Pseudo-spin order of Wigner crystals in multi-valley electron gases

Vladimir Calvera,[1,*] Steven A. Kivelson,[1,†] and Erez Berg[2,‡]

[1] *Department of Physics, Stanford University, Stanford, CA 94305, USA*
[2] *Department of Condensed Matter Physics, Weizmann Institute of Science, Rehovot 76100, Israel*

We study multi-valley electron gases in the low density ($r_s \gg 1$) limit. Here the ground-state is always a Wigner crystal (WC), with additional pseudo-spin order where the pseudo-spins are related to valley occupancies. Depending on the symmetries of the host semiconductor and the values of the parameters such as the anisotropy of the effective mass tensors, we find a striped or chiral pseudo-spin antiferromagnet, or a time-reversal symmetry breaking orbital loop-current ordered pseudo-spin ferromagnet. Our theory applies to the recently-discovered WC states in AlAs and in mono and bilayer transition metal dichalcogenides. We identify a set of interesting electronic liquid crystalline phases that could arise by continuous quantum melting of such WCs.

I. INTRODUCTION

The study of the two dimensional electron gas (2DEG), both experimentally and theoretically, has been extraordinarily fruitful [1]. From the theoretical perspective, it is one of the most basic problems in correlated electron physics, while because it can be realized and probed in semiconductor devices, it is of direct experimental significance. In its simplest realization, the electrons are associated with a unique minimum ("valley") in the host semiconductor's band dispersion, and can be treated (in the context of the effective mass approximation) as formally analogous to electrons in free space.

However, in many experimentally relevant cases, the host semiconductor has more than one relevant valley [2]. This results in a more structured version of the 2DEG in which the electrons carry a pseudo-spin index corresponding to the distinct valleys. This opens the possibility of a host of new collective states.

In the present paper, we analyze various natural multi-valley versions of the 2DEG in the low density (large r_s) limit. Here, the Coulomb interactions dominate the kinetic energy, and so for precisely the same reasons as in the simple 2DEG, the system forms an insulating, Wigner crystalline (WC) state [3]. However, when the electrons in question have a pseudo-spin degree of freedom, the nature of the pseudo-spin order in the WC remains to be resolved. Here, we address this issue for various natural valley structures (i.e. for different sorts of pseudo-spin symmetries). We show that the pseudo-spin order can be determined in an asymptotically controlled expansion in $\sqrt{1/r_s}$. (By contrast, the issue of spin order in the WC cannot be addressed in any order in $\sqrt{1/r_s}$, as it involves tunneling processes of order $\exp\left[-\alpha\sqrt{r_s}\right]$.) Analogous phenomena have been discussed in the context of 2DEGs with Rashba spin-orbit coupling, that can be viewed as having a continuous, rather than discrete, pseudo-spin degree of freedom [4; 5].

As addenda, we consider several extensions of these results: 1) We comment on the application of these ideas to moiré systems, in which commensurability effects between the WC and the moiré lattices can be tuned. 2) We briefly consider the simplest 3d version of the pseudo-spin ordering problem. 3) We consider some interesting possible multi-step quantum or thermal melting processes of a pseudo-spin ordered 2D WC that lead to electronic liquid crystalline states with various patterns of vestigial order.

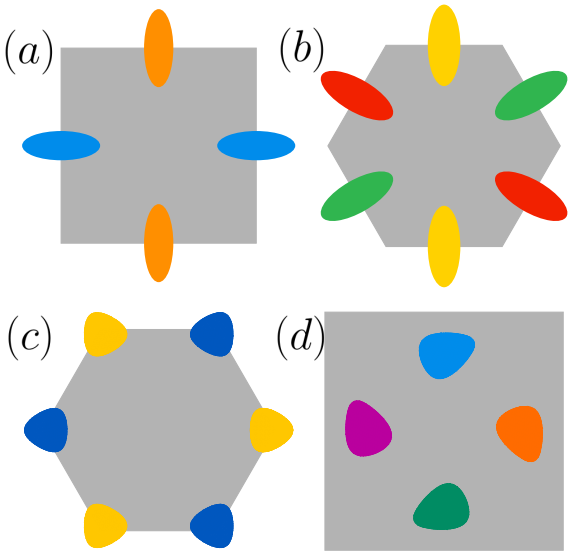


FIG. 1. Schematic representation of valleys on the Brillouin zone of 2d crystals. The distinct colors indicate distinct valleys (pseudo-spin polarizations) and the shapes are representative of the Fermi surface at some fixed, small electron density. Cases (a) and (b) correspond to valleys with anisotropic masses tensors (Sec. III). Case (c) correspond to a pair of valleys related by inversion symmetry and cubic corrections to the dispersion relations (Sec. IV). Case (d) is a more complicated example that could be realized on a system with C_4 rotation but no reflection plane. The same color scheme will be used in subsequent figures.

The paper is organized as follows. In Sec. II we introduce some notation and review the approximations used in this work. Sec. III considers the effect of mass anisotropy on the Wigner crystal and how the pseudo-spin degeneracy is broken when there are multiple valleys related by rotations. Sec. IV treats possible ways to lift the degeneracy between pairs of valleys related by

inversion symmetry. In Sec. V, we consider extensions of our considerations to external moiré potentials and to Wigner crystals in three dimensions. Then in Sec. VI we entertain various scenarios for melting of our Wigner crystal. The bulk of the paper is relatively technical, so in Sec. VII we summarize the principle findings.

II. BACKGROUND

A. Multi-valley semiconductors

Consider a semiconductor whose conduction band has minima at n_v distinct, symmetry-related momenta $\{\boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_{n_v}\}$ in the Brillouin zone. When we slightly dope the system we can use the effective mass approximation. Here we treat the conduction electrons with crystal momenta near $\boldsymbol{Q}_\alpha$ ($\boldsymbol{k} \approx \boldsymbol{Q}_\alpha$) as one flavor of a 2-dimensional electron gas with dispersion relation

$$T_\alpha(\mathbf{p}) = \sum_{a,b} \frac{1}{2} W^{(\alpha)}_{ab} p_a p_b \tag{1}$$

where in terms of $\varepsilon_{\boldsymbol{k}}$, the electron dispersion relation, the effective inverse mass tensor is defined as

$$W^{(\alpha)}_{ab} = \left. \frac{\partial^2 \varepsilon_{\boldsymbol{k}}}{\partial k_a \partial k_b} \right|_{\boldsymbol{k}=\boldsymbol{Q}_\alpha}. \tag{2}$$

We will refer to the electrons with lattice momentum near $\mathbf{Q}_\alpha$ to have *valley flavour* $\mathbf{Q}_\alpha$ or simply α. Here and henceforth we work in units where $\hbar = 1$.

The point-group (PG) symmetry of the semiconductor imposes the following constraints: If $\mathbf{g} \in PG$ acts as $\mathbf{r} \to R_{\mathbf{g}}\mathbf{r}$ in real space, then $\mathbf{g}$ must permute the valleys, $R_{\mathbf{g}} Q_\alpha = Q'_\alpha$ (modulo reciprocal lattice vectors), and the mass tensors are related by

$$W^{(\alpha)} = R_{\mathbf{g}} \cdot W^{(\alpha)} \cdot R^\top_{\mathbf{g}}. \tag{3}$$

Manifestly, the inverse mass tensors $W^{(i)}$ need not have the full PG symmetry of the lattice, but only the subgroup that leaves $\mathbf{Q}_\alpha$ invariant.

In this paper we consider electrons interacting via long-range Coulomb interactions that do not depend on the valley flavor [6] so

$$V(\boldsymbol{R} - \boldsymbol{R}') = \frac{e^2}{4\pi\epsilon} \frac{1}{\|\boldsymbol{R} - \boldsymbol{R}'\|}. \tag{4}$$

The true symmetries of the system are those of the lattice Hamiltonian which has a space group (discrete translations combined with the PG) and internal symmetries such as time reversal symmetry and spin rotation. Once we make the effective mass approximation, there is an emergent continuous translation symmetry ($\mathbb{R}^2$) and particle number conservation for each valley ($\mathrm{U}(1)_{(i)}$). Note that the diagonal subgroup of $\prod_{i=1}^{n_v} \mathrm{U}(1)_i$ should be identified with physical particle number conservation, and is thus an exact symmetry.

In the familiar case of a single valley at a high symmetry point in the Brillouin zone, the emergent symmetry includes the full $O(2)$ rotational symmetry of Euclidean space, but in most cases this does not happen. Indeed, the individual $W^{(i)}$'s do not have the full PG symmetry; typically, PG operations operate both on the spatial indices a and b, as well as involving permutation of the flavor indices, i. (In some cases the emergent symmetry is even larger, e.g. there is a U(2) symmetry when there are two different valleys related by inversion symmetry, but we will not consider this cases here.) For simplicity, we will neglect the possibility of non-symmorphic symmetries, so the effective mass Hamiltonian has the symmetry group

$$G_{EM} = (\mathbb{R}^2 \times \prod_{i=1}^{n_v} \mathrm{U}(1)_{(i)}) \rtimes PG \times G_{\text{int}}$$

where PG acts as $O(2)$ on $\mathbb{R}^2$ and by permutation on the product of U(1)'s. Here G_{int} are the internal symmetries with total particle conservation omitted, which can include spin-rotation and time-reversal symmetries.

B. Low-density limit

The discussion so far holds for electron densities low enough compared to atomic scales, thus justifying the effective mass approximation. There is another density scale that determines the balance between the kinetic energy and the Coulomb interactions, conventionally defined as $1/\pi a_{\text{eff}}^2$ where $a_{\text{eff}} = \frac{4\pi\varepsilon_0\hbar^2}{e^2 m_{\text{eff}}}$ is the effective Bohr radius of the system in which m_{eff} is the square root of the determinant of the effective mass tensor. We define the dimensionless quantity $r_{\text{s}} = (\pi n_{\text{el}} a_{\text{eff}}^2)^{-1/2}$ as usual. We focus on the $r_{\text{s}} \gg 1$ regime, i.e. $n_{\text{el}} \ll [\pi a_{\text{eff}}^2]^{-1}$.

As in the isotropic 2DEG, the kinetic energy per electron scales as $1/r_{\text{s}}^2$ while the Coulomb energy per electron scales as $1/r_{\text{s}}$ but the prefactors depend on the various mass anisotropies. For large values of r_{s} the Coulomb energy dominates forcing the electrons to form a triangular lattice [7]. For later convenience we define ν as the area per electron, which for the triangular lattice is $\nu = \frac{\sqrt{3}}{2} a^2$, in which, by appropriate choice of units, we will set the WC lattice constant $a = 1$.

At this point, there is an extensive degeneracy associated with the pseudo-spin orientation at every lattice site, the resolution of which is the primary goal of the present work. In the absence of spin-orbit coupling, there is also an extensive degeneracy associated with spin orientations, which we will ignore for the bulk of this paper. [8] Finally, there are the usual global degeneracies associated with the broken translation and rotation symmetries in the WC. The former of these is exact within effective mass approximation, but (as we will discuss in the con-

text of moiré systems) is really only a discrete symmetry when commensurability effects with the underlying lattice are included. However, even within the effective mass approximation, the rotational symmetry of the classical ($r_s \to \infty$) problem will be resolved to the discrete PG symmetry, when pseudo-spin ordering is considered; it is thus important to define the orientation of the WC in terms of an angle defined relative to the symmetry axes of the host crystal ($\theta_{\rm WC}$).

The following are the principle features of the band dispersion that could control the pattern of symmetry breaking in the large r_s limit according to the order in $1/r_s$ at which they arise:

1. Valleys with different mass tensors (order - $1/r_{\rm s}^{3/2}$);
2. Trigonal warping corrections to the dispersion relations (order - $1/r_{\rm s}^3$);
3. Berry curvature effects - although they are unimportant for the examples we have considered (order - $1/r_{\rm s}^3$ or $1/r_{\rm s}^{9/2}$).

We will proceed as follows: First, we consider the effect of feature 1 by studying the phonon zero point energy as a function of the mass anisotropy, $\lambda \equiv [m_L - m_T]/[m_L + m_T]$, for $0 < \lambda \leq 1$ (m_L and m_T are the eigenvalues of the mass tensor. When the valley lies on a reflection-symmetric axis of the BZ, we can identify m_L and m_T as the longitudinal and transversal masses, respectively.). For small λ, we obtain an effective statistical mechanics model from which the optimal pattern of pseudo-spin ordering can be derived. For more general λ, we evaluate the energy of this and other possible small period patterns of pseudo-spin order to argue that the same pattern of pseudo-spin order is optimal over the entire range of λ.

Next, we add the corrections to the dispersion relations (features 2 and 3) as small perturbations (in $1/r_{\rm s}$) to the previous states. We again obtain an effective statistical model for the pseudo-spins that determines more subtle features of the ground-state order.

C. Effective mass and harmonic approximations

The effective Hamiltonian has the form

$$H = \sum_j T_{\alpha_j}(\vec{p}_j) + \sum_{i>j} V(\vec{r}_i - \vec{r}_j - \vec{R}_i + \vec{R}_j), \quad (5)$$

where j labels an electron in valley α_j with momentum $\vec{p}_j$ at position $\vec{r}_j + \vec{R}_j$ ($\vec{R}_j$ is the classical WC ground state position).

To quadratic order in these displacements, we consider dimensionless position/momentum variables as $P_{ia} = \frac{p_{ia}}{\sqrt{E_\circ m^\star}}$ and $Q_{ia} = \sqrt{E_\circ m^\star} q_{ia}$ where x_{ia} is the electron displacement in the a direction away from the classical equilibrium position $\vec{R}_i$ and p_{ia} is the conjugate momentum. We use

$$m^\star = \frac{2m_L m_T}{m_L + m_T}; E_\circ = \sqrt{\frac{e^2\hbar^2}{4\pi\epsilon a^3 m^\star}} \propto r_{\rm s}^{-3/2} \quad (6)$$

for mass and energy scales, respectively. The effective dimensionless Hamiltonian is

$$h_{\rm eff} = \frac{H_{\rm eff}}{E_\circ} = \sum_i \frac{1}{2} P_i \overline{W}_i P_i + \sum_{i,j} \frac{1}{2} Q_i K_{ij} Q_j \quad (7)$$

where $\overline{W}_i$ is $W^{(\alpha_i)}/m^\star$ and K_{ij} is the dipole-dipole coupling matrix given by

$$\begin{aligned} (K_{ij})_{ab} &= \frac{1}{R_{ij}^3}\left(\delta_{ab} - 3\frac{(R_{ij})_a (R_{ij})_b}{R_{ij}^2}\right); i \neq j \\ (K_{ii})_{ab} &= \gamma\delta_{ab}; \quad \gamma = \frac{1}{2}\sum_{i\neq 0}\frac{1}{R_{i0}^3}. \end{aligned} \quad (8)$$

Where $\gamma \approx 5.5171$ for the triangular lattice. It is important to recognize that the effective mass tensor in Eq. 7 depends implicitly on the pseudo-spin ordering. Because the effective Hamiltonian is quadratic, it can still be diagonalized in terms of normal modes (phonons) with energies $\hbar\omega_\mu$. Thus, the pseudo-spin-configuration dependent zero point energy of the phonon per electron can be expressed as

$$e_{\rm ph} \equiv \frac{E_{\rm eff}}{N_{\rm el} E_\circ} = \frac{1}{2N_{\rm el}} \sum_\mu \omega_\mu \quad (9)$$

where $E_{\rm eff}$ is the phonon zero-point motion correction to the ground-state energy per electron. (Similar expressions can be obtained for the configuration dependent free energy, at low enough temperatures that the harmonic approximation is reliable.) The preferred pseudo-spin order is the pattern which minimizes $E_{\rm eff}$. (This is, in essence, a form of "order by disorder.")

Note that when the pseudo-spin pattern is periodic in space, we can label the normal modes by a crystal momentum and a band index. The crystal momenta live in a new Brillouin zone (BZ′) that is obtained by folding the original BZ of WC. The band index take $2B$ values, where B is the number of lattice sites in the unit cell of the pattern.

III. ANISOTROPIC MASS TENSORS

In this section, we calculate the effect on Wigner crystallization when the mass tensors are anisotropic and the valleys are related by rotations. Recall that we measure the anisotropy using the parameter

$$\lambda = \frac{m_L - m_T}{m_L + m_T} \quad (10)$$

where $m_L > m_T$ are the two eigenvalues of the effective mass tensor[9]. The outline of this section is as follows. First, we consider the small λ limit and derive an effective statistical model. Next, we use a variational approach to study other values of λ. Finally, we apply the results to cases relevant to experiments where there are $n_v = 1, 2$ or 3 valleys.

A. Small mass anisotropy: perturbative approach

We calculate the leading in $\lambda \ll 1$ behavior of the system by mapping the problem to a classical clock model on the triangular lattice with long-range angle-dependent interactions. See App. C 2 for details.

The phonon zero-point energy in Eq. 9 can be expanded in powers of λ as $e_{\rm ph} = \sum_{n=0}^{\infty} e_n \lambda^n$ with coefficients that depend on the pattern of orbital ordering. This, in turn, is most efficiently represented by identifying a phase variable, θ_j on each site, that indicates the orientation of the preferred axis of the effective mass tensor, and which can take on one of n_v discrete values, equally spaced in units of $2\pi/n_v$ starting from $\theta_{\rm WC}$ which is the angle between one of axis of the WC and the host semiconductor. The leading correction that depends on the valley pattern is $\lambda^2 e_2$, where

$$e_2 = -\frac{1}{N_{\rm el}} \sum_{\boldsymbol{k}} \Phi^{\dagger}_{2;\boldsymbol{k}} \tilde{J}(\boldsymbol{k}) \Phi_{2;\boldsymbol{k}}, \tag{11}$$

$\Phi_{2;\boldsymbol{k}}$ is the Fourier transform of $[\cos(2\theta_i), \sin(2\theta_i)]^{\top}$ and $\tilde{J}(\boldsymbol{k})$ is a symmetric matrix (given in App. C 2, Eq. C9) obtained from the phonon spectrum of the WC with isotropic effective mass.

We have evaluated $\tilde{J}(\boldsymbol{k})$ numerically and found that its largest eigenvalue over all $\boldsymbol{k}$ arises at the $\boldsymbol{M}, \boldsymbol{M}', \boldsymbol{M}''$ points. The eigenvector at $\boldsymbol{M} = \frac{2\pi}{\sqrt{3}}[1, 0]$ is $[1, 0]^{\top}$. Therefore, were we to ignore the discreteness of the phase variables, we would conclude that the optimal pattern would be $e^{2i\theta_i} = e^{i\mathbf{R}_i \cdot \mathbf{M}}$ (or the analogous state at $\boldsymbol{M}'$ or $\boldsymbol{M}''$). This is thus a lower bound on the energy of the discrete model. Taking the constraint on values of θ_j into account, we can achieve this optimal pattern with $\theta_{WC} = 0$ so long as the number of valleys, n_v, is even. The order for $n_v = 2$ is shown in Fig. 5. For n_v odd, no state saturates the lower bound, but one can construct states that are close to the ideal state. The case of $n_v = 1$ (which is equivalent to the fully polarized state for arbitrary n_v) is treated in Sec. III D 1. For $n_v = 3$ we can construct a close to optimal state with $e^{2i\theta_i} = \frac{-1+\sqrt{3}ie^{i\mathbf{R}_i \cdot \boldsymbol{M}}}{2}$. We have checked that this configuration has lower energy than any state with no more than three electrons per unit cell, although we have not proven that it has the absolute minimum energy.

B. Arbitrary mass anisotropy: Variational Approach

We study the $T = 0$ phases at arbitrary mass anisotropy λ using variational trial states. Specifically, we assume periodic pseudo-spin patterns with various sizes of unit cell.

The energy per electron of each trial state can be written as

$$e_{\rm ph} = \frac{\nu}{2} \int_{\boldsymbol{k} \in \mathrm{BZ}'} \frac{\mathrm{d}\boldsymbol{k}^2}{(2\pi)^2} \mathrm{Tr}\left[\sqrt{\Omega^2_{\boldsymbol{k}}}\right] \tag{12}$$

where BZ′ is the folded Brillouin zone with with size $1/B$, where B is the number of sites on the new unit cell. The dynamical matrix is $\Omega^2_{\boldsymbol{q}} = \sqrt{\mathcal{W}}\mathcal{K}(\boldsymbol{q})\sqrt{\mathcal{W}}$. Here $\mathcal{W}$ and $\mathcal{K}(\boldsymbol{q})$ are $2B \times 2B$ matrices that represent the inverse mass anisotropy and the elastic matrix within the pseudo-spin pattern unit cell. See App. B for details. For instance, the period-two stripe state from the small λ limit has $B = 2$, and BZ′ corresponds to the yellow rectangle in Fig. 2.

For the three cases considered in Sec. III D ($n_v = 1, 2, 3$), we found that the following is a good fit for the energy of the preferred states (depicted in Figs. 4,5,7, respectively):

$$e^{(n_v)}_{\rm ph}(\eta)_{\rm fit} = \sqrt{\frac{2}{1+\eta^2}} \left(A_0 + A_1\eta + A_2\eta^2 + A_3\eta^3\right) \tag{13}$$

where $\eta = \sqrt{\frac{m_L}{m_T}} \in [0, 1)$. The coefficients A_j are n_v dependent. Their values are displayed in Tab. I.

C. Physical quantities from variational states

We calculate the following measurable properties using the trial states:

- Broken symmetries;
- The phonon's dispersion relations[10];
- The optical conductivity.

In particular, an expression for the conductivity can be obtained from the Kubo formula,

$$\sigma_{ab}(\omega) = \frac{4\pi}{A} \sum_n \frac{\delta(\omega - \omega_n)}{\omega} \langle 0|J_a|n\rangle \langle n|J_b|0\rangle. \tag{14}$$

where A is the area of the sample and the sum is over excited states, $|n\rangle$, with energies $\omega_n > 0$. We use $\vec{J} = \sum_i \vec{J}_i$ for the current operator, where $\vec{J}_i = e\vec{v}_i = eW_i \cdot \vec{p}_i$ is the contribution to the current of the electron at site i. After some algebra (see App. E), the optical conductivity can be written as the sum of the contribution of each

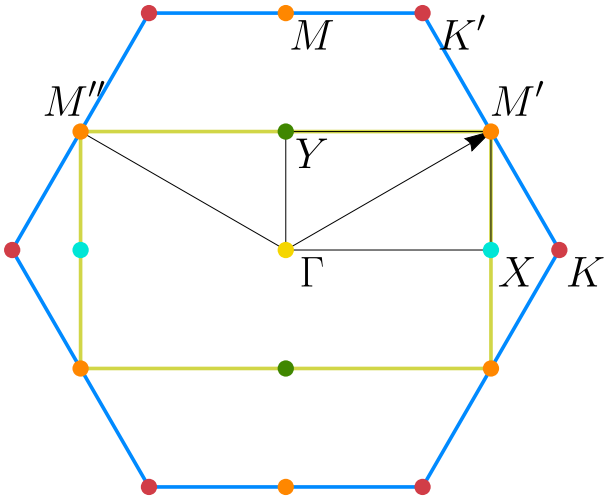


FIG. 2. Wigner crystal's Brillouin zone (boundary in blue) and folded Brillouin zone (boundary in yellow) for our candidate anisotropic Wigner crystals with period two stripe. (The black line corresponds to the path used in Fig. 6.)

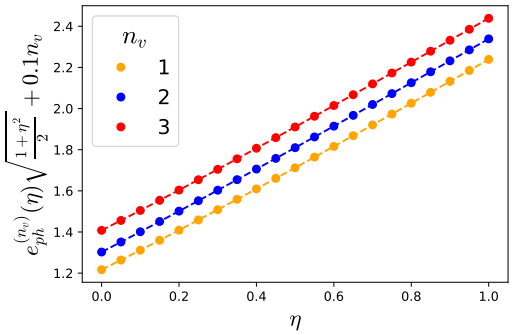


FIG. 3. Dependence of the normalized zero point phonon energy (Eq. 12) for the different candidate states with different number of valleys (n_v). The dots are obtained via numerical evaluation of the integral and the dashed line corresponds to the fit in Eq. 13. A sketch of the states are shown in Figs. 4($n_v = 1$), 5($n_v = 2$) and 7($n_v = 3$). (Note that we have introduced a shift to make the three curves distinguishable.)

optical phonon with crystal momentum zero (in the reduced BZ). Note that the above calculation is equivalent to the computation of the optical conductivity using the Kubo formula according to $\sigma(\omega) \equiv \lim_{\vec{q}\to\vec{0}} \sigma(\vec{q},\omega)$). This explains why we do not get a contribution from the acoustic modes, i.e., there is no $\delta(\omega)$ contribution. (In the presence of disorder, this acoustic contribution would appear as a low frequency feature associated with a "pinning mode.")

n_v	A_0	A_1	A_2	A_3	Fig. #
1	1.116	0.941	0.120	-0.0395	4
2	1.102	0.979	0.083	-0.0267	5
3	1.108	0.962	0.105	-0.0360	7

TABLE I. Fit parameters (Eq. 13) for the variational energy $e_{\rm ph}^{(n_v)}(\eta)_{\rm fit} = \sqrt{\frac{2}{1+\eta^2}}\left(A_0 + A_1\eta + A_2\eta^2 + A_3\eta^3\right)$ ($\eta = \sqrt{\frac{m_L}{m_T}} \in [0,1)$) of the states depicted in the respective figure.

D. Applications to 2d semiconductors

1. *One valley* ($n_v = 1$)

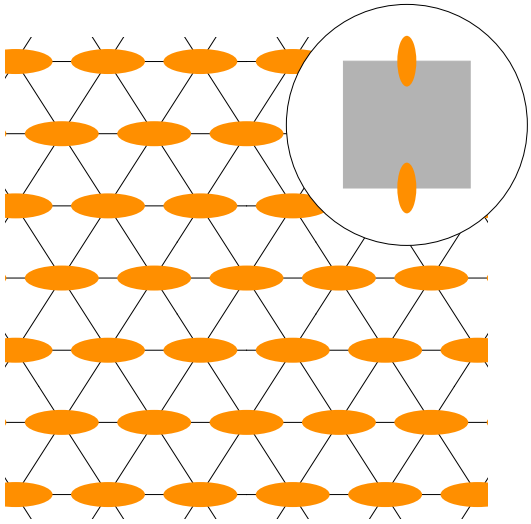


FIG. 4. Schematic real space charge density in a WC with a single anisotropic valley ($n_v = 1$). (Inset shows the orientation of the underlying crystal's BZ with respect to the WC.)

We start by considering the case of a single valley with an anisotropic effective mass, which is also relevant to the case of a system with multiple valleys in a fully valley polarized state. In this case, we only need to consider a unit cell with a single site. We find that the optimal orientation of the WC relative to the principal axes of the effective mass tensor is such that the long direction of the Fermi contour is perpendicular to one of the WC's axes. A depiction of the charge density distribution that results from the electron zero-point motion is shown in Fig. 4. A fit and plot for the energy of this variational state can be found in Tab. I and Fig. 3, respectively.

The phonon dispersion near $k = 0$ is *anisotropic*:

$$\omega_{ac,L}(\boldsymbol{k}) = \sqrt{kk_\circ\left(\cos^2(\theta_{\boldsymbol{k}}) + \eta^2\sin^2(\theta_{\boldsymbol{k}})\right)}$$
$$\omega_{ac,T}(\boldsymbol{k}) = \frac{\mathrm{v}k}{\sqrt{\cos^2(\theta_{\boldsymbol{k}}) + \eta^2\sin^2(\theta_{\boldsymbol{k}})}}. \quad (15)$$

where $\mathrm{v}^2 \approx 0.245(1-\lambda^2)$ and $\boldsymbol{k} = [k\cos(\theta_{\boldsymbol{k}}), k\sin(\theta_{\boldsymbol{k}})]$. Along symmetry directions, the lower branch is the transverse acoustic mode and the upper is the longitudinal acoustic mode, so we more generally label them L and T. The dispersion along various symmetry directions is displayed in Fig. 6a. There is no optic phonon mode, and consequently no optical response.

2. *Two valleys related by C_4 ($n_v = 2$)*

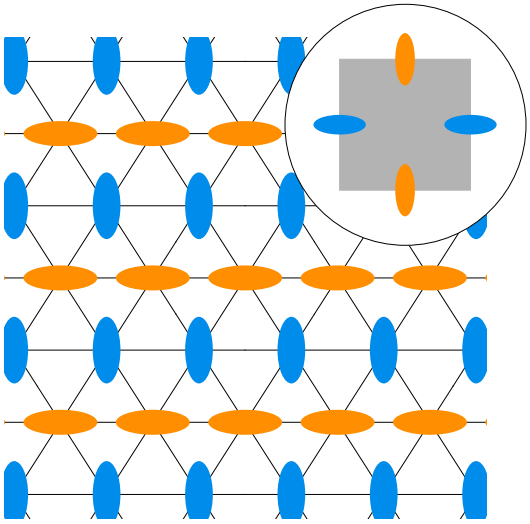


FIG. 5. Schematic real space charge density in a WC with two valleys related by C_4 symmetry ($n_v = 2$). (Inset shows the orientation of the underlying crystal's BZ with respect to the WC.)

In this section we consider a semiconductor with $PG = D_4$ and two pockets at $X = [\pi, 0]$ and $Y = [0, \pi]$, as shown in Fig. 1-a. The two valleys are related by C_4 symmetry. We consider $\eta^2 = m_T/m_L \in [0, 1)$. This section could be relevant to AlAs [11].

We have computed e_{ph} for all symmetry inequivalent valley ordered patterns with unit cells of up to six sites. The valley pattern with the lowest e_{ph} has the period-two valley-stripe order depicted in Fig. 5. A close competitor is the fully polarized state considered above. In particular, the ratio $\frac{e^{(2)}_{\mathrm{ph}} - e^{(1)}_{\mathrm{ph}}}{e^{(2)}_{\mathrm{ph}} + e^{(1)}_{\mathrm{ph}}}$ depends on the mass anisotropy and varies from 0 to approximately 6.4×10^{-2}. Our results coincide with the results of e_{ph} from Ref. [5] in the limit $\eta \to 0$, and with our results in Sec. III A in the $\eta \to 1$ limit.

There is an equal number of X and Y valley electrons per unit cell so there is no net valley polarization. However, the PG is broken to D_2 because the two flavours are at inequivalent positions in the unit cell.

As there are two electrons per unit cell, there are 4 phonon branches - two optical and two acoustic. Even though the valley pattern has only D_2 symmetry, the

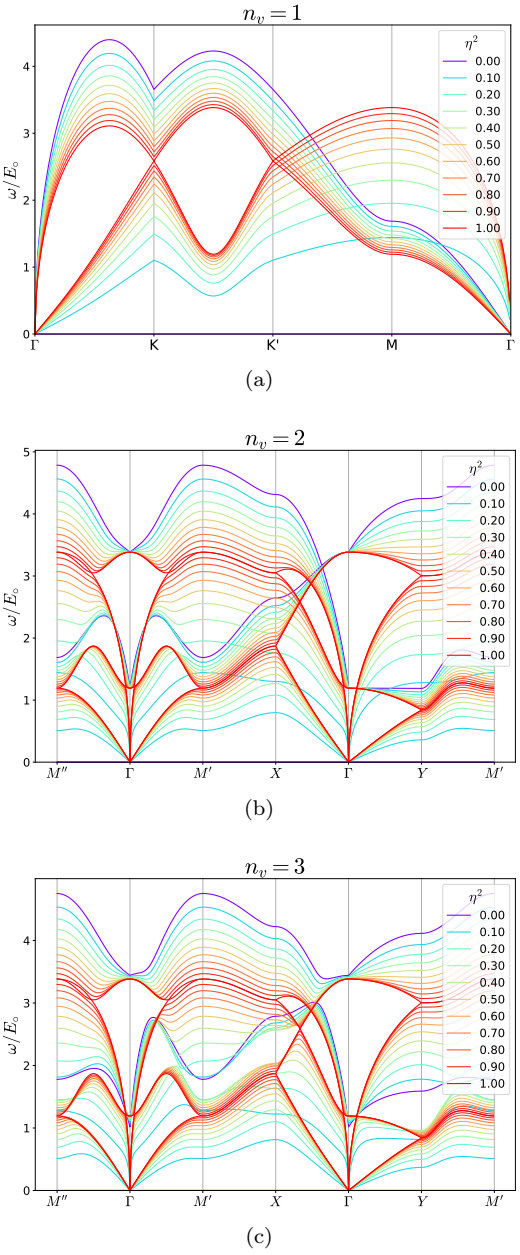


FIG. 6. Phonon band structure for the three relevant states for various values of $\eta^2 = \frac{m_T}{m_L}$. See Fig. 2 for the meaning of the momenta labels. The polarized state ($n_v = 1$) has one electron per unit cell, and so two acoustic branches, while the stripe states ($n_v = 2, 3$) have two electrons per unit cell, and hence two additional optic modes.

long wave-length acoustic modes are *isotropic*

$$\begin{aligned} \omega_{ac,L}(\boldsymbol{k}) &= \sqrt{(1-\lambda^2)\, k k_\circ} \\ \omega_{ac,T}(\boldsymbol{k}) &= \mathrm{v} k \end{aligned} \tag{16}$$

where $\mathrm{v}^2 \approx 0.245(1-\lambda^2)$ and $k_\circ = \frac{2\pi}{\sqrt{\nu}}$.

In Fig. 6b we show the phonon dispersion relations along high-symmetry directions for several values of $\eta^2 =$

m_T/m_L. Note that the phonon frequencies at $\boldsymbol{k} = \Gamma$ are independent of η.

The optical conductivity for the state depicted in Fig. 5 is the sum of contributions from the two optic modes, $\sigma(\omega) = \sigma^{(1)}(\omega) + \sigma^{(2)}(\omega)$:

$$\begin{aligned} \sigma^{(1)}_{ab}(\omega) &= \sigma_0 \lambda^2 \left[\delta_{ax}\delta_{bx}\delta(\omega-\omega_1)\right] \\ \sigma^{(2)}_{ab}(\omega) &= \sigma_0 \lambda^2 \left[\delta_{ay}\delta_{by}\delta(\omega-\omega_2)\right] \\ \omega_1 &\approx 1.19255 \\ \omega_2 &\approx 3.38394, \end{aligned} \tag{17}$$

$\sigma_0 = \frac{2\pi e^2 n_{\rm el}}{m^\star}\frac{1}{E_\circ}$ and ω is measured in units of $E_\circ$, and the imaginary part can be computed straightforwardly from Kramers-Kronig. Note that the existence of optic modes is directly related to the doubling of the WC unit cell; they would not arise, for example, in a fully valley polarized state.

3. *Three valleys related by C_3*

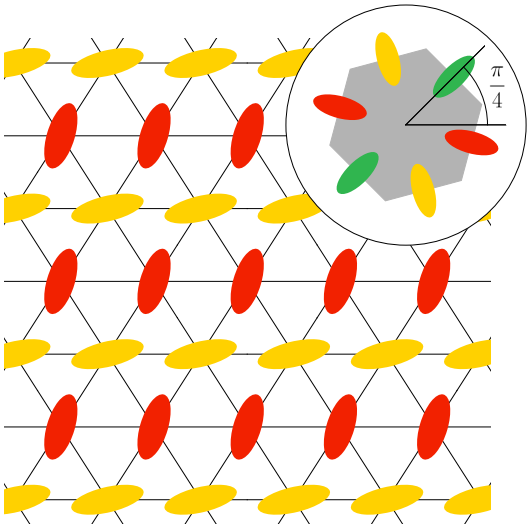


FIG. 7. Schematic charge density in a period-2 stripe valley ordered WC with three valleys ($n_v = 3$) related by C_3 symmetry. (Inset shows the orientation of the underlying crystal's BZ with respect to the WC.)

Next, we consider a system with $PG = D_6$ with three valleys related by threefold rotations. For instance, the Fermi pockets could be at the M points (the middle points of the edge of the hexagonal BZ), as shown in Fig. 1-b.

We considered unit cells with up to 3 electrons per unit cell. We find that the configuration with the lowest energy is a period-2 valley stripe pseudo-spin antiferromagnet. The fully polarized state is again a close competitor.

The valley order has momentum M (as for $n_v = 2$) but only electrons in two valleys out of the three valleys participate. The axes of the Wigner crystal no longer align with those of the underlying lattice. Instead, the horizontal axes of the WC lie preferentially at about 15° relative to one principal axes of the underlying crystal. Equivalently, the direction of the principal axes of the effective mass tensor of the unpopulated valley is oriented at 45° with respect to the horizontal axes of the WC. A cartoon of this state is shown in Fig. 7. The orientation is the equivalent to the proposed low energy state for the effective XY model $n_v = 3$ of Sec. III A.

The long wave-length longitudinal (L) and transverse (T) acoustic modes are *anisotropic*,

$$\begin{aligned} \omega_{ac,L}(\boldsymbol{k}) &= \mathrm{v}_L\sqrt{kk_\circ}\sqrt{1+\frac{\lambda}{2}\sin(2\varphi)} \\ \omega_{ac,T}(\boldsymbol{k}) &= \frac{\mathrm{v}_R k}{\sqrt{1+\frac{\lambda}{2}\sin(2\varphi)}} \end{aligned} \tag{18}$$

where $\mathrm{v}_{L/R} \propto \sqrt{1-\lambda^2}$ and $\boldsymbol{k} = k[\cos(\varphi), \sin(\varphi)]^\top$.

In Fig. 6c we show the phonon dispersion relation along high-symmetry directions for several values of $\eta^2 = m_T/m_L$. Note here that the optical phonon frequencies at $\boldsymbol{k} = \Gamma$ are η *dependent*.

This pseudo-spin order has inversion symmetry but lacks reflection symmetry. This shows up in the optical response as an η dependence of the transmission axes. The optical frequency is the sum of contributions of the two optical modes $\sigma(\omega) = \sigma^{(1)}(\omega) + \sigma^{(2)}(\omega)$,

$$\sigma^{(j)}_{ab}(\omega) = \sigma_0 D_j \left[v^{(j)}_a v^{(j)}_b \delta(\omega-\omega_j)\right] \tag{19}$$

here $\sigma_0 = \frac{2\pi e^2 n_{\rm el}}{m^\star}\frac{1}{E_\circ}$, ω_j is the frequency of optical mode j, $v^{(j)}_a = [\cos(\theta_j), \sin(\theta_j)]^T$ is a polarization vector, and D_j is the weight of the delta function. The η dependence of $\omega_j, \theta^{(j)}, D_j$ is plotted in Fig. 8.

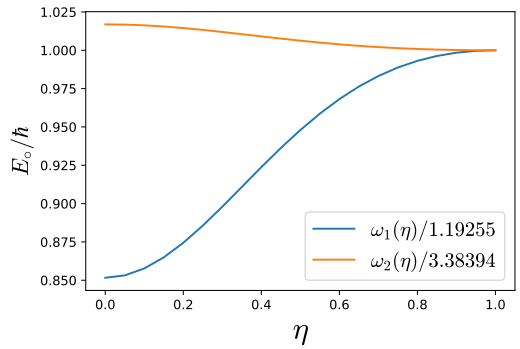
IV. INVERSION SYMMETRY BREAKING

In this section we consider possible effects that could break the degeneracy between pairs of valleys related by inversion symmetry. For this section, one should have in mind valleys at the $\pm\boldsymbol{K}$ points of a crystal with C_6 symmetry, as shown in Fig. 1-c. Our considerations may be relevant for monolayer and bilayer transition metal dichalcogenides [12; 13].

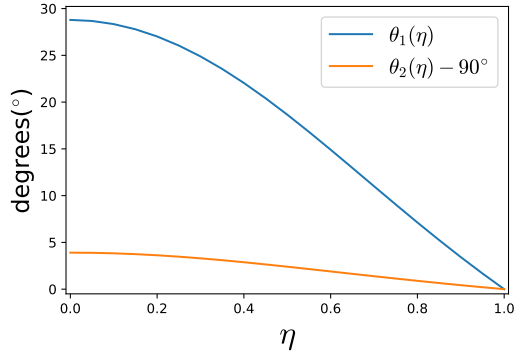
A. Trigonal warping

We start by considering trigonal warping, i.e. we go beyond the effective mass approximation and add a cubic in $\boldsymbol{p}$ term to the kinetic energy in Eq. 1:

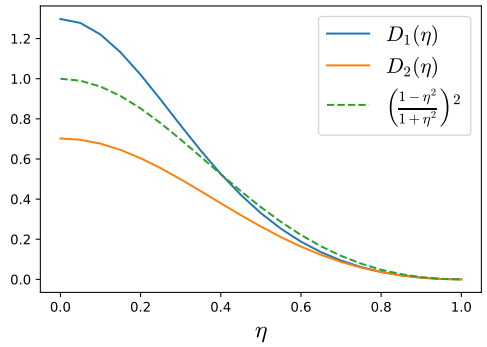
$$T_\alpha(\boldsymbol{p}) = \sum_{a,b}\frac{1}{2}W^{(\alpha)}_{ab}p_a p_b + \sum_{a,b,c}\frac{1}{3!}t^{(\alpha)}_{abc}p_a p_b p_c \tag{20}$$



(a)



(b)



(c)

FIG. 8. Dependence of parameters of the optical conductivity of stripe phase for $n_v = 3$. (*a*) phonon frequency normalized with respect to the $\eta = 1$ model ($E_\circ = \sqrt{\frac{e^2}{4\pi\epsilon a^3 m^\star}}$). (*b*) Deviation of the transmission axes from the isotropic limit ($\eta \to 1^-$). (*c*) weight of optical conductivity for the optical bands. The green line is a guide for comparison to the weight of optical conductivity of the $n_v = 2$ state.

If valleys α, α' are related by inversion, we must have $W^{(\alpha)} = W^{(\alpha')}$ and $t^{(\alpha)} = -t^{(\alpha')}$. Needless to say, in order to insure that the Hamiltonian is bounded below, we need to include higher order terms as well - at least to order p^4. However, since we will treat the effects of trigonal warping in perturbation theory, we do not need to include these higher order terms explicitly.

For simplicity, we consider a pair of C_3-symmetric valleys related by inversion. The dispersion relation with $p_x = p\cos(\theta_{\boldsymbol{p}})$ and $p_y = p\sin(\theta_{\boldsymbol{p}})$ simplifies to

$$T_\alpha(\boldsymbol{p}) = \frac{p^2}{2m} + tp^3\cos(3\theta_{\boldsymbol{p}} - 3\phi_\alpha), \tag{21}$$

where ϕ_α is an angle that fixes the orientation of the Fermi contour with respect to Wigner crystal coordinates. If α and α' label a pair of valleys related by inversion, then $3\phi_{\alpha'} = 3\phi_\alpha + \pi$. We rescale the position and momentum variables as in Sec. II C and use the harmonic approximation. The effective Hamiltonian in rescaled variables is

$$\begin{aligned} h_{\text{eff}} = &\sum_i \frac{1}{2}P_i^2 + \sum_{i,j}\frac{1}{2}Q_i K_{ij} Q_j \\ &+ \epsilon_3 \sum_i [\mathrm{e}^{3\mathrm{i}\phi_{\alpha_i}}(P_{i;x} - \mathrm{i}P_{i;y})^3 + h.c.] \end{aligned} \tag{22}$$

where $\epsilon_3 = t\sqrt{E_\circ (m^\star)^3}$ goes like $\epsilon_3 \propto r_{\text{s}}^{-3/4}$ as $r_{\text{s}} \to \infty$. Due to the C_3 symmetry, the next leading angle-dependent correction to the effective mass approximation is at order P^6 which has a prefactor that scales as $1/r_{\text{s}}^3$.

The details of the calculations can be found in App. C 3. Here we give an overview of the results. As we did in Sec. III A, we expand the phonon energy with trigonal warping $e_{\text{ph}}^{\triangle} = \sum_{n=0}^{\infty} e_n^{\triangle}\epsilon_3^n$ where each $e_n^{\triangle}$ in principle can depend on the pattern of valley ordering. We find that $e_0^{\triangle}$ is independent of the ordering pattern. $e_0^{\triangle} + \epsilon_3 e_1^{\triangle}$ is an effective potential for $P_{CM} = \frac{1}{N_{\text{el}}}\sum_i P_i$. Upon including a quartic term in P_{CM} for stability, the potential is optimized by $P_{CM} = 0$. (Were we to consider a system whose $\epsilon_3 = 0$ ground state lacks C_3 rotation symmetry, there could be a term $\epsilon_3 P_{CM}$ in the effective potential so there would be a non-zero $P_{CM} \propto \epsilon_3$. We can interpret this as a displacement of the band minima in the BZ of the underlying crystal.)

The next order correction to the energy is $\epsilon_3^2 e_2^{\triangle}$ with

$$e_2^{\triangle} = -\frac{1}{N_{\text{el}}}\sum_{\boldsymbol{k}} \Phi_{3;\boldsymbol{k}}^\dagger \tilde{J}_3(\boldsymbol{k})\Phi_{3;\boldsymbol{k}}, \tag{23}$$

where $\Phi_{3,\boldsymbol{k}}$ is the Fourier transform of $[\cos(3\phi_i), \sin(3\phi_i)]^\top$ and $\tilde{J}_3(\boldsymbol{k})$ is a symmetric matrix (given in App. C 3, Eq. C15) obtained from the phonon spectrum of the WC with isotropic effective mass. The largest eigenvalue of $\tilde{J}_3(\boldsymbol{k})$ is at $\boldsymbol{k} = 0$ with eigenvector $[0, 1]^\top$. Therefore, the state with lowest energy is a valley-ferromagnet ($\boldsymbol{k} = 0$) and $3\phi_i$ is constant and equal to $\pm\frac{\pi}{2}$, as shown schematically in Fig. 9.

B. Berry curvature

The Berry curvature can be thought of as a magnetic field in momentum space. Therefore, in a semi-classical

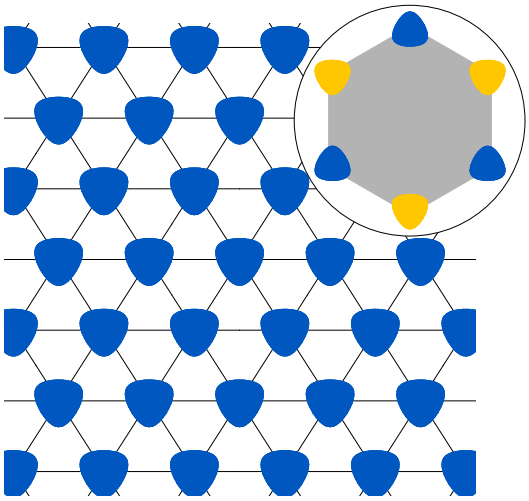


FIG. 9. Schematic real-space charge density of the valley-polarized state for a 2DEG with two valleys related by inversion symmetry at the $\pm\boldsymbol{K}$ points and with trigonal warping. (Inset shows the orientation of the underlying crystal's BZ with respect to the WC)

picture (like ours) we need to modify the commutation relations between the physical position and momentum variables (see e.g. Refs. [14; 15]) to

$$\begin{aligned} \{r_a^{\rm phys}, r_b^{\rm phys}\} &= \varepsilon_{ab}\mathcal{B}; \\ \{p_a^{\rm phys}, p_b^{\rm phys}\} &= 0; \\ \{r_a^{\rm phys}, p_b^{\rm phys}\} &= \delta_{ab}. \end{aligned} \tag{24}$$

where ε_{ab} is the Levi-Civita symbol. Here $\mathcal{B}$ is the Berry curvature, assumed to be constant for simplicity. We identify the canonical variables

$$\begin{aligned} r_a^{\rm canon} &= r_a^{\rm phys} + \frac{1}{2}\mathcal{B}\varepsilon_{ab}p_b^{phys}, \\ p_a^{\rm canon} &= p_a^{phys}. \end{aligned} \tag{25}$$

that are used to quantize the theory.

At large $r_{\rm s}$, the Coulomb energy dominates so the electrons form a triangular lattice. Next, we use the harmonic approximation as in Sec. II C to get the same Hamiltonian as Eq. 7 with Q_i and P_i replaced by $Q_i^{\rm phys} = \sqrt{E_\circ m^\star}q_i^{\rm phys}$ and $P_i^{\rm phys} = \frac{p_i^{\rm phys}}{\sqrt{E_\circ m^\star}}$. We then ought to write the Hamiltonian in terms of the (rescaled) canonical variables

$$\begin{aligned} Q_a^{\rm canon} &= Q_a^{\rm phys} + \bar{\mathcal{B}}\varepsilon_{ab}P_b^{\rm phys} \\ P_a^{\rm canon} &= P_a^{\rm phys} \end{aligned} \tag{26}$$

where $\bar{\mathcal{B}} = E_\circ m^\star \mathcal{B}$ is a dimensionless measure of the Berry curvature that scales as $r_{\rm s}^{-3/2}$ when $r_{\rm s} \to \infty$. From now on we will write $Q_i(P_i)$ for $Q_i^{\rm canon}(P_i^{\rm canon})$.

A pair of valleys related by inversion symmetry has the same mass tensor but opposite Berry curvature. To describe the latter, we introduce an Ising variable $\mu = \pm 1$ that we identify with the two possible valleys polarization within the pair so that we can write the Berry curvature of the pair of valleys as $\bar{\mathcal{B}} = \mu\left|\bar{\mathcal{B}}\right|$.

The dimensionless effective Hamiltonian is

$$\begin{aligned} h_{\rm eff} &= \frac{H_{\rm eff}}{E_\circ} = h_0 + \left|\bar{\mathcal{B}}\right|h_1 + \left|\bar{\mathcal{B}}\right|^2 h_2; \\ h_0 &= \sum_i \frac{1}{2}P_i^2 + \sum_{i,j}\frac{1}{2}Q_iK_{ij}Q_j; \\ h_1 &= -\sum_{ij}\frac{\mu_i(\varepsilon P)_iK_{ij}Q_j + \mu_jQ_jK_{ij}(\varepsilon P)_i}{2}; \\ h_2 &= \sum_{ij}\frac{\mu_i\mu_j(\varepsilon P)_iK_{ij}(\varepsilon P)_i}{2}; \end{aligned} \tag{27}$$

here the sums of i and j are over the sites of the triangular lattice. Note that the pseudo-spin pattern enters Eq. 27 via the sign of the Berry curvature (μ_i). εP denotes the product of the Levi-Civita tensor and the vector P.

Recall that we are interested in the large $r_{\rm s}$ limit. Therefore, we can treat h_1 and h_2 via degenerate perturbation theory over the ground state manifold of h_0. The extensive degeneracy stems from the μ_i variables that do not enter h_0.

The leading effect of the Berry curvature to the energy is of order $r_{\rm s}^{-3}$ and corresponds to the expectation value of $E_\circ\left|\bar{\mathcal{B}}\right|h_1$ on the ground state of h_0. It turns out that $\langle h_1\rangle = 0$ so that the leading effect of the Berry curvature on the energy comes with a prefactor of $E_\circ\left|\bar{\mathcal{B}}\right|^2 \propto r_{\rm s}^{-9/2}$ which is parametrically smaller than the effect of trigonal warping that is generically present. Therefore, we will not calculate these higher order corrections of the Berry curvature to the energy.

V. EXTENSIONS

We now consider two extensions of our methods to slightly more complicated systems:

A. External moiré potential

While in typical semiconductors, commensurability effects with the underlying lattice are negligible, if the period is extended - as is the case in moiré materials - such effects can be significant. Consider adding an external periodic potential that is commensurate with the Wigner crystal. Still keeping terms that to quadratic order in displacements about the classical ground-state, this corresponds to adding $\frac{1}{2}K_{\rm moiré}\hat{x}_i^2$ at every site of the WC. The elastic matrix gets shifted by a diagonal matrix

$$K_{\rm moiré}(\boldsymbol{q}) = K(\boldsymbol{q}) + \kappa_{\rm moiré}\mathbb{I};$$

We will treat $\kappa_{\text{moiré}} = \frac{K_{\text{moiré}}}{E_\circ^2 m^\star}$ as an additional parameter. In the case of isotropic mass tensors, the (normalized) phonon energies are

$$\omega_{\text{moiré},A}(\boldsymbol{q}) = \sqrt{\omega_A^2(\boldsymbol{q}) + \kappa_{\text{moiré}}}, \quad A = 1, 2.$$

In particular, as expected, this gaps the acoustic modes.

We have redone the above calculations of the preferred orbital ordering in the presence of a commensurability potential for a subset of the various cases considered above. For instance, over the range of parameters we have explored, for the case of a tetragonal symmetry considered Sec. III A in the limit of small mass anistropy ($\lambda \ll 1$), we find that the pseudo-spin configuration with least energy is still the period-two stripe.

We have also repeated the calculations of Sec. IV A in the presence of trigonal warping and no mass anisotropy, and again we find that commensurability effects do not change the character of the preferred orbital order. To reach this conclusion, we evaluated the kernel of Eq. 23 with the modified dispersion for $\kappa_{\text{moiré}} \in \{1/2, 1, 2\}$ and, additionally, performed a large $\kappa_{\text{moiré}}$ expansion in App. C 3. In both regimes the configuration with least energy corresponds to $e^{3i\phi_i} = \pm i$.

Needless to say, the biggest effects of a moiré potential (which we do not explore here) is felt if the system is slightly incommensurate. Here, rather than favoring a WC with an incommensurate period, the preferred state is a commensurate WC with a small concentration of interstitials or vacancies. In a recent study [16], it was shown that a small concentration of interstitials can have an outsized effect on the magnetic (spin) order of a WC.

B. 3 dimensional Wigner crystals

We can adapt our considerations to the 3-dimensional electron gases with multiple valleys. In the following, we go through the analysis and make emphasis on the distinctions with the 2-dimensional problem.

At large r_s, Coulomb interactions dominate and the system minimizes the energy by forming a WC with body-centered cubic structure [17] with lattice vectors $[\frac{1}{2}, \frac{1}{2}, \frac{1}{2}], [\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}], [\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}]$. The leading correction in $1/r_s$ to the energy is again the zero-point phonon energy, which depends on the pseudo-spin pattern. If W is the inverse mass tensor, we take $m^\star = \frac{3}{\text{Tr}[W]}$ as the mass scale (which is equal for every valley related by the PG action) and the mass anisotropy as

$$\begin{aligned}\lambda &= \sqrt{\frac{1}{6}\,\text{Tr}\left[(m^\star W - \mathbb{1})^2\right]} \\ &= \frac{\sqrt{w_1^2 + w_2^2 + w_3^2 - w_1 w_2 - w_2 w_3 - w_1 w_3}}{w_1 + w_2 + w_3}\end{aligned} \tag{28}$$

where w_i are the eigenvalues of W.

For small λ, we can derive an effective model for the valley pseudo-spin as in Sec. III A. The pseudo-spins are generalized clock variables with states in a group S. The elements of S correspond to symmetry operations in the PG of the host crystal that permute the valleys. For instance, consider a system with cubic symmetry and valleys at $[\pi, 0, 0], [0, \pi, 0], [0, 0, \pi]$. Then, $S = \mathbb{Z}_3$ corresponds to 3-fold rotation along the $[1, 1, 1]$ axis.

The effective energy per electron is better understood in terms of quadrupolar moments Q_i. These are related to the inverse mass tensor of electron at site i by $W_i = (\mathbb{1} + \lambda Q_i)/m^\star$. At every site, we associate to $\mathbf{g}_i \in S$ the quadrupolar moment obtained by applying $\mathbf{g}$ to a reference Q_{WC}. (Q_{WC} plays the role of θ_{WC} from Sec. III.)

The leading correction in λ to the energy is $\lambda^2 e_2$, where

$$e_2 = -\frac{1}{N_{\text{el}}} \sum_{\boldsymbol{k}} \tilde{Q}^\dagger_{\boldsymbol{k}} \tilde{R}(\boldsymbol{k}) \tilde{Q}_{\boldsymbol{k}}. \tag{29}$$

The sum is over the 3-dimensional BZ. $\tilde{Q}_{\boldsymbol{k}}$ is the Fourier transform of Q_i (seen as a 5 component vector) and $\tilde{R}(\boldsymbol{k})$ is a symmetric matrix (the expression can be found in Sec. C 6) obtained from the phonon spectrum of the 3d WC with $\lambda = 0$.

We have evaluated $\tilde{R}(\boldsymbol{k})$ in a $12 \times 12 \times 12$ momentum grid and found that the largest eigenvalue over all $\boldsymbol{k}$ arises at $\boldsymbol{N} = [\pi, \pi, 0]$ (and symmetry-related momenta) and has $e_2 \approx -0.01283$. The corresponding eigenvector ($\tilde{Q}_{\boldsymbol{k}} = \delta_{\boldsymbol{k},\boldsymbol{N}} \tilde{Q}^*$) is interpreted in terms of the mass tensors as follows. There is pseudo-spin ordering with a two-site unit cell (the new lattice vectors are $[1, 1, 1]$, $[\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}]$, $[\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}]$). Equivalently, the pseudo-spins are constant on the $x + y = k$ planes but oscillate in the $[1, 1, 0]$ direction. The principal axes of the inverse mass tensors W (and therefore of Q) are along the $x-y, x+y, z$ directions of the WC. Q_i has eigenvalues

$$\begin{aligned} q^{(A)}_{x+y} &= 0; q^{(A)}_{x-y} = +\sqrt{3}; q^{(A)}_z = -\sqrt{3} \\ q^{(B)}_{x+y} &= 0; q^{(B)}_{x-y} = -\sqrt{3}; q^{(B)}_z = +\sqrt{3} \end{aligned} \tag{30}$$

where $q^{(Y)}_d$ is the eigenvalue of Q_i (as a matrix) along the d direction of the electron on sublattice Y.

As in the 2d case, the allowed Q_i configurations are determined by the host semiconductor. If we restrict to configurations with two sites per unit cell, we can write $\tilde{Q}_{\boldsymbol{k}} = \cos(\gamma)\delta_{\boldsymbol{k},\boldsymbol{0}} \tilde{Q}_\Gamma + \sin(\gamma)\delta_{\boldsymbol{k},\boldsymbol{Q}} \tilde{Q}_{\boldsymbol{Q}}$. Here γ is an angle, $\tilde{Q}_\Gamma$ and $\tilde{Q}_{\boldsymbol{Q}}$ are (properly normalized) symmetric traceless matrices, and $\boldsymbol{Q}$ is a crystal momentum such that $2\boldsymbol{Q}$ is a reciprocal lattice vector. In a BCC crystal, there are 8 such momenta: Γ, $\boldsymbol{N}$ (plus other 5 symmetry related momenta) and $\boldsymbol{H} = [2\pi, 0, 0]$.

For instance, let us go back to the example of a host crystal with cubic symmetry and valleys at $[\pi, 0, 0]$ and related momenta. We optimize over states with unit cells with two sites and find that the lowest energy configuration has $e_2 \approx -0.01267$ and corresponds to a pseudo-spin ferromagnet with the longitudinal axes of mass tensor

aligned with one of the nearest-neighbour directions (e.g. $[1,1,0]$). Note that the energy of this configuration is larger than the energy of the configuration we discussed above but that configuration is not allowed by the host crystal band structure.

VI. POSSIBILITIES FOR THERMAL AND QUANTUM MELTING

Even for the case of a single, isotropic valley, the nature of the transitions from the WC to the symmetric fluid phase is unsettled. The presence of long-range Coulomb interactions preclude a single first-order transition [18], making it likely that the melting is always a multi-step transition. For the thermal melting, a two step Halperin-Nelson scenario with an intermediate hexatic phase is certainly possible [19]. The nature of the quantum melting (as a function of decreasing r_s) is still less well understood.

In the WC in a host with C_4 symmetry, the pseudo-spin order breaks the C_4 down to C_2 (Sec. III D 2). It is thus plausible that the orientational order survives for a range of T and/or r_s beyond the point at which the translational order of the WC is melted. This naturally leads to the expectation that there is a two-stage melting process, via a Kosterlitz-Thouless transition from a pseudo-spin ordered WC to an electronic Ising-nematic fluid, followed by an Ising transition to the isotropic fluid phase. Other possiblities are that there could be an orbital ordering transition within the WC phase from the period two striped phase we found at large r_s and $T = 0$ to a uniform pseudo-spin ordered WC phase at smaller r_s or larger T. It is also possible to imagine the existence of a smectic phase intermediate between the WC and the nematic phases.

In the WC in a host with C_6 symmetry, similar considerations apply. Here, the rotation symmetry in the pseudo-spin ordered WC is broken from C_6 to C_2, so the corresponding nematic phase would be an electronic three-state-Potts nematic fluid. Moreover, because the orbital order also breaks chiral symmetry, there is the possibility of vestigial chiral order once the WC is melted - either coexisting with nematic order or as a pure chiral fluid phase.

In the case of AlAs, there is experimental evidence that as a function of increasing r_s, there is a transition to a fully pseudo-spin polarized Ising-nematic fluid at a smaller value of r_s than the critical value for WC (or more directly insulating state) formation[11]. It is probably not possible for there to be a continuous transition from such a state to the pseudo-spin stripe ordered WC state we found for large r_s. However, as already mentioned, it is not impossible to imagine that there is a transition within the WC phase from a stripe ordered WC phase (as in Fig. 5) at large r_s to a fully pseudo-spin polarized WC (as in Fig. 4) for somewhat smaller r_s. Another possibility is that there may be a small extrinsic energy splitting between valleys (e.g. from a small strain) which would tip the delicate energy balance in favor of the fully polarized phase.

VII. SUMMARY

The "pure" 2DEG is one of the paradigmatic problems in condensed matter physics. Moreover, the accuracy of the effective mass approximation in many lightly doped semiconductors permits this problem to be experimentally realized with high fidelity in carefully constructed semiconductor devices. Experimentally observable signatures of a WC phase at large r_s have been studied in devices of ever increasing quality [20–25]. However, even in the context of the effective mass approximation, there are typically some additional features of the 2DEG in real devices that are both important for interpreting experiments, and also conceptually interesting as extensions of the basic theory. Specifically, depending on the symmetry and details of the band-structure of the semiconductor in question, the effective mass tensor can be anisotropic, there can be additional pseudo-spin degrees of freedom associated with distinct valleys of the underlying band-structure, and there can be forms of symmetry breaking that rely on corrections to the effective mass approximation (for example, trigonal warping).

Here we have analyzed several robust effects of this additional solid-state richness on the nature of the WC phase at large r_s. Notice that despite pertaining to the crystalline phase, these are all quantum effects (forms of "order by disorder") in that they all enter through the form of the effective electron kinetic energy operator. Some of our salient results are: 1) When there is a valley pseudo-spin degree of freedom ($n_v > 1$), pseudo-spin ordering always occurs at a temperature scale that is smaller than the classical ordering scale only by a factor of $r_s^{-1/2}$ or r_s^{-2}. 2) In a tetragonal semiconductor, with two valleys as shown in Fig. 1-a, the valley ordered state is the period 2 striped pseudo-spin antiferromagnet state shown in Fig. 5 that also breaks rotational symmetry down to C_2 - i.e. it has a nematic component. This case is potentially relevant to the insulating state of AlAs 2DEGS [26]. 3) For a hexagonal semiconductor with three valleys that transform trivially under time-reversal, as shown in Fig. 1-b, the ordered state, shown in Fig. 7, is again a period 2 stripe pseudo-spin antiferromagnet state, but one with both nematic and chiral components. 4) For a hexagonal semiconductor with two valleys that transform into one another under time-reversal, as shown in Fig. 1-c, the ground-state is the orbital (valley) ferromagnet shown in Fig. 9, which is both chiral and breaks time-reversal symmetry. We thus expect this state to have orbital loop-current ordering. The same valley-polarized state is obtained in the case of a moiré potential with a commensurate density of electrons. Such a state may be realized in transition metal dichalcogenide heterobilayers, where generalized Wigner

crystal states were recently discovered at commensurate fillings [27–29].

Long though this paper may be, our analysis is not exhaustive. There are a variety of other possible valley geometries, for example the one shown in Fig. 1-d, that could give rise to still other forms of orbitally ordered crystalline states - but these can be straightforwardly analyzed by the methods introduced here. Since the valleys transform into one another, in all cases there is strong pseudo-spin orbit coupling, i.e. there typically do not exist even approximate symmetries that operate on the pseudo-spins while leaving the spatial coordinates invariant. (Indeed, it is worth noting that even if the effective mass tensor is isotropic, corrections to the effective mass approximation always break rotational symmetry, so the WC is always orientationally pinned with respect to the symmetry axes of the underlying semiconductor.) If there is strong spin-orbit coupling in the underlying semiconductor, this can lead to strong coupling between the electron spin and pseudo-spin in the effective theory. This, in turn, could lead to interesting consequences for spin ordering in the WC that we have not yet explored.

Finally, we note that there are presumably many possible ways that the orbital order we have established in the WC phase could impress itself on the proximate quantum or thermal phases reached by melting or partially melting the WC. We have speculated a bit about this in Sec. VI, but this largely remains unexplored territory.

ACKNOWLEDGEMENTS

We acknowledge useful discussions with Akshat Pandey and with Ilya Esterlis who also provided helpful comments on the manuscript. This work was supported in part by the National Science Foundation under Grant DMR-2000987 at Stanford (SAK and VC). EB was supported by the European Research Council (ERC) under grant HQMAT (Grant Agreement No. 817799), the Israel-US Binational Science Foundation (BSF), and the Minerva Foundation. This work was performed in part at Aspen Center for Physics, which is supported by National Science Foundation grant PHY-1607611.

* fvcalvera@stanford.edu
† kivelson@stanford.edu
‡ erez.berg@weizmann.ac.il.

[1] Tsuneya Ando, Alan B. Fowler, and Frank Stern, "Electronic properties of two-dimensional systems," Rev. Mod. Phys. **54**, 437–672 (1982).
[2] M Shayegan, EP De Poortere, O Gunawan, YP Shkolnikov, E Tutuc, and K Vakili, "Two-dimensional electrons occupying multiple valleys in alas," physica status solidi (b) **243**, 3629–3642 (2006).
[3] E. Wigner, "On the interaction of electrons in metals," Phys. Rev. **46**, 1002–1011 (1934).
[4] Erez Berg, Mark S. Rudner, and Steven A. Kivelson, "Electronic liquid crystalline phases in a spin-orbit coupled two-dimensional electron gas," Phys. Rev. B **85**, 035116 (2012).
[5] P. G. Silvestrov and O. Entin-Wohlman, "Wigner crystal of a two-dimensional electron gas with a strong spin-orbit interaction," Phys. Rev. B **89**, 155103 (2014).
[6] We ignore short range interactions that depend on the valley flavour because these corrections are suppressed in the low density limit with respect to the contributions we take into account.
[7] Lynn Bonsall and A. A. Maradudin, "Some static and dynamical properties of a two-dimensional wigner crystal," Phys. Rev. B **15**, 1959–1973 (1977).
[8] The extensive spin-related degeneracies that arise in the absence of exchange interactions persist even in the presence of spin-orbit coupling so long as inversion and/or time-reversal symmetry are preserved; where these symmetries are broken, spin-valley locking can lift this degeneracy even in the absence of exchange interactions.
[9] We use the subscripts L and T because in the examples we have in mind (Figs. 1(a) and 1(b)) the two masses correspond to the longitudinal and transversal masses.
[10] See App. D for details on how we calculate the phonon dispersion relations near Γ.
[11] Md. S. Hossain, M. K. Ma, K. A. Villegas-Rosales, Y. J. Chung, L. N. Pfeiffer, K. W. West, K. W. Baldwin, and M. Shayegan, "Spontaneous valley polarization of itinerant electrons," Phys. Rev. Lett. **127**, 116601 (2021).
[12] Tomasz Smoleński, Pavel E Dolgirev, Clemens Kuhlenkamp, Alexander Popert, Yuya Shimazaki, Patrick Back, Xiaobo Lu, Martin Kroner, Kenji Watanabe, Takashi Taniguchi, *et al.*, "Signatures of wigner crystal of electrons in a monolayer semiconductor," Nature **595**, 53–57 (2021).
[13] You Zhou, Jiho Sung, Elise Brutschea, Ilya Esterlis, Yao Wang, Giovanni Scuri, Ryan J Gelly, Hoseok Heo, Takashi Taniguchi, Kenji Watanabe, *et al.*, "Bilayer wigner crystals in a transition metal dichalcogenide heterostructure," Nature **595**, 48–52 (2021).
[14] M. C. Chang and Q Niu, "Berry curvature, orbital moment, and effective quantum theory of electrons in electromagnetic fields," J. Phys. Condens. Matter **20**, 193202 (2008).
[15] Di Xiao, Ming-Che Chang, and Qian Niu, "Berry phase effects on electronic properties," Rev. Mod. Phys. **82**, 1959–2007 (2010).
[16] Kyung-Su Kim, Chaitanya Murthy, Akshat Pandey, and Steven A Kivelson, "Interstitial-induced ferromagnetism in a two-dimensional wigner crystal," arXiv preprint arXiv:2206.07191 (2022).
[17] K Fuchs, "A quantum mechanical investigation of the cohesive forces of metallic copper," Proceedings of the Royal Society of London. Series A-Mathematical and Physical Sciences **151**, 585–602 (1935).
[18] Boris Spivak and Steven A. Kivelson, "Phases intermediate between a two-dimensional electron liquid and wigner crystal," Phys. Rev. B **70**, 155114 (2004).

[19] B. I. Halperin and David R. Nelson, "Theory of two-dimensional melting," Phys. Rev. Lett. **41**, 121–124 (1978).
[20] C.-C. Li, J. Yoon, L. W. Engel, D. Shahar, D. C. Tsui, and M. Shayegan, "Microwave resonance and weak pinning in two-dimensional hole systems at high magnetic fields," Phys. Rev. B **61**, 10905–10909 (2000).
[21] V. J. Goldman, M Santos, M Shayegan, and J. E. Cunningham, "Evidence for two-dimentional quantum wigner crystal," Phys. Rev. Lett. **65**, 2189–2192 (1990).
[22] C.-C. Li, L. W. Engel, D. Shahar, D. C. Tsui, and M. Shayegan, "Microwave conductivity resonance of two-dimensional hole system," Phys. Rev. Lett. **79**, 1353–1356 (1997).
[23] E. Y. Andrei, G. Deville, D. C. Glattli, F. I. B. Williams, E. Paris, and B. Etienne, "Observation of a magnetically induced wigner solid," Phys. Rev. Lett. **60**, 2765–2768 (1988).
[24] R. L. Willett, H. L. Stormer, D. C. Tsui, L. N. Pfeiffer, K. W. West, and K. W. Baldwin, "Termination of the series of fractional quantum hall states at small filling factors," Phys. Rev. B **38**, 7881–7884 (1988).
[25] Jongsoo Yoon, C. C. Li, D. Shahar, D. C. Tsui, and M. Shayegan, "Wigner crystallization and metal-insulator transition of two-dimensional holes in gaas at $B = 0$," Phys. Rev. Lett. **82**, 1744–1747 (1999).
[26] Md S Hossain, MK Ma, KA Villegas Rosales, YJ Chung, LN Pfeiffer, KW West, KW Baldwin, and Mansour Shayegan, "Observation of spontaneous ferromagnetism in a two-dimensional electron system," Proceedings of the National Academy of Sciences **117**, 32244–32250 (2020).
[27] Yang Xu, Song Liu, Daniel A Rhodes, Kenji Watanabe, Takashi Taniguchi, James Hone, Veit Elser, Kin Fai Mak, and Jie Shan, "Correlated insulating states at fractional fillings of moiré superlattices," Nature **587**, 214–218 (2020).
[28] Emma C Regan, Danqing Wang, Chenhao Jin, M Iqbal Bakti Utama, Beini Gao, Xin Wei, Sihan Zhao, Wenyu Zhao, Zuocheng Zhang, Kentaro Yumigeta, *et al.*, "Mott and generalized wigner crystal states in wse2/ws2 moiré superlattices," Nature **579**, 359–363 (2020).
[29] Xiong Huang, Tianmeng Wang, Shengnan Miao, Chong Wang, Zhipeng Li, Zhen Lian, Takashi Taniguchi, Kenji Watanabe, Satoshi Okamoto, Di Xiao, *et al.*, "Correlated insulating states at fractional fillings of the ws2/wse2 moiré lattice," Nature Physics **17**, 715–719 (2021).
[30] W. J. Carr, "Energy, specific heat, and magnetic properties of the low-density electron gas," Phys. Rev. **122**, 1437–1446 (1961).
[31] M. H. Cohen and F. Keffer, "Dipolar sums in the primitive cubic lattices," Phys. Rev. **99**, 1128–1134 (1955).

Appendix A: Elastic matrix

The elastic matrix can be written as a sum of two parts (set $e^2 = m^\star a_c^{1/2}$ in the expressions of Ref. [7])

$$K(\boldsymbol{k}) = \mathcal{B}^{<}(\boldsymbol{k}) + \mathcal{B}^{>}(\boldsymbol{k}) \tag{A1}$$

where each term is defined as

$$\mathcal{B}^{<}_{ab}(\mathbf{k}) = \left[\sum_{\boldsymbol{G}\in\Lambda^\vee} (\mathcal{F}_{ab}(\boldsymbol{G}+\boldsymbol{k}) - \mathcal{F}_{ab}(\boldsymbol{G}))\right] \tag{A2}$$

$$\mathcal{B}^{>}_{ab}(\mathbf{k}) = \left[\sum_{\boldsymbol{R}\in\Lambda} 2\sin^2(\boldsymbol{k}\cdot\boldsymbol{R}/2)\mathcal{E}_{ab}(\boldsymbol{R})\right] \tag{A3}$$

$$\mathcal{F}_{ab}(\boldsymbol{q}) = (aq)_a(aq)_b \frac{\sqrt{\nu'}}{q} \operatorname{Erfc}\left(\sqrt{\frac{\pi q^2}{\nu'}}\right) \tag{A4}$$

$$\mathcal{E}_{ab}(\boldsymbol{R}) = a^2 \frac{\partial^2}{\partial R_a \partial R_b} \frac{\sqrt{\Omega}}{R} \operatorname{Erfc}\left(\sqrt{\frac{\pi R^2}{\nu}}\right) \tag{A5}$$

with $\nu = \frac{\sqrt{3}}{2} = (n_{\rm el}a^2)^{-1}$ and $\nu' = (2\pi)^2/\nu$. We denote the WC lattice without the origin by $\Lambda_\times$ and the reciprocal lattice without the origin by $\Lambda^\vee_\times$. The sums converge fast because they are exponentially suppressed in the distance from the center of the respective lattice.

Below we calculate the dispersion relation around high-symmetry points. These have interesting features that show up in the density of states.

1. Dispersion relations

a. $\boldsymbol{k} = \boldsymbol{\Gamma} + \boldsymbol{q}$*:* the elastic matrix is

$$K(\boldsymbol{q})_{ab} = \frac{\xi_0}{\xi}\xi_a\xi_b + \kappa^\Gamma_{abcd}\xi_c\xi_d + \mathrm{O}(\xi^3) \tag{A6}$$

where $\xi = aq$, $\xi_0 = k_\circ a$ and

$$\kappa^\Gamma_{abcd} = \kappa^\Gamma_\parallel \frac{\delta_{ac}\delta_{bd} + \delta_{ad}\delta_{bc}}{2} + \kappa^\Gamma \frac{\epsilon_{ac}\epsilon_{bd} + \epsilon_{ad}\epsilon_{bc}}{2}. \tag{A7}$$

To leading order in q, the eigenvalues are

$$\begin{aligned} K^+(\boldsymbol{q}) &= \xi_0\xi + \kappa_\parallel\xi^2 + \mathrm{O}(\xi^3) \\ K^-(\boldsymbol{q}) &= \kappa^\Gamma\xi^2 + \mathrm{O}(\xi^3) \end{aligned} \tag{A8}$$

so to leading order

$$\det(K(\boldsymbol{q})) = \kappa^\Gamma\xi_0\xi^3. \tag{A9}$$

This means that even in the presence of anisotropic mass tensors, there will be always at least one gapless mode at $\boldsymbol{q} = 0$. From our numerics, we find $\kappa^\Gamma_\parallel \approx -1.2253$ and $\kappa^\Gamma \approx +0.24506$.

b. $\boldsymbol{k} = \boldsymbol{M}_2 + \boldsymbol{q}$*:* we have an inverted (quadratic) band and a saddle point. The Little group is D_2 instead of the full D_6. To quadratic order in q, the elastic matrix reads

$$
\begin{aligned}
(K(\boldsymbol{M}_2+\boldsymbol{q}))_{ab} &= k^{M_2}_{ab} + \kappa^{M_2}_{abcd}\xi_c\xi_d \\
k^{M_2} &= \begin{pmatrix} 1.42218 & 0 \\ 0 & 11.451 \end{pmatrix} \\
\kappa^{M_2} &= \begin{pmatrix} 1.58534 & 0 & 0 & -0.251287 \\ 0 & -1.13164 & -1.13164 & 0 \\ 0 & -1.13164 & -1.13164 & 0 \\ -1.35849 & 0 & 0 & -0.629068 \end{pmatrix}
\end{aligned}
\tag{A10}
$$

where the indices of κ^{M_2} are ordered as $\{xx, xy, yx, yy\}$.

The eigenvalues to quadratic order in q are

$$
\begin{aligned}
K^-(M_2+\boldsymbol{q}) &= k^{M_2}_{11} + \kappa^{M_2}_{1;1}\xi_x^2 + \kappa^{M_2}_{1;4}\xi_y^2 \\
K^+(M_2+\boldsymbol{q}) &= k^{M_2}_{22} + \kappa^{M_2}_{4;1}\xi_x^2 + \kappa^{M_2}_{4;4}\xi_y^2
\end{aligned}
$$

note that K^- is a saddle point while K^+ is a band top.

c. $\boldsymbol{k} = \pm\boldsymbol{K} + \boldsymbol{q}$, the little group is C_3

$$
\begin{aligned}
K(\boldsymbol{K}+\boldsymbol{q}) &= k_K\tau^0 + \xi^2\left(\kappa_0^K\tau^0 + \kappa_1^K\tau^3_\theta\right) \\
\tau^3_\theta &= (\tau^3\cos(2\theta) + \tau^1\sin(2\theta)) \\
k_K &= 6.683 \\
\kappa_0^K &= -0.1794 \\
\kappa_1^K &= -0.2691 \\
\frac{\kappa_0^K}{\kappa_1^K} &= \frac{2}{3}, \text{ up to numerical precision}
\end{aligned}
\tag{A11}
$$

where θ is defined by $\boldsymbol{q} = (q\cos(\theta), q\sin(\theta))$. The dispersion relations are

$$
K^\pm(\boldsymbol{K}+\boldsymbol{q}) = k_K + \xi^2\left(\kappa_0^K \mp \kappa_1^K\right) \tag{A12}
$$

Appendix B: Details about the Variational Approach

Our trial states are of the form $|\Psi\rangle = |\Psi_v(\{\alpha_i\})\rangle \otimes |\Psi_{\text{ph}}(\{\alpha_i\})\rangle$ where $|\Psi_v(\{\alpha_i\})\rangle = \bigotimes_i |\alpha_i\rangle_i$ is a product state and $|\Psi_{\text{ph}}(\{\alpha_i\})\rangle$ is the ground state of the phonon Hamtiltonian with $\hat{W}_{ab;i} \to \langle\alpha_i|\hat{W}_{ab;i}|\alpha_i\rangle =: W_{ab;i}$. We consider periodic states $|\Psi_v(\{\alpha_i\})\rangle$ with possibly more than one (WC) lattice site per unit cell. We specify the larger unit cells by new lattice vectors $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ which must be integer combinations of the WC lattice vectors $\boldsymbol{a}_1 = a(1,0)$ and $\boldsymbol{a}_2 = \frac{a}{2}(1,\sqrt{3})$. In addition to the valley polarization patter, we also need to fix the overall orientation of the WC lattice, or equivalently, an orientation of the mass tensors with respect to the axes of the WC. We denote this angle by θ_{WC}.

We define the Fourier transform (FT) in terms of the WC lattice and interpret the larger unit cell as folding the BZ to a smaller one. Our FT is taken to be

$$
\begin{aligned}
Q_{\boldsymbol{k};a} &= \frac{1}{\sqrt{N_{\text{el}}}}\sum_i \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_i} Q_{i;a}, \\
P_{\boldsymbol{k};a} &= \frac{1}{\sqrt{N_{\text{el}}}}\sum_i \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_i} P_{i;a},
\end{aligned}
\tag{B1}
$$

and the commutation relations read

$$
\begin{aligned}
[Q_{\boldsymbol{k};a}, P^\dagger_{\boldsymbol{k}';b}] &= \frac{\mathrm{i}\delta_{ab}}{N_{\text{el}}}\sum_i \mathrm{e}^{\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}_i} \\
&= \mathrm{i}\delta_{ab}\sum_{\boldsymbol{G}} \delta_{\boldsymbol{k}+\boldsymbol{G},\boldsymbol{k}'}
\end{aligned}
\tag{B2}
$$

where the last sum is over reciprocal lattice vectors. Note that $(Q_{\boldsymbol{k}})^\dagger = (Q_{-\boldsymbol{k}})$ and $(P_{\boldsymbol{k}})^\dagger = (P_{-\boldsymbol{k}})$ and $\boldsymbol{k}$ is understood as a momentum on the first Brillouin zone of the WC (BZ).

Next we plug in the definitions of the FT into the Hamiltonian in Eq. 7. The terms that depend of the position operators reduce to

$$
\frac{H_{\hat{r}}}{N_{\text{el}}} = \int_{\boldsymbol{k}\in\text{BZ}} \frac{\mathrm{d}\boldsymbol{k}^2}{\nu'}\sum_{ab}\frac{1}{2}Q^\dagger_{\boldsymbol{k};a}K(\boldsymbol{k})_{ab}Q_{\boldsymbol{k};b}. \tag{B3}
$$

where $\nu' = \frac{4\pi^2}{\nu}$ is the area of BZ and $K(\boldsymbol{k})$ is the elastic matrix which corresponds to the Fourier transform of the dipole-dipole interactions. See App. A for the explicit expressions.

The $\hat{p}$ part of H is off-diagonal in momentum space. The larger unit cell spanned by $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ has an area that is an integer multiple of the original unit cell: $\Omega' = B\times\Omega$. When we take the FT of $W_{ab;i}$ we will generally find weight at B different momenta in the BZ that are closed under addition modulo reciprocal lattice vectors. We denote this set by $\mathcal{L} = \{\boldsymbol{L}_1, \ldots, \boldsymbol{L}_B\}$. Then

$$
\begin{aligned}
\frac{H_{\hat{p}}}{N_{\text{el}}} &= \int_{\boldsymbol{k}\in\text{BZ}} \frac{\mathrm{d}\boldsymbol{k}^2}{\nu'}\sum_{ab}\sum_{\boldsymbol{L}\in\mathcal{L}}\frac{\tilde{W}(\boldsymbol{L})_{ab}}{2B}P^\dagger_{\boldsymbol{k}+\boldsymbol{L};a}P_{\boldsymbol{k};b}; \\
&= \int_{\boldsymbol{k}\in\text{BZ}'} \frac{\mathrm{d}\boldsymbol{k}^2}{\nu'}\sum_{\alpha\beta}\mathcal{W}_{\alpha\beta}\frac{\mathcal{P}^\dagger_{\boldsymbol{k};\alpha}\mathcal{P}_{\boldsymbol{k};\beta}}{2};
\end{aligned}
\tag{B4}
$$

where $\alpha = (a,l)$ is a composite index with $a = x, y$ and $l = 1, \ldots, B$. The new variables are $\mathcal{Q}_{\boldsymbol{k};(a,l)} := Q_{\boldsymbol{k}+\boldsymbol{L}_l;a}$, $\mathcal{P}_{\boldsymbol{k};(a,l)} := P_{\boldsymbol{k}+\boldsymbol{L}_l;a}$, $\mathcal{W}_{(a,l)(b,l')} = \tilde{W}(\boldsymbol{L}_l - \boldsymbol{L}_{l'})_{ab}$ and

$\mathcal{K}_{(a,l)(b,l')}(\boldsymbol{k}) = K(\boldsymbol{k}+\boldsymbol{L}_l)_{ab}\delta_{ll'}$. Using the fact that $\mathcal{W}$ is positive-definite, we write $\mathcal{W} = \mathcal{V}^2$ with $\mathcal{V}$ positive-definite. Then we change variables to $\tilde{\mathcal{P}} = \mathcal{V}\mathcal{P}$ and $\tilde{\mathcal{Q}} = \frac{1}{\mathcal{V}}\mathcal{Q}$. The frequency matrix is $\Omega_{\boldsymbol{k}}^2 = \mathcal{V}\mathcal{K}(\boldsymbol{k})\mathcal{V}$ and the energy becomes

$$e_{\rm ph} = \frac{E_{\rm ph}}{N_{\rm el}} = \frac{1}{2}\int_{\boldsymbol{k}\in \mathrm{BZ}'} \frac{\mathrm{d}\boldsymbol{k}^2}{\nu'} \mathrm{Tr}\left[\sqrt{\Omega_{\boldsymbol{k}}^2}\right]. \tag{B5}$$

1. Overview of calculation

Our calculations of the variational states were done as follows

1. We chose a super-cell ($\boldsymbol{b}_1$ and $\boldsymbol{b}_2$) with a pattern of valley polarization.

2. We numerically evaluate the zero-point energy in Eq. 12, $e_{\rm ph}$, for several values of η and an evenly spaced grid of $\theta_{\rm WC}$. The numerical evaluation used an evenly spaced momentum grid of 24×24 momentum points that respect the PG of the triangular lattice to approximate the integral over the BZ.

3. For a given η, we determine the optimal $\theta^\star_{\rm WC}(\eta)$ that minimizes $e_{\rm ph}$. We always find that for a given valley-polarization pattern, the value $\theta^\star_{\rm WC}(\eta)$ does not depend on η.

4. For each η, we compare $e_{\rm ph}$ for the valley patterns (at their optimal $\theta_{\rm WC}$) and find the ones with the lowest $e_{\rm ph}$. We always find that the order of $e_{\rm ph}$ as a function of valley patterns is independent of $\eta > 0$.

5. To confirm the lowest energy trial state, we recalculate $e_{\rm ph}$ for the two configurations with lowest $e_{\rm ph}$ for a finer η grid.

Appendix C: Perturbative calculations around the isotropic electron gas

1. Correlation functions for isotropic case

We record some correlation functions for the isotropic case. We start with

$$\begin{aligned} h &= \sum_i \frac{P_i^2}{2} + \sum_{ij}\frac{Q_i K_{ij} Q_j}{2} \\ &= \sum_{\boldsymbol{k}} \frac{P_{\boldsymbol{k}}^\dagger P_{\boldsymbol{k}} + Q_{\boldsymbol{k}}^\dagger K(\boldsymbol{k}) Q_{\boldsymbol{k}}}{2} \end{aligned} \tag{C1}$$

We then introduce the harmonic modes as follows. For $K(\boldsymbol{k})v_{\boldsymbol{k};A} = \omega_{\boldsymbol{k}A}^2 v_{\boldsymbol{k};A}$ where A is the "band index". Then we introduce oscillator variables as follows

$$\begin{aligned} Q_{\boldsymbol{k}} &= \sum_A \sqrt{\frac{1}{2\omega_{\boldsymbol{k};A}}} v_{\boldsymbol{k};A}\left(\alpha^\dagger_{\boldsymbol{k};A} + \alpha_{-\boldsymbol{k};A}\right), \\ P_{\boldsymbol{k}} &= \sum_A \mathrm{i}\sqrt{\frac{\omega_{\boldsymbol{k};A}}{2}} v_{\boldsymbol{k};A}\left(\alpha^\dagger_{\boldsymbol{k};A} - \alpha_{-\boldsymbol{k};A}\right). \end{aligned} \tag{C2}$$

Then ($\langle O\rangle \equiv \mathrm{Tr}\left[O\mathrm{e}^{-\beta h}\right]/\mathrm{Tr}\left[\mathrm{e}^{-\beta h}\right]$)

$$\begin{aligned} \langle P_{i;a}(\tau)P_{j;b}\rangle &= \frac{1}{N_{\rm el}}\sum_{\boldsymbol{q}} \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_{ij}}\left\langle P^\dagger_{\boldsymbol{q};a}(\tau)P_{\boldsymbol{q};b}\right\rangle \\ \left\langle P^\dagger_{\boldsymbol{q};a}(\tau)P_{\boldsymbol{q};b}\right\rangle &= \sum_A \frac{\omega_{\boldsymbol{k};A}}{2} v_{\boldsymbol{k};A;a}v_{\boldsymbol{k};A;b}G_b(\omega_{\boldsymbol{k};A},\tau) \\ G_b(\omega,\tau) &= \frac{\cosh(\omega(\beta/2-\tau))}{\sinh\left(\frac{\beta\omega}{2}\right)}; 0\le\tau\le\beta. \end{aligned} \tag{C3}$$

2. Mapping to effective clock model

The aim of this section is to obtain the mapping used in Sec. III A to map the valley problem at small mass anisotropy ($\lambda \ll 1$) to an effective classical XY model with strong pinning potential. In this section, we are measuring energy in units of $E_\circ$.

We first write the effective Hamiltonian in Eq. 7 as $h = h_0 + \epsilon h_1$ with $\epsilon = \lambda/2$. h_0 is the Hamiltonian for the isotropic problem and h_1 is of the form

$$\begin{aligned} h_1 &= \sum_i P_i D_i P_i; \\ D_i &= \overline{W_i} - 1 = \cos(2\theta_i)\tau^3 + \sin(2\theta_i)\tau^1, \end{aligned}$$

here θ_i is the relative angle of the long axes of the Fermi surface of valley α_i with the horizontal axes. τ^i is i-th Pauli matrix.

Consider now the partition function of the valley pseudo-spins and the position degrees of freedom

$$Z = \sum_{\{\theta_i\}} Z[\{\theta_i\}];$$

where $Z[\{\theta_i\}]$ is the partition function of the position degrees of freedom in the presence of a valley pattern $\{\theta_i\}$. As ϵ is small, we can write

$$Z[\{\theta_i\}] = Z_0 \times \exp(-\beta\,\delta F\,[\{\theta_i\}])$$

where Z_0 is the partition function of the isotropic problem $Z_0 = \mathrm{Tr}\left[\mathrm{e}^{-\beta h_0}\right]$ and δF can be expanded in powers of ϵ so that $\delta F\,[\{\theta_i\}] = \sum_{n=0}^\infty \epsilon^n F_n[\{\theta_i\}]$. The leading correction with $\{\theta_i\}$ dependence is

$$-\beta F_2 = \frac{1}{2}\int_0^\beta \mathrm{d}\tau \int_0^\tau \mathrm{d}\tau' \left\langle h_1(\tau)h_1(\tau')\right\rangle. \tag{C4}$$

The $\{v_i\}$ dependence is hidden in the definition of h_1. We expand h_1 in terms of the momentum variables, evaluate the four point correlation functions and take the Fourier transform to arrive at

$$F_2 = -\frac{T}{2}\sum_{\boldsymbol{k},\alpha,\alpha'} \delta_{\boldsymbol{k},\boldsymbol{q}+\boldsymbol{q}'}\,\mathrm{Tr}\Big[\tilde{D}_{\boldsymbol{k}}\Pi_\alpha \tilde{D}^\dagger_{\boldsymbol{k}}\Pi_{\alpha'}\Big]\mathcal{G}_\beta(\omega_\alpha,\omega_{\alpha'}) \tag{C5}$$

here $\alpha = (\boldsymbol{q}, A)$ and A is the 'band' label of the phonon at momentum $\boldsymbol{q}$. $\Pi_\alpha = v_\alpha^\top v_\alpha$ is a projector where $v_\alpha = [\cos(\theta_\alpha), \sin(\theta_\alpha)]^\top$ is the eigenvector of $K(\boldsymbol{q}_\alpha)$ with eigenvalue ω_α^2. $\tilde{D}_{\boldsymbol{k}} = \frac{1}{N_{\rm el}}\sum_r \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_i} D_i$. $\mathcal{G}_\beta(\omega_\alpha,\omega_{\alpha'})$ is given by

$$\mathcal{G}_\beta(\omega_1,\omega_2) = \frac{\mathrm{sinhc}(\frac{\beta(\omega_1+\omega_2)}{2}) + \mathrm{sinhc}(\frac{\beta(\omega_1-\omega_2)}{2})}{\mathrm{sinhc}(\frac{\beta\omega_1}{2})\,\mathrm{sinhc}(\frac{\beta\omega_2}{2})},$$

where $\mathrm{sinhc}(x) = \sinh(x)/x$. If we write $\tilde{D}_{\boldsymbol{k}} = \Phi^{(1)}_{2;\boldsymbol{k}}\tau^3 + \Phi^{(1)}_{2;\boldsymbol{k}}\tau^1$, then $\Phi^{(1)}_{q;\boldsymbol{k}}$ and $\Phi^{(2)}_{q;\boldsymbol{k}}$ are the Fourier transform of $\cos(2\theta_i)$ and $\sin(2\theta_i)$, respectively. In terms of this variables, the correction to the free energy is

$$F_2 = -\frac{1}{4}\sum_{\boldsymbol{k}} \Phi^\dagger_{2;\boldsymbol{k}}\underline{K}(\boldsymbol{k},\beta)\Phi_{2;\boldsymbol{k}} \tag{C6}$$

where $\Phi_{2;\boldsymbol{k}} = [\Phi^{(1)}_{2;\boldsymbol{k}}, \Phi^{(2)}_{2;\boldsymbol{k}}]^\top$. The kernel can be expanded as $\underline{K} = K_1\tau^0 + K_2\tau^3 + K_3\tau^1$ with

$$\begin{aligned}
K_1(\boldsymbol{k},\beta) &= \frac{T}{N_{\rm el}}\sum_{\boldsymbol{k},\alpha,\alpha'} \delta_{\boldsymbol{k},\boldsymbol{q}+\boldsymbol{q}'}\mathcal{G}_\beta(\omega_\alpha,\omega_{\alpha'});\\
K_2(\boldsymbol{k},\beta) &= \frac{T}{N_{\rm el}}\sum_{\boldsymbol{k},\alpha,\alpha'} \delta_{\boldsymbol{k},\boldsymbol{q}+\boldsymbol{q}'}\mathcal{G}_\beta(\omega_\alpha,\omega_{\alpha'})\cos(2(\theta_\alpha+\theta_{\alpha'}));\\
K_3(\boldsymbol{k},\beta) &= \frac{T}{N_{\rm el}}\sum_{\boldsymbol{k},\alpha,\alpha'} \delta_{\boldsymbol{k},\boldsymbol{q}+\boldsymbol{q}'}\mathcal{G}_\beta(\omega_\alpha,\omega_{\alpha'})\sin(2(\theta_\alpha+\theta_{\alpha'})).
\end{aligned} \tag{C7}$$

The effective energy (at zero temperature) can be derived by using the relation $E = \frac{\partial\beta F}{\partial\beta}\big|_{T=0}$. We use the identity $\frac{\partial}{\partial\beta}\mathcal{G}_\beta(\omega_1,\omega_2)\big|_{T=0} = \frac{\omega_1\omega_2}{\omega_1+\omega_2}$ to write the leading contribution to the energy as

$$E_2 = -\frac{1}{4}\sum_{\boldsymbol{k}} \Phi^\dagger_{2;\boldsymbol{k}}\underline{K}(\boldsymbol{k})\Phi_{2;\boldsymbol{k}} \tag{C8}$$

where

$$\begin{aligned}
\underline{K} &= K_1 + K_2\tau^3 + K_3\tau^1;\\
K_1(\boldsymbol{k}) &= \frac{1}{N_{\rm el}}\sum_{\boldsymbol{k},\alpha,\alpha'} \delta_{\boldsymbol{k},\boldsymbol{q}+\boldsymbol{q}'}\frac{\omega_\alpha\omega_{\alpha'}}{\omega_\alpha+\omega_{\alpha'}};\\
K_2(\boldsymbol{k}) &= \frac{1}{N_{\rm el}}\sum_{\boldsymbol{k},\alpha,\alpha'} \delta_{\boldsymbol{k},\boldsymbol{q}+\boldsymbol{q}'}\frac{\omega_\alpha\omega_{\alpha'}}{\omega_\alpha+\omega_{\alpha'}}\cos(2(\theta_\alpha+\theta_{\alpha'}));\\
K_3(\boldsymbol{k}) &= \frac{1}{N_{\rm el}}\sum_{\boldsymbol{k},\alpha,\alpha'} \delta_{\boldsymbol{k},\boldsymbol{q}+\boldsymbol{q}'}\frac{\omega_\alpha\omega_{\alpha'}}{\omega_\alpha+\omega_{\alpha'}}\sin(2(\theta_\alpha+\theta_{\alpha'}));\\
\tilde{J}(\boldsymbol{k}) &\equiv \frac{1}{16}\underline{K}(\boldsymbol{k}).
\end{aligned} \tag{C9}$$

Eq. 11 in the main text is obtained by diving Eq. C8 by the number of electrons ($N_{\rm el}$).

We numerically evaluate $\underline{K(\boldsymbol{k})}$ over a 12×12 grid in momentum space. Fig. 10 shows each component as well as the largest eigenvalue ($K_1 + |K_2 + \mathrm{i}K_3|$) at each $\boldsymbol{k}$.

3. Correction to the Energy from inversion pairs

We want to calculate the effect of trigonal warping on the degeneracy between two valleys related by inversion. As mentioned in Sec. IV, the effective Hamiltonian (Eq. 22) is $h_{\rm eff} = h_0 + \epsilon_3 h_3$ where h_0 is the phonon Hamil-

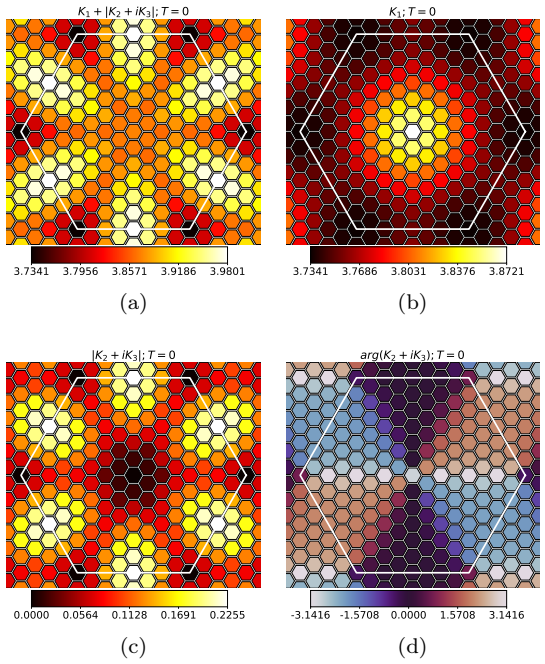


FIG. 10. Plots of the components of the kernel in Eq. C9 used to evaluate the effective energy of the system.

tonian in the absence of trigonal warping, $\epsilon_3 \ll 1$ and

$$h_3 = \sum_i \mathrm{e}^{+3\mathrm{i}\phi_i}(P_{i;x}-\mathrm{i}P_{i;y})^3+\mathrm{e}^{-3\mathrm{i}\phi_i}(P_{i;x}+\mathrm{i}P_{i;y})^3 \quad \text{(C10)}$$

We want to consider a perturbative expansion of the energy (per electron) with trigonal warping $e^{\triangle}_{\mathrm{ph}} = \sum_{n=0}^{\infty} \epsilon_3^n e_n^{\triangle}$. $e_0^{\triangle}$ is the energy of the system in the absence of trigonal warping. We calculate $e_n^{\triangle}$ using many-body perturbation theory.

First, we rewrite Eq. C10 in momentum space

$$h_3 = \frac{1}{N_{\mathrm{el}}^{1/2}} \sum_{\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3} \left[\Phi^+_{3;\boldsymbol{q}} \prod_{l=1}^{3} (\zeta^* \cdot P^\dagger_{\boldsymbol{k}_l}) + h.c. \right] \quad \text{(C11)}$$

here $\zeta = [1, \mathrm{i}]$ and $\Phi^+_{3;\boldsymbol{q}}$ is the Fourier transform of $\mathrm{e}^{3\mathrm{i}\phi_i}$ with $\boldsymbol{q} = \boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3$.

Let's start with $e_1^{\triangle}$. This corresponds to evaluating matrix elements of h_3/N_{el} between unperturbed ground states. These states correspond to states with different center of mass momenta ($P_{CM} = \frac{1}{N_{\mathrm{el}}} \sum_i P_i = \frac{1}{\sqrt{N_{\mathrm{el}}}} P_{\mathbf{0}}$) and the respective ground state of the oscillators at the other crystal momenta. The only terms in Eq. C11 that contribute are the ones with an odd number of $\boldsymbol{k}_l = \mathbf{0}$ and $\boldsymbol{q} = 0$. If we write $\Phi^+_{3;\mathbf{0}} = |\Phi_{3;\mathbf{0}}|\mathrm{e}^{\mathrm{i}\Theta_{3;\mathbf{0}}}$ and $P_{CM} = |P_{CM}|[\cos(\theta_{CM}), \sin(\theta_{CM})]^\top$, then

$$\begin{aligned} e_1^{\triangle} &= 2|\Phi_{3;\mathbf{0}}||P_{CM}|^3 \cos(3\theta_{CM} - \Theta_{3;\mathbf{0}}) \\ &\quad + 3|\Phi_{3;\mathbf{0}}||P_{CM}| \left[\mathrm{e}^{\mathrm{i}(\Theta_{3;\mathbf{0}} - \theta_{CM})} \mathcal{A}^* + c.c. \right] \\ \mathcal{A} &= \frac{1}{2N_{\mathrm{el}}} \sum_{\boldsymbol{k}_1,\boldsymbol{k}_2} \left\langle (\zeta \cdot P_{\boldsymbol{k}_1})(\zeta \cdot P_{\boldsymbol{k}_2}) \right\rangle \end{aligned} \quad \text{(C12)}$$

Let's assume that the ground state of h_0 preserves the C_3 symmetry of the WC so that $\mathcal{A}$ vanishes (to see this, under $C_3 : \zeta \cdot P_{\boldsymbol{k}} \to \mathrm{e}^{\frac{2\pi\mathrm{i}}{3}} \zeta \cdot P_{C_3\boldsymbol{k}}$. Then, after relabeling the moment, we see that $\mathcal{A} \to \mathcal{A}\mathrm{e}^{\frac{4\pi\mathrm{i}}{3}}$ which implies that $\mathcal{A} = 0$.). We need to add a $\frac{u}{4}|P|^4_{CM}$ to the energy ($u > 0$) to stabilize the system. The effective potential is then $e_0^{\triangle} + \epsilon_3 e_1^{\triangle} + \frac{u}{4}|P|^4_{CM}$ is minimized at $P_{CM} = 0$ as long as ϵ_3 is small because there is no linear term in P_{CM}.

Next, we evaluate $e_2^{\triangle}$ using perturbation theory and ignore the terms involving P_{CM} as these terms are higher order in ϵ_3 than the ones already present in $e_0^{\triangle} + \epsilon_3 e_1^{\triangle} + \frac{u}{4}|P|^4_{CM}$:

$$e_2^{\triangle} = -\frac{3!}{2^3 N_{\mathrm{el}}^2} \sum_{\boldsymbol{k}\alpha_1,\alpha_2,\alpha_3} \delta_{\boldsymbol{k},\sum_\alpha \boldsymbol{q}_\alpha} \frac{\omega_{\alpha_1}\omega_{\alpha_2}\omega_{\alpha_3}}{\omega_{\alpha_1}+\omega_{\alpha_2}+\omega_{\alpha_3}} \left| \Phi^+_{3;\boldsymbol{k}} \mathrm{e}^{-\mathrm{i}\sum_\alpha \theta_\alpha} + \Phi^-_{3;\boldsymbol{k}} \mathrm{e}^{+\mathrm{i}\sum_\alpha \theta_\alpha} \right|^2 \quad \text{(C13)}$$

As in Sec. C 2, we can expand the absolute value and write the final answer as

$$e_2^{\triangle} = -\frac{3!}{2^4 N_{\mathrm{el}}} \sum_{\boldsymbol{k}} \Phi^\dagger_{3;\boldsymbol{k}} \underline{K}^{(3)}(\boldsymbol{k}) \Phi_{3;\boldsymbol{k}} \quad \text{(C14)}$$

with $\Phi_{3;\boldsymbol{k}}$ the Fourier transform of $[\cos(3\theta_i), \sin(3\theta_i)]^\top$ and the components of kernel is

$$\begin{aligned} \underline{K}^{(3)} &= K_1^{(3)}\tau^0 + K_2^{(3)}\tau^3 + K_3^{(3)}\tau^1 \\ K_1^{(3)}(\boldsymbol{k}) &= \frac{1}{N_{\mathrm{el}}^2} \sum_\alpha \delta_{\boldsymbol{k},\sum_\alpha \boldsymbol{q}_\alpha} \frac{\omega_{\alpha_1}\omega_{\alpha_2}\omega_{\alpha_3}}{\omega_{\alpha_1}+\omega_{\alpha_2}+\omega_{\alpha_3}}; \\ K_2^{(3)}(\boldsymbol{k}) &= \frac{1}{N_{\mathrm{el}}^2} \sum_\alpha \delta_{\boldsymbol{k},\sum_\alpha \boldsymbol{q}_\alpha} \frac{\omega_{\alpha_1}\omega_{\alpha_2}\omega_{\alpha_3}}{\omega_{\alpha_1}+\omega_{\alpha_2}+\omega_{\alpha_3}} \cos\left(2\sum_\alpha \theta_\sigma\right); \\ K_3^{(3)}(\boldsymbol{k}) &= \frac{1}{N_{\mathrm{el}}^2} \sum_\alpha \delta_{\boldsymbol{k},\sum_\alpha \boldsymbol{q}_\alpha} \frac{\omega_{\alpha_1}\omega_{\alpha_2}\omega_{\alpha_3}}{\omega_{\alpha_1}+\omega_{\alpha_2}+\omega_{\alpha_3}} \sin\left(2\sum_\alpha \theta_\sigma\right); \\ \tilde{J}_3(\boldsymbol{k}) &\equiv \frac{3!}{2^4} \underline{K}^{(3)}(\boldsymbol{k}). \end{aligned} \quad \text{(C15)}$$

We numerically evaluate the kernel $\underline{K}^{(3)}(\boldsymbol{k})$ in a 12×12 momentum grid. The results are shown in Fig. 11. The largest eigenvalue is at $\boldsymbol{k} = 0$ with eigenvector $[0, 1]^\top$. This corresponds to a valley-ferromagnetic state with $\mathrm{e}^{3\mathrm{i}\phi_i} = \pm\mathrm{i}$.

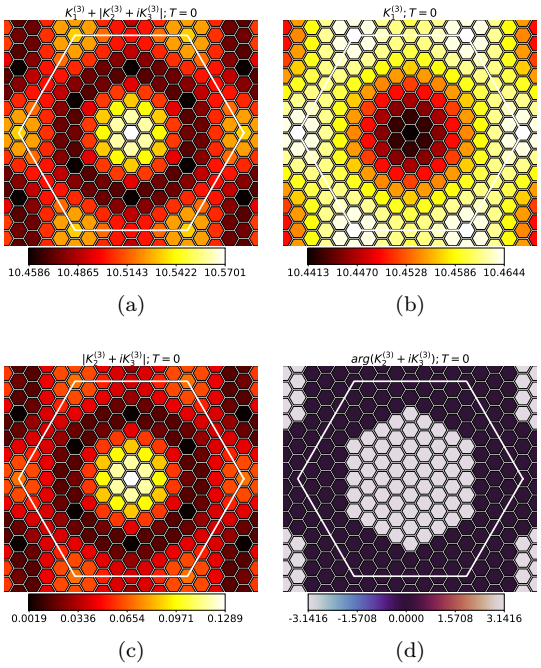


FIG. 11. Plots of the Kernels in Eq. C15 at $T = 0$.

In the absence of C_3 symmetry of the ground state of h_0, $\mathcal{A}$ in general would be non-zero. The potential for P_{CM} is now minimized by a $P^*_{CM} \propto \epsilon_3 |\Phi_{3;\mathbf{0}}|$ and $\theta_{CM} = \Theta_{3;\mathbf{0}} - \arg(\mathcal{A})$. The value of the potential at this minima is negative and proportional to $\epsilon_3^2 |\Phi_{3;\mathbf{0}}|^2$. This favours having a large $|\Phi_{3;\mathbf{0}}|$, i.e. a valley ferromagnetic order as does $e_2^{\triangle}$ when there is C_3 rotation symmetry. Therefore, we hope that even in the absence of C_3 symmetry, there could be a valley ferromagnet that breaks inversion symmetry and time-reversal symmetry. A main new feature in the absence of C_3, is that average crystal momentum (P_{CM}) is shifted, i.e. the band minima is displaced from its position at small r_s.

4. Berry curvature

Consider the Hamiltonian in Eq. 27 of the main text. We treat h_1 and h_2 as perturbations to h_0. The leading order correction to the energy arises at first order in $\bar{\mathcal{B}}_0$ and corresponds to $\bar{\mathcal{B}}_0 \langle h_1 \rangle$, where $\langle h_1 \rangle$ is the expectation value of h_1 on the ground state of h_0. This can be computed as

$$\begin{aligned}\langle h_1 \rangle &= -\sum_{ij} \frac{\langle \mu_i (\epsilon P)_i K_{ij} Q_j + \mu_j Q_j K_{ij} (\epsilon P)_i \rangle}{2} \\ &= \frac{\mathrm{i}}{4} \sum_{ij;aba'} (\mu_i \epsilon_{aa'} \delta_{a'b} \delta_{ij} (K_{ij})_{ab} - \mu_j \delta_{ba'} \delta_{ij} (K_{ij})_{ba} \epsilon_{aa'}) \\ &= \frac{\mathrm{i}}{4} \sum_{i;a} (\mu_i \epsilon_{ab} (K_{ii})_{ab} - \mu_j \epsilon_{ab} (K_{ii})_{ba}) = 0.\end{aligned} \tag{C16}$$

5. Moiré corrections

If we modify the dispersion relation to $\omega \to \sqrt{\kappa_{\text{moiré}} + \omega^2}$ and assume that $\kappa_{\text{moiré}}$ is large compared to the phonon bandwidth, we can perform a perturbative expansion in $\kappa_{\text{moiré}}$ for the kernel. In particular,

$$\begin{aligned}\frac{\prod_{j=1,2,3} \sqrt{1+\epsilon x_j}}{\sum_{j=1,2,3} \sqrt{1+\epsilon x_j}} &= \frac{1}{3} + \frac{\epsilon}{9}(x_1 + x_2 + x_3) - \frac{5\epsilon^2}{108}\left(x_1^2 - x_2 x_1 - x_3 x_1 + x_2^2 + x_3^2 - x_2 x_3\right) \\ &+ \frac{\left(34(x_1^3 + x_2^3 + x_3^3) - 21(x_2 x_1^2 + x_3 x_1^2 + x_2^2 x_1 + x_3^2 x_1 + x_2 x_3^2 + x_2^2 x_3) + 24 x_1 x_2 x_3\right)\epsilon^3}{1296} + \mathrm{O}(\epsilon^4)\end{aligned} \tag{C17}$$

Then we want to replace $\epsilon = 1/\kappa_{\text{moiré}}$ and identify the x's with the ω^2's and evaluate the averages. First, note that when the energy does not appear the average of $\cos(\cdots)$ and $\sin(\cdots)$ is zero because the two eigenvectors at the same $\boldsymbol{q}$-point have $\theta_2 - \theta_1 = \pi/2$. Secondly, the only terms that depend on $\boldsymbol{k}$ is the one that corresponds to that is the average of something proportional to $\omega_\alpha^2 \omega_{\alpha'}^2 \omega_{\alpha''}^2$. In all the others, we can use the $\delta_{\sum \boldsymbol{q}, \boldsymbol{k}}$ to get rid of the $\boldsymbol{k}$ dependence. So we can write

$\underline{K}^{(3)}(\boldsymbol{k}) = K' \mathbb{I} + \frac{1}{\kappa^2_{\text{moiré}}} \underline{\delta K}^{(3)}(\boldsymbol{k}) + \mathrm{O}(\kappa^{-3}_{\text{moiré}})$ with

$$\begin{aligned}\delta K_1^{(3)}(\boldsymbol{k}) &= \frac{1}{N_{\text{el}}^2} \sum_\alpha \delta_{\boldsymbol{k}, \sum_\alpha \boldsymbol{q}_\alpha} \frac{\omega_\alpha^2 \omega_{\alpha'}^2 \omega_{\alpha''}^2}{54} \\ \delta K_2^{(3)}(\boldsymbol{k}) &= \frac{1}{N_{\text{el}}^2} \sum_\alpha \delta_{\boldsymbol{k}, \sum_\alpha \boldsymbol{q}_\alpha} \frac{\omega_\alpha^2 \omega_{\alpha'}^2 \omega_{\alpha''}^2}{54} \cos\left(2 \sum_\alpha \theta_\sigma\right) \\ \delta K_3^{(3)}(\boldsymbol{k}) &= \frac{1}{N_{\text{el}}^2} \sum_\alpha \delta_{\boldsymbol{k}, \sum_\alpha \boldsymbol{q}_\alpha} \frac{\omega_\alpha^2 \omega_{\alpha'}^2 \omega_{\alpha''}^2}{54} \sin\left(2 \sum_\alpha \theta_\sigma\right)\end{aligned} \tag{C18}$$

we again evaluate $\underline{\delta K}^{(3)}(\boldsymbol{k})$ numerically and find similar behaviour to Fig. 11. In particular, the eigenvector with the largest eigenvalue is the same as for $\underline{K}^{(3)}(\boldsymbol{k})$.

6. Extension to 3D

The above calculation easily generalizes to 3d. In particular, we consider electrons with inverse mass tensor W. Then define $m^\star = 3\left(\mathrm{Tr}[W]\right)^{-1}$ and $W_i m^\star = \mathbb{I} + \lambda Q_i$ with Q_i a traceless symmetric matrix.

We obtain the same expression for the free energy and energy as in Sec. C 2 with the following modifications. $A \in \{1,2,3\}$, the sum is over the 3 dimensional BZ.

$$\delta E = -\frac{\lambda^2}{32}\sum_{\boldsymbol{k}} \Phi^\dagger_{\boldsymbol{k}} \underline{Y}(\boldsymbol{k}) \Phi_{\boldsymbol{k}} \tag{C19}$$

where $\Phi_{\boldsymbol{k}}$ is a five component vector with components $\mathrm{Tr}\left[\tilde{D}_{\boldsymbol{k}} S^a\right]$, where $\{S^a, a = 1,2,3,4,5\}$ is a orthonormal basis for 3 by 3 symmetric traceless matrices. The kernel is

$$\underline{Y}(\boldsymbol{k})_{ab} = \frac{1}{N_{\mathrm{el}}}\sum_{\boldsymbol{k},\alpha,\alpha'} \delta_{\boldsymbol{k},\boldsymbol{q}+\boldsymbol{q}'} \frac{\omega_\alpha \omega_{\alpha'}}{\omega_\alpha + \omega_{\alpha'}} \mathrm{Tr}\left[\Pi_\alpha S^a \Pi_{\alpha'} (S^b)^\dagger\right] \tag{C20}$$

In here ω_α and Π_α are the eigenvalues and projectors of the matrix $K(\boldsymbol{k})$. We used the expressions of Refs. [30; 31] to evaluate the kernel. When we evaluate the sum, we should not use the value of ω_α when $\alpha = (\boldsymbol{0}, A)$ but instead use an average around $\boldsymbol{q} = \boldsymbol{0}$ because $K(\boldsymbol{q})$ is discontinuous at $\boldsymbol{q} = 0$ due to the long-range nature of Coulomb interactions. We approximate $K(\boldsymbol{q} \to \boldsymbol{0})_{ab} = K_0 \frac{q_a q_b}{q^2}$ so that only the longitudinal model contributes. Then taking the average over all directions result in a zero contribution. Therefore, we simply omit the terms with $\boldsymbol{q} = \boldsymbol{0}$ or $\boldsymbol{q}' = \boldsymbol{0}$.

We approximate the kernel by an average over a $12 \times 12 \times 12$ grid of the first BZ of the FCC crystal. We find that the largest eigenvalue of $\underline{Y}(\boldsymbol{k})$ is at $\boldsymbol{k} = \boldsymbol{N} = (\pi, \pi, 0)$ (and 5 other symmetry related momenta) with eigenvalue, eigenvector pair

$$\Lambda \approx 0.4106, V = \frac{1}{2}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix}.$$

At Γ, the diagonal matrices have eigenvalue 0.3773 while the off-diagonal ones have 0.4054.

The kernel in the main text is $\tilde{J}(\boldsymbol{k}) = \frac{1}{32}\underline{Y}(\boldsymbol{k})$.

Appendix D: Dispersion relations

In this section we give the details of calculation of small wavelength dispersion relation of the phonons. We focus on the frequency matrix around $\mathbf{q} = 0$ for the case of period 2 stripes.

$$\Omega^2_{\mathbf{q}} = \mathcal{V} \cdot \mathcal{K}(\mathbf{q}) \cdot \mathcal{V} \tag{D1}$$

$$\mathcal{V}^2 = \mathcal{W} = \frac{1}{2}\begin{pmatrix} W_A + W_B & W_A - W_B \\ W_A - W_B & W_A + W_B \end{pmatrix} \tag{D2}$$

$$\mathcal{K}(\boldsymbol{q}) = \begin{pmatrix} K(\boldsymbol{q}) & 0 \\ 0 & K(\boldsymbol{q}+\boldsymbol{M}) \end{pmatrix} \tag{D3}$$

At $\boldsymbol{q} = 0$, there are two zero eigenvalues. To leading order in q, $K(\boldsymbol{q}) = q_0 q \hat{q}_a \hat{q}_b$ and $K(\boldsymbol{M}+\boldsymbol{q}) = K(\boldsymbol{M})$. We then expect that the order q perturbation will lift the degeneracy. First order degenerate perturbation theory tell us to compute the matrix elements of the perturbation in the basis of the degenerate states. In this case, we can take the basis as $\mathbf{e}_1 = \mathcal{V}^{-1}(1,0,0,0)$ and $\mathbf{e}_2 = \mathcal{V}^{-1}(0,1,0,0)$. Define $\tilde{M}$ as

$$\tilde{M}_{ij} = (\mathcal{W}^{-1})_{ij}; \quad i,j \in \{1,2\}.$$

so that $\mathbf{e}_i \cdot \mathbf{e}_j = \tilde{M}_{ij}$. Let's assume this matrix is positive definite. Then we need to find the eigenvalues of

$$\tilde{\Omega}^2_{\boldsymbol{q}} = \frac{q_0}{q}\frac{1}{\sqrt{\tilde{M}}} \cdot \begin{pmatrix} q_x q_x & q_x q_y \\ q_y q_x & q_y q_y \end{pmatrix} \cdot \frac{1}{\sqrt{\tilde{M}}}.$$

but as $\det\left(\tilde{\Omega}^2_{\boldsymbol{q}}\right) = 0$, there is a zero eigenvalue and the other one is given by the trace:

$$\begin{aligned} \tilde{\omega}^2_{\boldsymbol{q},L} &= \frac{q_0}{q} q \cdot \frac{1}{\tilde{M}} \cdot q \\ \tilde{\omega}^2_{\boldsymbol{q},T} &= 0. \end{aligned} \tag{D4}$$

Also note that because $\det\left(\Omega^2_{\boldsymbol{q}}\right) \propto q^3$ for q small, we can find the leading order behaviour of $\omega^2_{\boldsymbol{q},T}$ to be q^2. To see this note that as a function of $\boldsymbol{q}$, the frequency of the optical phonons is

$$\omega^2_{\boldsymbol{q},op,i} = \omega^2_{0,op,i} + \mathcal{O}(q^\alpha) \quad ; i \in \{1,2\}, \omega^2_{0,op,i} \neq 0.$$

Then we calculate the determinant of $\Omega^2_{\boldsymbol{q}}$ in two different ways

$$\omega^2_{\boldsymbol{q},L}\omega^2_{\boldsymbol{q},T}\omega^2_{\boldsymbol{q},op,1}\omega^2_{\boldsymbol{q},op,2} = \det\left(\Omega^2_{\boldsymbol{q}}\right) = \det(\mathcal{K}(\boldsymbol{q}))\det(\mathcal{W}),$$

then expanding to leading order in q

$$\omega^2_{\boldsymbol{q},T} = \frac{\det(K(\boldsymbol{q}))}{\omega^2_{\boldsymbol{q},L}} \frac{\det(K(\boldsymbol{M}))\det(\mathcal{W})}{\omega^2_{0,op,1}\omega^2_{0,op,2}}$$

From here we can see that if $\tilde{M}$ is isotropic, the dispersion relation of both acoustic phonons is isotropic for $q \ll 1$.

For a symmetric matrix

$$J = \begin{pmatrix} A & B \\ B & A \end{pmatrix} \Rightarrow J^{-1} = \begin{pmatrix} (A - BA^{-1}B)^{-1} & \star \\ \star & (A - BA^{-1}B)^{-1} \end{pmatrix}$$

assuming A and B are invertible matrices.

If we identify $A = \mathcal{W}_\Gamma = \frac{1}{2}(W_A + W_B)$ and $B = \mathcal{W}_M = \frac{1}{2}(W_A - W_B)$ we have

$$\tilde{M}^{-1} = W_\Gamma - W_M W_\Gamma^{-1} W_M.$$

In particular, for the square lattice $W_\Gamma = \tau^0$ and $W_M = \pm\lambda\tau^3$ so $\tilde{M}^{-1} = (1-\lambda^2)\tau^0$. We can calculate the other frequencies by noting that $\mathcal{W} = \mathbb{1} + \lambda\tau^3\mu^1 = \frac{2\eta^{-\tau^3\mu^1}}{\eta+1/\eta}$ and $\mathcal{K}(0) = \frac{(1+\mu^3)}{2}(A + B\tau^3)$. Then

$$\Omega_0^2 = \frac{2\eta}{1+\eta^2}\frac{(\eta^{\tau^3\mu^1} + \mu^3)}{2}(A + B\tau^3) \Rightarrow \omega_{0;i}^2 = k_{ii}^{M_2},$$

in other words the optical frequencies of the phonons are independent of η.

Appendix E: Optical conductivity

In this Appendix, we simplify the expression for the optical conductivity that is used in Sec. III D 2 and Sec. III D 3. First, note that

$$\begin{aligned} J^a(\boldsymbol{q})/e &= \sum_{i;b}(W_i)_{ab}\cdot p_i^b \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\mathbf{R}_i} && \text{(E1)}\\ &= \sum_{i;b}\sum_{\boldsymbol{L}\in\mathcal{L}}\tilde{W}(-\boldsymbol{L})_{ab}\mathrm{e}^{\mathrm{i}(\boldsymbol{L}+\boldsymbol{q})\cdot\mathbf{R}_i}p_i^b && \text{(E2)}\\ &= \sqrt{N_{\mathrm{el}}}\sum_{\boldsymbol{L}\in\mathcal{L};b}\tilde{W}(-\boldsymbol{L})_{ab}P_{\boldsymbol{L}+\boldsymbol{q};b}\sqrt{E_\circ m^\star} && \text{(E3)}\\ &= \sqrt{\frac{E_\circ}{m^\star}}\sqrt{N_{\mathrm{el}}}\,[\mathcal{W}\mathcal{P}_{\boldsymbol{q}}]_{(a,0)} && \text{(E4)} \end{aligned}$$

Then the calculation reduces to calculate the matrix elements $\langle m|\mathcal{P}_{\boldsymbol{k}}|0\rangle$ between the ground state $|0\rangle$ and the excited states $|m\rangle$. At every momenta $\boldsymbol{k} \in BZ'$ the Hamiltonian in units of $E_\circ$ is

$$\mathrm{H}_{\boldsymbol{k}}/E_\circ = \frac{1}{2}\mathcal{P}_{\boldsymbol{k}}^\dagger\mathcal{W}\mathcal{P}_{\boldsymbol{k}} + \frac{1}{2}\mathcal{Q}_{\boldsymbol{k}}^\dagger\mathcal{K}(\boldsymbol{k})\mathcal{Q}_{\boldsymbol{k}}.$$

Next we diagonalize the frequency matrix as $\Omega_{\boldsymbol{k}}^2 = \mathcal{O}_{\boldsymbol{k}}^\dagger D_{\boldsymbol{k}}^2\mathcal{O}_{\boldsymbol{k}}$ with $\mathcal{O}_{\boldsymbol{k}}$ orthogonal and $D_{\boldsymbol{k}}$ diagonal. Then, we make the canonical change of variables $\acute{\mathcal{P}}_{\boldsymbol{k}} = \mathcal{O}_{\boldsymbol{k}}\mathcal{V}\mathcal{P}_{\boldsymbol{k}}$ and $\acute{\mathcal{Q}}_{\boldsymbol{k}} = \mathcal{O}_{\boldsymbol{k}}(\mathcal{V})^{-1}\mathcal{Q}_{\boldsymbol{k}}$ to rewrite

$$\mathrm{H}_{\boldsymbol{k}}/E_\circ = \frac{1}{2}\acute{\mathcal{P}}_{\boldsymbol{k}}^\dagger\acute{\mathcal{P}}_{\boldsymbol{k}} + \frac{1}{2}\acute{\mathcal{Q}}_{\boldsymbol{k}}^\dagger D_{\boldsymbol{k}}^2\acute{\mathcal{Q}}_{\boldsymbol{k}}.$$

Thus,

$$\langle 0|(\acute{\mathcal{P}}_{\boldsymbol{k}})_\beta|n\rangle = \sum_\alpha \delta_{\omega_n, E_\circ D_{\boldsymbol{k};\alpha\alpha}}\mathrm{i}\sqrt{\frac{D_{\boldsymbol{k};\alpha\alpha}}{2}}\delta_{\alpha\beta}$$

which further implies

$$\langle 0|(\acute{\mathcal{P}}_{\boldsymbol{k}})_\beta|n\rangle\ \overline{\langle 0|(\acute{\mathcal{P}}_{\boldsymbol{k}})_\gamma|n\rangle} = \sum_\alpha \delta_{\omega_n, E_\circ D_{\boldsymbol{k};\alpha\alpha}}\frac{D_{\boldsymbol{k};\alpha\alpha}}{2}\delta_{\alpha\beta}\delta_{\alpha\gamma}.$$

Then we multiply by $\delta(\omega-\omega_n)/\omega$ and sum over n

$$\begin{aligned} \sum_n \frac{\delta(\omega-\omega_n)}{\omega}\langle 0|(\acute{\mathcal{P}}_{\boldsymbol{k}})_\beta|n\rangle\ \overline{\langle 0|(\acute{\mathcal{P}}_{\boldsymbol{k}})_\gamma|n\rangle} &= \sum_{\alpha,n}\frac{\delta(\omega-\omega_n)}{\omega}\delta_{\omega_n, E_\circ D_{\boldsymbol{k};\alpha\alpha}}\frac{D_{\boldsymbol{k};\alpha\alpha}}{2}\delta_{\alpha\beta}\delta_{\alpha\gamma}\\ &= \sum_{\alpha,n}\frac{\delta(\omega-E_\circ D_{\boldsymbol{k};\alpha\alpha})}{E_\circ D_{\boldsymbol{k};\alpha\alpha}}\delta_{\omega_n, E_\circ D_{\boldsymbol{k};\alpha\alpha}}\frac{D_{\boldsymbol{k};\alpha\alpha}}{2}\delta_{\alpha\beta}\delta_{\alpha\gamma}\\ &= \frac{1}{2E_\circ^2}\sum_\alpha \delta(\omega/E_\circ - D_{\boldsymbol{k},\alpha\alpha})\delta_{\beta\alpha}\delta_{\gamma\alpha}\\ &= \frac{1}{2E_\circ^2}\delta(\omega/E_\circ - D_{\boldsymbol{k}})_{\beta\gamma}, \end{aligned} \tag{E5}$$

where we defined $\delta(\omega - X) \equiv \sum_x \delta(\omega-x)\Pi_x$ where X is a Hermitian matrix, x are the eigenvalues of X and Π_x are projections on the subspace with eigenvalue x.

The last step is to conjugate the above equality by $(\mathcal{O}\mathcal{V})^{-1}$ to go from $\acute{\mathcal{P}}$ to $\mathcal{P}$:

$$\begin{aligned} \sum_n \frac{\delta(\omega-\omega_n)}{\omega}\langle 0|(\mathcal{P}_{\boldsymbol{k}})_\beta|n\rangle\ \overline{\langle 0|(\mathcal{P}_{\boldsymbol{k}})_\gamma|n\rangle} &= \frac{1}{2E_\circ^2}\left[\frac{1}{\mathcal{V}}\mathcal{O}_{\boldsymbol{k}}^\dagger\delta(\omega/E_\circ - D_{\boldsymbol{k}})\mathcal{O}_{\boldsymbol{k}}\frac{1}{\mathcal{V}}\right]_{\beta\gamma}\\ &= \frac{1}{2E_\circ^2}\left[\frac{1}{\mathcal{V}}\delta(\omega/E_\circ - \Omega_{\boldsymbol{k}})\frac{1}{\mathcal{V}}\right]_{\beta\gamma} \end{aligned} \tag{E6}$$

Finally,

$$\begin{aligned}\sigma_{ab}(\omega,\boldsymbol{q}) &= \frac{4\pi}{A}\sum_n \frac{\delta(\omega-\omega_n)}{\omega}\,\langle 0|J_a(\boldsymbol{q})|n\rangle\,\overline{\langle 0|J_b(\boldsymbol{q})|n\rangle} \\ &= \frac{4\pi e^2}{A}\frac{N_{\rm el}}{2m^\star E_\circ}\left[\mathcal{W}\frac{1}{\mathcal{V}}\delta(\omega/E_\circ-\Omega_{\boldsymbol{q}})\frac{1}{\mathcal{V}}\mathcal{W}\right]_{(a,0),(b,0)} \\ &= \left(\frac{2\pi e^2 n_{\rm el}}{m^\star}\right)\frac{1}{E_\circ}\left[\mathcal{V}\delta(\omega/E_\circ-\Omega_{\boldsymbol{q}})\mathcal{V}\right]_{(a,0),(b,0)} \end{aligned} \tag{E7}$$

where we used that $n_{\rm el} = N_{\rm el}/A$.