. The neural network encompass two convolutional layers with 32 channels of kernel size 3×3 and 2 channels of kernel size 3×3, followed by two fully connected layers with the first one containing 1024 neurons and the second 512 neurons. At last, the neural network output the total absolute vorticity.
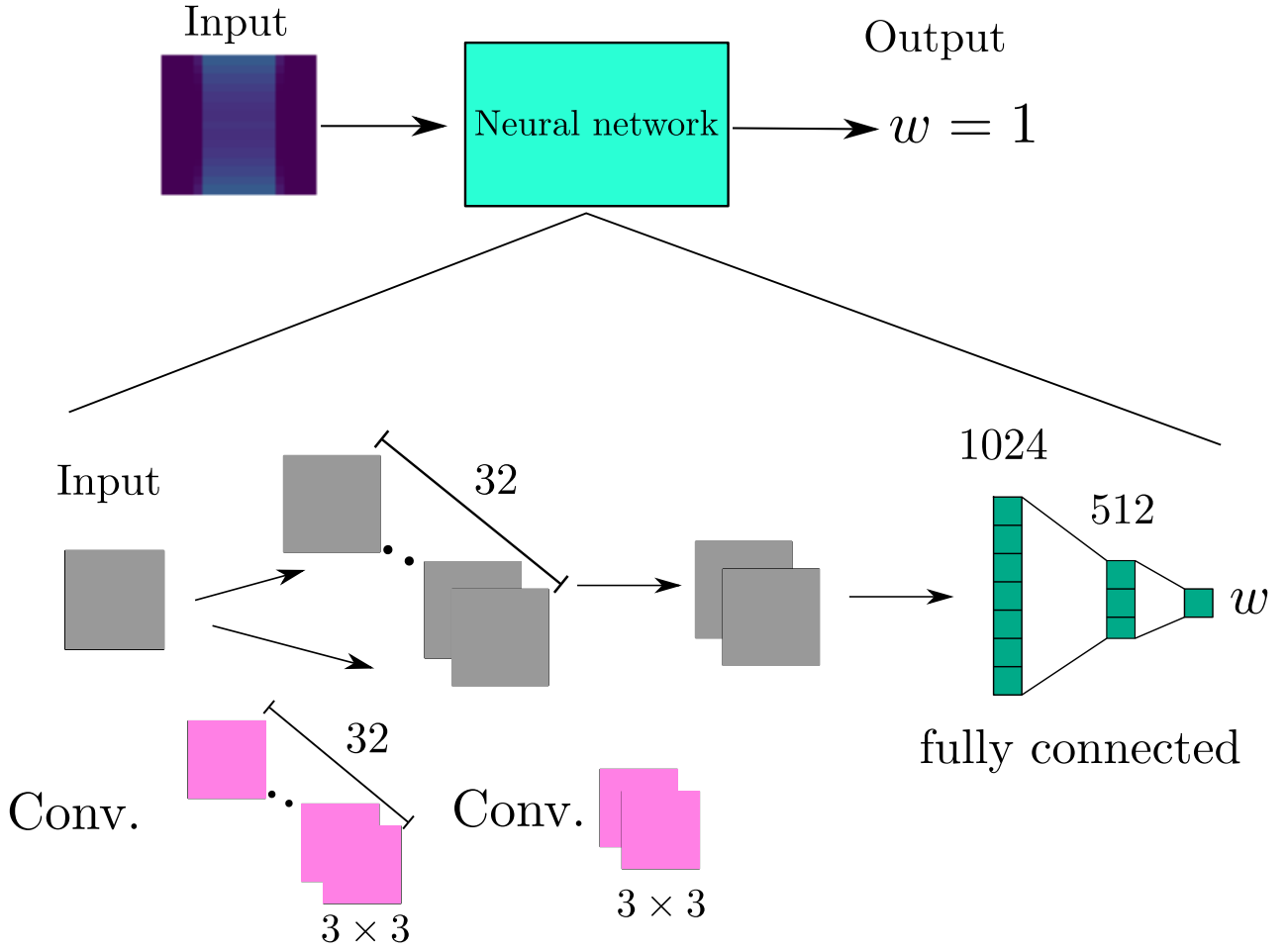
Figure 3: Schematic of the machine learning workflow and the structure of the convolutional neural network.

We split the Dirac and parabolic model data set into 2 part, 4000 determinant samples with the total absolute vorticity $w \in (0, 6)$, randomly choose 70% of them for training data set and the remainder as (i) the first validation data set which was not used during training process; 1000 samples with the total absolute vorticity $w \in (7, 9)$ (ii) as the second validation data set whose vorticity is unseen in training process. (iii) Meanwhile we generate 4000 determinant samples from graphene with $t \in (2000, 4000)$ as test data set. The test result is presented in Table (1).

	Validation(i)	Validation(ii)	Test(iii)
Accuracy	100%	98.3%	98.0%

Table 1: Performance of the neural network with respect to different validation and test data sets.